



**UNIVERSIDADE ESTADUAL PAULISTA**  
**"JÚLIO DE MESQUITA FILHO"**  
Câmpus Bauru

Felipe Rodrigues Perche Mahlow

# **Aplicações de Aprendizado de Máquina Clássico e Quântico em Informação Quântica**

Bauru  
2025

Felipe Rodrigues Perche Mahlow

## **Aplicações de Aprendizado de Máquina Clássico e Quântico em Informação Quântica**

Exame de Defesa de Doutorado, apresentado ao Programa de Pós-Graduação em Ciência da Computação, da Faculdade de Ciências da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de Bauru.

Financiadora: CAPES - Proc. 88887.607339/2021-00

Orientador: Prof. Dr. Felipe Fernandes Fanchini

Coorientador: Prof. Dr. Daniel Zini Rossatto


M214a Mahlow, Felipe Rodrigues Perche  
Aplicações de Aprendizado de Máquina Clássico e Quântico em  
Informação Quântica / Felipe Rodrigues Perche Mahlow. -- Bauru,  
2025  
156 p.

Tese (doutorado) - Universidade Estadual Paulista (UNESP),  
Faculdade de Ciências, Bauru  
Orientador: Felipe Fernandes Fanchini  
Coorientador: Daniel Zini Rossato

1. Aprendizado de Máquinas. 2. Informação Quântica. 3.  
Inteligência Artificial. 4. Computação Quântica. I. Título.

**ATA DA DEFESA PÚBLICA DA TESE DE DOUTORADO DE FELIPE RODRIGUES PERCHE MAHLOW, DISCENTE DO PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO, DA FACULDADE DE CIÊNCIAS - CÂMPUS DE BAURU.**

Aos 25 dias do mês de fevereiro do ano de 2025, às 10h, por meio de Videoconferência, realizou-se a defesa de TESE DE DOUTORADO de FELIPE RODRIGUES PERCHE MAHLOW, intitulada **Aplicações de Aprendizado de Máquina Clássico e Quântico em Informação Quântica**. A Comissão Examinadora foi constituída pelos seguintes membros: Prof. Dr. FELIPE FERNANDES FANCHINI (Orientador(a) - Participação Virtual) do(a) Departamento de Física / Faculdade de Ciências Unesp Câmpus de Bauru, Prof. Dr. KELTON AUGUSTO PONTARA DA COSTA (Participação Virtual) do(a) FC / UNESP Bauru SP, Prof. Dr. MARCOS CÉSAR DE OLIVEIRA (Participação Virtual) do(a) Departamento de Física / Universidade Estadual de Campinas, Prof. Dr. REGINALDO DE JESUS NAPOLITANO (Participação Virtual) do(a) Instituto de Física de São Carlos / Universidade de São Paulo, Prof. Dr. LUIZ GUSTAVO ESMENARD ARRUDA (Participação Virtual) do(a) LARC - Laboratório de Arquitetura e Redes de Computadores do Departamento de Engenharia de Computação e Sistemas Digitais / Escola Politécnica da Universidade de São Paulo (PCS-EPUSP). Após a exposição pelo doutorando e arguição pelos membros da Comissão Examinadora que participaram do ato, de forma presencial e/ou virtual, o discente recebeu o conceito final APROVADO. Nada mais havendo, foi lavrada a presente ata, que após lida e aprovada, foi assinada pelo(a) Presidente(a) da Comissão Examinadora.

Documento assinado digitalmente  
 FELIPE FERNANDES FANCHINI  
Data: 14/03/2025 12:21:12-0300  
Verifique em <https://validar.iti.gov.br>

Prof. Dr. FELIPE FERNANDES FANCHINI

Felipe Rodrigues Perche Mahlow

## **Aplicações de Aprendizado de Máquina Clássico e Quântico em Informação Quântica**

Exame de Defesa de Doutorado, apresentado ao Programa de Pós-Graduação em Ciência da Computação, da Faculdade de Ciências da Universidade Estadual Paulista "Júlio de Mesquita Filho", Câmpus de Bauru.

Financiadora: CAPES - 88887.607339/2021-00

Comissão Examinadora

**Prof. Dr. Felipe Fernandes Fanchini**

Universidade Estadual Paulista (Unesp) - Câmpus de Bauru  
Orientador

**Prof. Dr. Kelton Augusto Pontara da Costa**

Universidade Estadual Paulista (Unesp) - Câmpus de Bauru

**Prof. Dr. Marcos César de Oliveira**

Universidade Estadual de Campinas (Unicamp)

**Prof. Dr. Reginaldo de Jesus Napolitano**

Universidade de São Paulo (USP) - Câmpus de São Carlos

**Dr. Luiz Gustavo Esmenard Arruda**

Universidade de São Paulo (USP) - Câmpus de São Paulo

Bauru  
18 de janeiro de 2025

*Dedico esta Tese a minha mãe - Augusta.*

*Espero que um dia possa lhe retribuir.*

# Agradecimentos

A princípio, agradeço à Universidade Estadual Paulista - Unesp, por ser a minha mãe acadêmica. Sou extremamente grato pelas oportunidades que aqui me foram proporcionadas e sinto um orgulho tremendo ao me autodenominar Unespiano! Agradeço a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo financiamento que me permitiu levar adiante esta pesquisa. Meus queridos professores do departamento de Física/ Ciência da Computação, que me foram exemplos de seres humanos e profissionais.

Agradeço à toda a minha grande família. Tenho cinco irmãos, pelos quais nutro muito amor, além de mais seis sobrinhos que são os meus xodós e os quais estou ansioso para assistir crescendo e se tornarem o que desejam em seus mais ambiciosos sonhos. Aos meus amigos, por todo o carinho e companheirismo, toda à diversão e rizadas. Agradeço em especial aos dois maiores alicerces da minha vida, duas mulheres que foram cruciais na minha jornada para chegar até aqui: A primeira é a minha mãe e a segunda a Raissa, minha companheira. Sem elas nada disso seria possível.

Sou grato à alguns grandes seres humanos que me ajudaram a chegar até aqui, como meus professores da educação básica (profs. Dino, Bárbara e Dugan) e alguns divulgadores científicos. Também em especial, agradeço ao meu orientador/amigo, o professor Felipe Fanchini, por todas as oportunidades as quais têm me proporcionado, além de todo o seu tempo e paciência empregados na minha trajetória. Por fim, agradeço a CAPES, PROEX-Unesp e Associação de Universidades do Grupo de Montevideu pelas oportunidades que me proporcionaram de trabalhar com grupos de pesquisa no exterior, onde eu tive o privilégio de conviver com seres humanos incríveis, que hoje tenho o orgulho de chamar de meus amigos.

Adicionando a essa imensa lista de agradecimentos, não poderia deixar de olhar para trás e refletir sobre a minha trajetória. Foi no terceiro ano do ensino médio, ainda morando no bairro de Pirituba em São Paulo, que decidi que queria ser cientista. Lembro-me vividamente do dia em que fui a uma feira de profissões na USP. Lá, conversando com um professor de física, ele me explicou que, para seguir essa carreira, eu precisaria fazer mestrado e doutorado. Foi naquele instante que plantei a semente do meu sonho e tracei esse objetivo para a minha vida. Hoje, dez anos depois, é surreal pensar que aquele plano traçado por aquele jovem cheio de aspirações está finalmente se concretizando. O tempo voa; parece que foi ontem que tudo começou. A vida realmente passa em um piscar de olhos, e é por isso que quero, mais do que nunca, aproveitar este momento para viver o agora, para sentir e valorizar cada pessoa e cada passo que me trouxeram até aqui. Obrigado.

O doutorado está se encerrando, mas a ciência está só começando!

*"É necessário sempre acreditar que o sonho é possível  
Que o céu é o limite e você, truta, é imbatível  
Que o tempo ruim vai passar, é só uma fase  
Que o sofrimento alimenta mais a sua coragem  
Que a sua família precisa de você  
Lado a lado se ganhar, pra te apoiar se perder"  
(Racionais MCs. - A Vida é Desafio)*

# Resumo

Esta tese explora as interseções entre Aprendizado de Máquina e Informação Quântica, investigando como ferramentas de inteligência artificial podem ser aplicadas para resolver problemas fundamentais em sistemas quânticos complexos. O trabalho foca em dois desafios principais: a classificação de fases quânticas e a predição de sincronização quântica. Na classificação de fases quânticas, apresentamos três abordagens complementares. A primeira explora a independência de modelo, demonstrando como é possível treinar um classificador em dados provenientes de um modelo de Hamiltoniano e utilizá-lo para classificar fases em um modelo distinto. A segunda abordagem emprega algoritmos de Aprendizado de Máquina Quântico (QML) com seleção de atributos baseada na técnica SHAP (Shapley Additive Explanations), otimizando a escolha de features relevantes para melhorar a precisão dos classificadores. Por fim, introduzimos o uso de redes generativas adversárias (GANs) para criar dados sintéticos, permitindo ampliar e diversificar os conjuntos de dados e, assim, aumentar a eficiência e a precisão das classificações. No segundo desafio, a predição de sincronização quântica, aplicamos um algoritmo de aprendizado supervisionado para prever a ocorrência de sincronização ou antissincronização em sistemas quânticos. Baseando-nos em dados iniciais da evolução temporal do sistema, demonstramos como identificar padrões que indicam a dinâmica futura desses fenômenos com alta confiabilidade. Esta tese contribui apresentando soluções inovadoras e um conjunto abrangente de ferramentas que conectam inteligência artificial à compreensão de sistemas quânticos. Os resultados obtidos oferecem insights valiosos para a compreensão e manipulação desses sistemas.

**Palavras-chave:** Informação Quântica. Aprendizado de Máquina Quântico. Fases Quânticas. Sincronização Quântica.

# Abstract

This thesis explores the intersections between Machine Learning and Quantum Information, investigating how artificial intelligence tools can be applied to solve fundamental problems in complex quantum systems. The work focuses on two main challenges: quantum phase classification and quantum synchronization prediction. In quantum phase classification, we present three complementary approaches. The first explores model independence, demonstrating how a classifier can be trained on data from one Hamiltonian model and used to classify phases in a distinct model. The second approach employs Quantum Machine Learning (QML) algorithms with feature selection based on the SHAP (Shapley Additive Explanations) technique, optimizing the choice of relevant features to improve classifier accuracy. Finally, we introduce the use of Generative Adversarial Networks (GANs) to create synthetic data, enabling the expansion and diversification of datasets and, consequently, increasing the efficiency and accuracy of classifications. For the second challenge, the prediction of quantum synchronization, we apply a supervised learning algorithm to predict the occurrence of synchronization or anti-synchronization in quantum systems. Using initial data from the system's temporal evolution, we demonstrate how to identify patterns that indicate the future dynamics of these phenomena with high reliability. This thesis contributes by presenting innovative solutions and a comprehensive set of tools that connect artificial intelligence to the understanding of quantum systems. The results provide valuable insights into the comprehension and manipulation of these systems.

**Keywords:** Quantum Information. Quantum Machine Learning. Quantum Phases. Quantum Synchronization.

# Lista de ilustrações

Figura 1 – Ilustração do funcionamento de uma Support Vector Machine (SVM). Autoria própria. Inspirado em (ALCALDE, 2017). . . . .	43
Figura 2 – Estrutura básica de uma Rede Adversarial Generativa (GAN). O gerador cria dados sintéticos, que são avaliados pelo discriminador. O feedback do discriminador melhora o desempenho do gerador em iterações subsequentes. . . . .	46
Figura 3 – Representação gráfica da esfera de Bloch . . . . .	58
Figura 4 – Áreas de Estudo dentro de Aprendizado de Máquina Quântico. Do autor, com base em (SCHULD; PETRUCCIONE, 2018) . . . . .	71
Figura 5 – Tentativas dos algoritmos de classificar o diagrama de fases de $\mathcal{H}_1$ usando dados fragmentados do próprio modelo. . . . .	88
Figura 6 – Tentativas dos algoritmos de classificar o diagrama de fases de $\mathcal{H}_2$ usando dados fragmentados do próprio modelo. . . . .	89
Figura 7 – Tentativas dos algoritmos de classificar o diagrama de fases de $\mathcal{H}_3$ usando dados fragmentados do próprio modelo. . . . .	90
Figura 8 – Estudo da relação entre a quantidade de dados utilizados no conjunto de treinamento versus a precisão obtida pelos quatro classificadores analisados. <b>a)</b> Refere-se aos resultados para $\mathcal{H}_2$ , <b>b)</b> é a precisão obtida na análise para $\mathcal{H}_2$ e <b>c)</b> apresenta essa relação para $\mathcal{H}_3$ . . . . .	90
Figura 9 – Componente PCA 1 vs. Componente PCA 2, a cor representa a fase do ponto de dados correspondente. E as figuras (a) e (e) referem-se ao sistema descrito por $\mathcal{H}_1$ , (b) e (f) referem-se a $\mathcal{H}_2$ , (c) e (g) referem-se a $\mathcal{H}_3$ , (d) e (h) são os dados dos três sistemas. As figuras (a), (b), (c) e (d) são os dados brutos, e (e), (f), (g), (h) são os dados após a realização da transformação. . . . .	92
Figura 10 – A figura (a) mostra a previsão de fases do modelo $\mathcal{H}_1$ usando o aprendizado dos modelos $\mathcal{H}_2$ e $\mathcal{H}_3$ . A figura (b) mostra a previsão de fases do modelo $\mathcal{H}_2$ usando o aprendizado dos modelos $\mathcal{H}_1$ e $\mathcal{H}_3$ , e a figura (c) mostra a previsão de fases do modelo $\mathcal{H}_3$ usando o aprendizado dos modelos $\mathcal{H}_1$ e $\mathcal{H}_2$ . Os pontos brancos nos painéis (a) e (b) e as linhas verticais no painel (c) são as fronteiras teóricas das fases fornecidas na literatura, para os modelos $\mathcal{H}_1$ (CHEN; HIDA; SANCTUARY, 2003), $\mathcal{H}_2$ (KITAZAWA; NOMURA; OKAMOTO, 1996) e $\mathcal{H}_3$ (LÄUCHLI; SCHMID; TREBST, 2006), respectivamente. O nome em preto representa a fase real naquele local do modelo, e a cor representa a fase prevista. As fases com transparência (LD em (a) e trimer em (c)) são aquelas que não estavam presentes na base de treinamento do algoritmo. . . . .	93

Figura 11 – Precisão obtida em função do valor de $k$ ao classificar $\mathcal{H}_1$ usando dados de $\mathcal{H}_2$ e $\mathcal{H}_3$ para treinamento, ao classificar $\mathcal{H}_2$ usando dados de $\mathcal{H}_1$ e $\mathcal{H}_3$ para treinamento e ao classificar $\mathcal{H}_3$ usando dados de $\mathcal{H}_1$ e $\mathcal{H}_2$ para treinamento. a) Variação da precisão para $k$ até $k = 3000$ , enquanto b) é um quadro na região entre $0 < k < 100$ . . . . .	94
Figura 12 – Classificação do diagrama de fases XXZ usando modelos de aprendizado de máquina supervisionado com dados mínimos. Os resultados ilustram a precisão da previsão de fase alcançada pelos modelos SVM, Random Forest, Perceptron Multicamadas e $k$ -Nearest Neighbors, utilizando apenas cinco pontos de dados por fase. . . . .	102
Figura 13 – Classificação do diagrama de fases Bilinear usando modelos de aprendizado de máquina supervisionado com dados mínimos. A figura mostra a precisão da previsão de fase para os modelos SVM, Random Forest, Perceptron Multicamadas e $k$ -Nearest Neighbors, treinados com cinco pontos de dados por fase. . . . .	103
Figura 14 – Desempenho dos algoritmos SVM, Random Forest, Perceptron Multicamadas e $k$ -Nearest Neighbors em termos de acurácia na classificação de fases do modelo XXZ, em função do número de amostras de treinamento por fase. . . . .	103
Figura 15 – Desempenho dos algoritmos SVM, Random Forest, Perceptron Multicamadas e $k$ -Nearest Neighbors na classificação de fases do modelo Bilinear em função do número de amostras de treinamento por fase. . . . .	104
Figura 16 – Impacto da adição de dados sintéticos gerados por uma CTGAN no desempenho do algoritmo KNN para os modelos XXZ e Bilinear. . . . .	105
Figura 17 – Ranking de importância de características pelo SHAP para um modelo ANNNI de 8 sítios utilizando o Support Vector Classifier. O gráfico mostra a importância relativa das diferentes características na classificação das fases do modelo. . . . .	111
Figura 18 – Ranking de importância de características pelo SHAP para um modelo ANNNI de 12 sítios. O gráfico revela as características mais significativas para a classificação de fases neste sistema maior. . . . .	112
Figura 19 – Diagrama de fases do modelo ANNNI, delineado por meio da classificação utilizando o VQC (Topo) e o QSVM (Base). Apresentado com cadeias de 8 sítios (esquerda) e 12 sítios (direita). Os marcadores denotam as previsões de fase de acordo com o modelo, enquanto as linhas representam as transições teóricas. . . . .	114
Figura 20 – Precisão em função do número de características para QSVM e VQC. . . . .	114

- Figura 21 – (a) Primeiro modelo de colisão simulando dissipação local para o qubit do sistema  $s_2$ : (i) O qubit do sistema  $s_2$  interage com o qubit do ambiente  $e_n$ . (ii) Os qubits do sistema  $s_1$  e  $s_2$  interagem diretamente entre si. (iii) Os qubits do sistema  $s_1$  e  $s_2$  evoluem livremente. (iv) O qubit do ambiente que interagiu com o qubit do sistema  $s_2$  na etapa (i) é traçado, e o mesmo ciclo é repetido com o próximo qubit do ambiente  $e_{n+1}$ . (b) Segundo modelo de colisão simulando dissipação global para os qubits do sistema  $s_1$  e  $s_2$ : (i) O qubit do sistema  $s_1$  interage com o qubit do ambiente  $e_n$ . (ii) O qubit do sistema  $s_2$  interage com o mesmo qubit do ambiente  $e_n$ . (iii) Os qubits do sistema  $s_1$  e  $s_2$  interagem diretamente. (iv) Os qubits do sistema  $s_1$  e  $s_2$  evoluem livremente, e o qubit do ambiente que interagiu com os qubits do sistema na etapa (i-ii) é traçado para continuar o ciclo com o próximo qubit do ambiente  $e_{n+1}$ . . . . . 119
- Figura 22 – O primeiro modelo de colisão: (a) O coeficiente de Pearson  $C_{12}$ , calculado para os últimos 100 valores esperados dos qubits do sistema  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , em função da intensidade de interação direta  $\lambda$  entre eles e do *detuning* entre suas autoenergias  $\omega_1/\omega_2$ , para a intensidade de acoplamento sistema-ambiente  $J = 0.1$ . (b) Valores previstos de  $C_{12}$  fornecidos pelo regressor  $k$ -*NN* usando os primeiros 100 valores esperados do par de qubits em função de  $\lambda$  e  $\omega_1/\omega_2$  para  $J = 0.1$ . . . . . 125
- Figura 23 – Erro médio absoluto (MAE) para o coeficiente de Pearson  $C_{12}$  no primeiro modelo de colisão em função do erro percentual que é adicionado aos valores esperados calculados dos qubits  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ . Quatro linhas diferentes na legenda correspondem aos casos de fornecimento ao regressor  $k$ -*NN* de 5, 10, 50 e 100 pares de valores esperados dos qubits para previsão. . . . . 126
- Figura 24 – O segundo modelo de colisão: (a) O coeficiente de Pearson  $C_{12}$ , calculado para os últimos 100 valores esperados dos qubits do sistema  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , em função da intensidade de interação direta  $\lambda$  entre eles e da *detuning* entre suas autoenergias  $\omega_1/\omega_2$ . (b) Valores previstos de  $C_{12}$  fornecidos pelo regressor  $k$ -*NN* usando os primeiros 100 valores esperados dos qubits em função de  $\lambda$  e  $\omega_1/\omega_2$ . . . . . 127
- Figura 25 – Erro médio absoluto (MAE) para o coeficiente de Pearson  $C_{12}$  no segundo modelo de colisão em função do erro percentual adicionado aos valores esperados calculados  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ . Quatro linhas diferentes na legenda correspondem aos casos em que fornecemos ao regressor  $k$ -*NN* de 5, 10, 50 e 100 pares de valores esperados dos qubits para previsão. . . . . 128

Figura 26 – O modelo de equação mestra: (a) O coeficiente de Pearson $C_{12}$ , calculado para os últimos 100 valores esperados dos qubits $\langle \sigma_{s_1}^x \rangle$ e $\langle \sigma_{s_2}^x \rangle$ , em função da intensidade de interação de troca $f_{12}$ entre eles e da <i>detuning</i> entre suas autoenergias $\omega_1/\omega_2$ . (b) Valores previstos de $C_{12}$ produzidos pelo regressor $k$ - <i>NN</i> usando os primeiros 100 valores esperados do par de qubits em termos de $f_{12}$ e $\omega_1/\omega_2$ . . . . .	129
Figura 27 – Erro médio absoluto (MAE) para o coeficiente de Pearson $C_{12}$ no modelo de equação mestra em função do erro percentual adicionado aos valores esperados calculados dos qubits $\langle \sigma_{s_1}^x \rangle$ e $\langle \sigma_{s_2}^x \rangle$ . Quatro linhas diferentes na legenda correspondem aos casos de fornecimento ao regressor $k$ - <i>NN</i> de 5, 10, 50 e 100 pares de valores esperados dos qubits para previsão. . . . .	131
Figura 28 – Erro médio absoluto (MAE) para o coeficiente de Pearson $C_{12}$ em função da probabilidade de erro no estado inicial fornecido ao regressor $k$ - <i>NN</i> para o modelo de equação mestra (ME), modelo de colisão local (LCM) e modelo de colisão global (GCM) usando 100 valores esperados iniciais. . . . .	132
Figura 29 – Erro médio absoluto (MAE) para o coeficiente de Pearson $C_{12}$ em função da temperatura do ambiente para o modelo de equação mestra (ME), modelo de colisão local (LCM) e modelo de colisão global (GCM), considerando 100 valores esperados iniciais dos qubits. . . . .	133
Figura 30 – Dinâmica dos valores esperados locais de spin dos dois qubits para o modelo de colisão local (LCM) (superior), modelo de colisão global (GCM) (meio) e modelo de equação mestra (ME) (inferior), que exibem tanto sua dinâmica pré-sincronização quanto o surgimento da transição dinâmica para o comportamento (anti-)síncrono. . . . .	135

# Lista de tabelas

Tabela 1 – Fases contidas nos diagramas correspondentes aos três Hamiltonianos analisados. Os círculos preenchidos indicam que a fase quântica está presente no modelo. Os números sobrescritos nos círculos representam a quantidade de dados calculados para cada fase de cada modelo. . . . .	85
Tabela 2 – Precisão obtida pelo algoritmo $k$ -NN na classificação das fases quânticas quando há variação no tamanho da cadeia de spins. . . . .	95
Tabela 3 – Desempenho na classificação dos algoritmos QSVM e VQC em sistemas de 8 e 12 sítios, utilizando as 5 características mais impactantes determinadas pelo algoritmo SHAP. . . . .	113

# Lista de abreviaturas e siglas

AM	Aprendizado de Máquina
APS	<i>American Physical Society</i>
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
DT	<i>Decision Tree</i>
IA	Inteligência Artificial
k-NN	<i>k-Nearest Neighbors</i>
MAE	<i>Mean Absolute Error</i>
ML	<i>Machine Learning</i>
MLP	<i>Multi-Layer Perceptron</i>
MSE	<i>Mean Squared Error</i>
NMSE	<i>Normalized Mean Squared Error</i>
PCA	<i>Principal Component Analysis</i>
QIS	<i>Quantum Information Science</i>
QML	<i>Quantum Machine Learning</i>
RNAs	Redes Neurais Artificiais
RNQ	Redes Neurais Quânticas
SHAP	<i>Shapley Additive Explanations</i>
SSL	<i>Self-Supervised Learning</i>
SVM	<i>Support Vector Machine</i>

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>18</b>
<b>1.1</b>	<b>Hipótese</b>	<b>20</b>
<b>1.2</b>	<b>Objetivos</b>	<b>21</b>
<b>1.3</b>	<b>Estrutura da Tese</b>	<b>21</b>
<b>1.4</b>	<b>Declaração de Autoria</b>	<b>22</b>
<b>2</b>	<b>MACHINE LEARNING</b>	<b>23</b>
<b>2.1</b>	<b>Contextualização Histórica</b>	<b>24</b>
<b>2.2</b>	<b>Conceitos Fundamentais</b>	<b>27</b>
2.2.1	Supervisionado X Não Supervisionado	27
2.2.2	Regressão e Classificação	27
2.2.3	Funções Custo	29
2.2.4	<i>Overfitting</i> e <i>Underfitting</i>	30
2.2.5	Métricas de Avaliação e Validação	31
2.2.5.1	Métricas na Classificação	32
2.2.5.2	Métricas na Regressão	33
2.2.5.3	Validação Cruzada	34
2.2.5.4	Desafios na Avaliação	34
2.2.6	Hiperparâmetros e Otimização	35
2.2.7	Pré-Processamento de Dados	36
<b>2.3</b>	<b>Exemplos de Algoritmos</b>	<b>37</b>
2.3.1	<i>k</i> -Nearest Neighbors	37
2.3.2	Perceptrons Multi Camadas	39
2.3.3	Árvores de Decisão	41
2.3.4	Máquinas de Vetores de Suporte	43
2.3.5	Análise do Componente Principal	45
<b>2.4</b>	<b>Redes Adversariais Generativas (GANs)</b>	<b>46</b>
2.4.1	Conditional Tabular GANs (CTGANs)	47
<b>2.5</b>	<b>Explainable AI</b>	<b>47</b>
2.5.1	<i>Shapley Additive Explanations</i>	48
<b>2.6</b>	<b>Tendências Futuras</b>	<b>50</b>
<b>3</b>	<b>INFORMAÇÃO QUÂNTICA</b>	<b>52</b>
<b>3.1</b>	<b>Uma Breve Introdução Histórica</b>	<b>52</b>
<b>3.2</b>	<b>Conceitos Fundamentais</b>	<b>54</b>
3.2.1	Notação de Dirac	55

3.2.2	Superposição . . . . .	55
3.2.3	O Qubit . . . . .	56
3.2.4	Esfera de Bloch e bases para um Qubit . . . . .	57
3.2.5	Operações em um Qubit . . . . .	59
3.2.6	Operações com dois Qubits . . . . .	60
3.2.6.1	Produto Tensorial . . . . .	60
3.2.6.2	Portas de Dois Qubits . . . . .	61
3.2.6.3	Emaranhamento . . . . .	61
3.2.7	Operador Densidade . . . . .	62
3.2.8	Medidas e Observáveis . . . . .	63
3.2.8.1	População e Coerência . . . . .	65
3.2.8.2	Operação Traço Parcial . . . . .	65
3.2.9	Hamiltoniano . . . . .	66
<b>3.3</b>	<b>Sistemas Quânticos Abertos . . . . .</b>	<b>67</b>
3.3.1	Evolução Temporal . . . . .	67
3.3.2	Equação Mestra de Lindblad . . . . .	68
<b>3.4</b>	<b>Aplicações e Perspectivas Futuras . . . . .</b>	<b>69</b>
<b>4</b>	<b>QUANTUM MACHINE LEARNING . . . . .</b>	<b>70</b>
<b>4.1</b>	<b>Contextualização Histórica . . . . .</b>	<b>71</b>
<b>4.2</b>	<b><i>Quantum Speedup</i> . . . . .</b>	<b>72</b>
<b>4.3</b>	<b>Codificação da Informação . . . . .</b>	<b>73</b>
4.3.1	<i>Basis Encoding</i> . . . . .	73
4.3.2	<i>Amplitude Encoding</i> . . . . .	74
4.3.3	Pauli Feature Map . . . . .	75
<b>4.4</b>	<b>Algoritmos Quânticos Variacionais . . . . .</b>	<b>76</b>
<b>4.5</b>	<b>Redes Neurais Quânticas . . . . .</b>	<b>77</b>
<b>4.6</b>	<b>Métodos de Kernel . . . . .</b>	<b>77</b>
4.6.1	Máquina de Vetores de Suporte Quântica . . . . .	79
<b>4.7</b>	<b>Perspectivas Futuras . . . . .</b>	<b>79</b>
<b>5</b>	<b>CLASSIFICANDO FASES QUÂNTICAS . . . . .</b>	<b>81</b>
<b>5.1</b>	<b>Classificador Independente de Modelo . . . . .</b>	<b>81</b>
5.1.1	Transições de Fase . . . . .	83
5.1.1.1	Transições de Fase Clássicas . . . . .	83
5.1.1.2	Transições de Fase Quânticas . . . . .	84
5.1.2	Os Três Modelos . . . . .	84
5.1.3	Resultados . . . . .	86
5.1.3.1	Os Bancos de Dados . . . . .	86
5.1.3.2	Método de Generalização . . . . .	87

5.1.3.3	Independência de Modelo . . . . .	91
5.1.4	Conclusão . . . . .	95
<b>5.2</b>	<b>Classificando Fases Quânticas com Dados Escassos . . . . .</b>	<b>97</b>
5.2.1	Modelos Físicos . . . . .	98
5.2.1.1	Cadeia XXZ . . . . .	98
5.2.1.2	Cadeia Bilinear Biquadrática . . . . .	99
5.2.2	Estrutura dos Dados . . . . .	100
5.2.3	Resultados . . . . .	101
5.2.4	Conclusão . . . . .	105
<b>5.3</b>	<b>Classificando Fases Quânticas com QML . . . . .</b>	<b>107</b>
5.3.1	O Modelo ANNNI . . . . .	107
5.3.2	Métodos . . . . .	109
5.3.2.1	Seleção de Características com SHAP . . . . .	109
5.3.3	Resultados . . . . .	109
5.3.3.1	Seleção de Características . . . . .	110
5.3.3.2	Resultados de Classificação . . . . .	111
5.3.4	Conclusão . . . . .	113
<b>6</b>	<b>ESTIMANDO A SINCRONIZAÇÃO QUÂNTICA . . . . .</b>	<b>116</b>
<b>6.1</b>	<b>Modelos Físicos . . . . .</b>	<b>118</b>
6.1.1	Modelos de Colisão . . . . .	118
6.1.2	Modelo de Equação Mestra . . . . .	120
<b>6.2</b>	<b>Estrutura dos dados . . . . .</b>	<b>121</b>
<b>6.3</b>	<b>Resultados . . . . .</b>	<b>123</b>
6.3.1	Modelos de Colisão . . . . .	123
6.3.2	Modelo de Equação Mestra . . . . .	129
6.3.3	Possíveis Mecanismos de Erro Físico . . . . .	131
<b>6.4</b>	<b>Conclusão . . . . .</b>	<b>133</b>
<b>6.5</b>	<b>Apêndice - Comportamento em tempos iniciais e surgimento da sincronização . . . . .</b>	<b>134</b>
<b>7</b>	<b>CONCLUSÃO DA TESE . . . . .</b>	<b>136</b>
<b>7.1</b>	<b>Direcionamentos para Trabalhos Futuros . . . . .</b>	<b>137</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>138</b>

# 1 Introdução

Inteligência Artificial (IA) é um ramo da Ciência da Computação dedicado a entender entidades inteligentes. Ao contrário de ciências como a filosofia e a psicologia, a IA se interessa pela construção e modelagem dessas inteligências enquanto as compreende, permitindo que contribuam para as mais diversas aplicações na sociedade moderna (RUSSELL; NORVIG, 1995). Embora ainda estejamos nos primórdios do desenvolvimento da IA, ao entrar na terceira década do século XXI, já é possível observar resultados importantes, como o reconhecimento de voz e imagem, sistemas de recomendação, diagnósticos médicos e automação de veículos. Esses avanços têm transformado significativamente o cotidiano humano (GOODFELLOW; BENGIO; COURVILLE, 2016).

Em ciência, algoritmos de IA têm alcançado grande sucesso, sendo amplamente utilizados em diversas áreas, como a ciência da computação, medicina, biologia e estudos sociais. A física, com seu foco nas leis fundamentais que governam o universo, representa um campo de extrema complexidade que desafia nossa compreensão. Muitos dos padrões que regem os fenômenos quânticos ainda escapam à nossa intuição. Esse cenário abre possibilidades para que algoritmos de IA, com sua habilidade de identificar padrões em dados complexos, possam contribuir para a compreensão de sistemas de forma inovadora. Isso de fato já vem sendo feito em diversas áreas, como cosmologia, física de partículas, computação quântica e ciência de materiais (CARLEO et al., 2019). A atual relevância da IA em ciência foi destacada pelos Prêmios Nobel de 2024. O Prêmio Nobel de Física foi concedido a John J. Hopfield e Geoffrey Hinton por suas descobertas fundamentais que possibilitaram o aprendizado de máquina com redes neurais artificiais. No campo da química, David Baker, Demis Hassabis e John Jumper foram laureados por suas contribuições no design computacional de proteínas e na predição de suas estruturas utilizando IA.

Contudo, dois fatores principais representam barreiras ao avanço da área. O primeiro é o enorme volume de dados a ser processado, amplificado pela crescente adoção de dispositivos conectados à Internet das Coisas (IoT), que coletam dados em tempo real de forma massiva. O segundo é o fim da Lei de Moore, que previa o aumento exponencial na quantidade de transistores em chips e, conseqüentemente, no poder de processamento. Com os transistores modernos se aproximando de limites físicos na escala nanométrica, novas soluções estão sendo exploradas, como a adoção de arquiteturas heterogêneas e o desenvolvimento de chips especializados em IA para superar as limitações dos sistemas clássicos. Uma das principais apostas para enfrentar esses desafios é a computação quântica, que utiliza fenômenos da mecânica quântica, como o emaranhamento e a superposição, para ir além das limitações dos computadores clássicos (WOSSNIG, 2021).

Desde o surgimento da teoria quântica no início do século XX, avanços significativos permitiram a explicação de fenômenos fundamentais e impulsionaram inúmeras inovações tecnológicas, como semicondutores, supercondutores e superfluidos. Nas décadas de 1970 e 1980, surgiram questionamentos sobre a natureza da informação quântica, levando ao nascimento da Ciência da Informação Quântica (QIS – Quantum Information Science, em inglês), que combina aspectos de física, computação e teoria da informação (NIELSEN; CHUANG, 2000). Nesse contexto, o estudo do processamento de informação utilizando os princípios da informação quântica, bem como a investigação das propriedades e dos sistemas quânticos, emerge como uma área de extrema relevância para o desenvolvimento científico e tecnológico. Além disso, essas pesquisas contribuem para o aprimoramento de tecnologias disruptivas, como computadores quânticos, sensores quânticos de alta precisão e comunicações seguras, que têm o potencial de transformar profundamente a sociedade moderna.

Tanto o aprendizado de máquina quanto a computação quântica representam avanços notáveis para a humanidade. Isso levou à emergência do campo de Aprendizado de Máquina Quântico (Quantum Machine Learning – QML), que combina princípios da mecânica quântica com técnicas de aprendizado de máquina clássico (BIAMONTE et al., 2017). Essa área emergente visa aproveitar fenômenos quânticos, como a superposição e o emaranhamento, para resolver problemas computacionais complexos de forma mais eficiente, abrangendo desafios como otimização, simulação de sistemas moleculares e análise de grandes volumes de dados (SCHULD; PETRUCCIONE, 2018).

Nesta tese, iremos explorar tanto a utilização de algoritmos de aprendizagem de máquina para detectar padrões e fornecer *insights* sobre o comportamento de sistemas quânticos complexos, quanto nos aprofundar nessa intersecção entre aprendizado de máquina (ML) e computação quântica, utilizando algoritmos de aprendizado de máquina quântico (QML) para realizar a classificação de fases quânticas. O estudo das fases quânticas e do fenômeno de sincronização quântica permite a exploração de aspectos fundamentais da dinâmica de sistemas quânticos, ao mesmo tempo em que fornece uma plataforma para avançar na compreensão e desenvolvimento de técnicas baseadas em IA e QML.

Durante o desenvolvimento deste trabalho, utilizaremos e nos aprofundaremos em uma ampla gama de algoritmos, tanto clássicos quanto quânticos, cobrindo diferentes abordagens de aprendizado de máquina e ferramentas de análise de dados. Entre os métodos clássicos, destacam-se os algoritmos supervisionados para identificação de padrões e classificação, e técnicas de explicação como o SHAP (Shapley Additive Explanations), que nos permitirão avaliar a relevância de diferentes features no contexto dos sistemas quânticos. Também exploraremos redes generativas adversárias (GANs) para a criação de dados sintéticos, visando melhorar a eficiência e a precisão das classificações.

No contexto de aprendizado de máquina quântico, utilizaremos algoritmos como Quantum Support Vector Machines e Redes Neurais Quânticas. Esses algoritmos não apenas nos

permitirão abordar o problema de classificação de fases quânticas, mas também fornecerão novos caminhos para investigar como a interação entre QML e sistemas físicos pode contribuir para avanços teóricos e práticos. Ao combinar esses métodos, esperamos desenvolver uma abordagem robusta e integrada para explorar as fronteiras entre a informação quântica e a inteligência artificial, contribuindo para o progresso dessa área emergente.

É importante destacar que, embora os exemplos explorados ao longo desta tese sejam provenientes da física, o foco central deste trabalho não reside na resolução de problemas físicos específicos, mas sim no desenvolvimento, aplicação e análise crítica de técnicas de aprendizado de máquina clássico e quântico aplicadas no contexto de informação quântica. A física surge aqui como um domínio rico e desafiador que fornece contextos relevantes e complexos nos quais essas técnicas podem ser testadas e avaliadas, mas não é, por si só, o objeto principal de investigação. Nosso interesse se concentra, portanto, nos métodos, nos algoritmos e nas estratégias de modelagem e interpretação dos dados. E como eles podem ser úteis no contexto estudado.

Além disso, esta tese deve ser compreendida como uma coleção de investigações interconectadas (em quatro trabalhos distintos), cada uma dedicada a explorar diferentes aspectos do uso de inteligência artificial clássica e quântica na análise de sistemas quânticos complexos. Ao invés de buscar uma única narrativa linear com um único problema central, reunimos contribuições que se complementam e reforçam mutuamente, com ênfase contínua na avaliação do desempenho, interpretabilidade e potencial das técnicas aplicadas. Essa abordagem permite não apenas profundidade técnica, mas também flexibilidade na análise de diferentes cenários, contribuindo assim com os campos de pesquisa estudados.

## 1.1 Hipótese

O aprendizado de máquina, tanto em sua forma clássica quanto em sua vertente quântica, pode ser aplicado ao estudo de sistemas quânticos complexos com o objetivo de identificar padrões, inferir estruturas subjacentes e realizar previsões relevantes. Essas técnicas permitem explorar grandes volumes de dados e dinâmicas de alta dimensionalidade que, em muitos casos, escapam da análise analítica tradicional ou exigiriam simulações computacionais extremamente custosas. Em especial, métodos de aprendizado de máquina são capazes de lidar com a não linearidade e a incerteza inerente aos sistemas quânticos, oferecendo caminhos alternativos para sua caracterização e compreensão. Dessa forma, o uso dessas técnicas não apenas complementa abordagens tradicionais, mas também abre novas perspectivas para o avanço do conhecimento em ciência e tecnologia quânticas.

## 1.2 Objetivos

**Objetivo Central:** Demonstrar através de aplicações práticas, a possibilidade de reconhecer padrões e fazer inferências relevantes para a Informação/ Computação Quântica, utilizando algoritmos de aprendizagem de máquina clássicos e quânticos.

Trabalhar este objetivo central permite que seja desenvolvido um denso ferramental em aprendizagem de máquina, tanto de uma perspectiva teórica quanto de sua aplicação prática, assim como em teoria de Informação Quântica.

### **Objetivos Secundários:**

- Demonstrar a possibilidade de classificar fases quânticas utilizando algoritmos de aprendizagem de máquina clássicos e quânticos;
- Demonstrar a possibilidade de classificar fases quânticas mesmo no caso de dados escassos e como a geração de dados sintéticos pode contribuir com o processo;
- Indicar a possibilidade de prever comportamentos de sincronização/ antisincronização/ ausência destes em sistemas quânticos.

Os objetivos secundários tratam-se das aplicações a serem estudadas e serão extensamente abordadas em capítulos individuais ao decorrer da tese, como indicado na próxima seção.

## 1.3 Estrutura da Tese

A seguir, será apresentado um breve resumo sobre os capítulos que serão apresentados nesta tese.

A tese se inicia com três capítulos voltados para a contextualização e fundamentação teórica. No capítulo 2 (Machine Learning), nos aprofundamos no domínio do aprendizado de máquina, estabelecendo os conceitos fundamentais e as preliminares matemáticas necessárias para a tese. O Capítulo 3 (Informação Quântica) fornece uma visão geral da Ciência da Informação Quântica, traçando suas raízes históricas e princípios fundamentais relevantes nas problemáticas estudadas. Em seguida, o capítulo 4 (*Quantum Machine Learning*) irá fornecer o ferramental básico sobre o estado da arte do aprendizado de máquina quântico, uma disciplina extremamente nova que trabalha nos campos das duas anteriores. Estas seções servem de plano de fundo para nossa exploração das interseções entre QIS e Machine Learning, e ajudarão o leitor a compreender conceitos fundamentais das problemáticas a serem atacadas nas aplicações.

A partir do quarto capítulo, entramos no campo das aplicações, que são as contribuições desta tese para a literatura acadêmica. O Capítulo 5 (Classificando Fases Quânticas) possui três seções principais, sendo cada, uma diferente abordagem e um artigo acadêmico a parte,

onde abordaremos diferentes Hamiltonianos a serem estudados e diferentes técnicas para a classificação das fases. Em 5.1 demonstramos o conceito de independência de modelo, onde um classificador treinado com dados de dois modelos distintos é capaz de prever as fases quânticas de um terceiro. Em 5.2, mostramos a possibilidade de realizar tais predições mesmo em situações em que os dados são extremamente escassos, indicando como melhorar a predição desses modelos utilizando dados sintéticos. Por fim em 5.3 mostramos a possibilidade de realizar a classificação das fases utilizando algoritmos quânticos de aprendizagem de máquina e como realizar a seleção de features para aumentar a precisão dos mesmos e reduzir o custo computacional usando o algoritmo SHAP. Os resultados, incluindo especificidades do banco de dados e métodos para todos os casos são apresentados em detalhe. O Capítulo 6 (Estimando a Sincronização Quântica) centra-se na predição da sincronização quântica, incluindo modelos de colisão e de equação mestra. Descrevemos as estruturas de dados empregadas e apresentamos os resultados dos nossos esforços, concluindo com insights e orientações para pesquisas futuras.

Por fim, no capítulo 7 (Conclusão) resumimos nossas descobertas, destacando os principais resultados da tese. Concluímos com reflexões sobre a pesquisa realizada e propomos possíveis caminhos para futuras investigações.

## 1.4 Declaração de Autoria

Esta tese é baseada nos seguintes artigos de pesquisa:

- Mahlow, F., Luiz, F.S., Malvezzi, A.L. et al. Model-independent quantum phases classifier. *Sci Rep* 13, 14411 (2023). <https://doi.org/10.1038/s41598-023-33301-0>
- Mahlow, F., Çakmak, B., Karpat, G., Yalçinkaya, İ., Fanchini, F. F. (2024). Predicting the onset of quantum synchronization using machine learning. *Physical Review A*, 109(5), 052411. (**Editors' Suggestion**)
- Santos, G., Mahlow, F., Prado, P., Pexe, G., Rattighieri, L., Fanchini, F. Quantum Phases Classification Using Quantum Machine Learning with SHAP-Driven Feature Selection. 2025. <https://doi.org/10.48550/arXiv.2504.10673>.

## 2 Machine Learning

Há centenas de anos a humanidade se pergunta se é possível programar uma máquina para desenvolver pensamento próprio, havendo diversos exemplos de mitos de criação de vida artificial espalhadas ao redor do mundo, tais como Galatea e Pandora na mitologia grega (GOODFELLOW; BENGIO; COURVILLE, 2016). Atualmente, utilizamos inúmeros dispositivos e algoritmos que se utilizam de técnicas ou algoritmos de inteligência artificial, porém, estes são constituídos de uma forma muito diferente do que inicialmente foram idealizados. No começo do desenvolvimento da IA, programadores criaram aplicações utilizando *hard-coded knowledge*, ou seja, programavam algoritmos utilizando regras rígidas de comportamento, através de inúmeros *if (se)*, *else (senão)*, afim de ajustar o *input* (dados de entrada) a um *output* (dados de saída). Este tipo de abordagem pode ser eficaz em diversas situações, como por exemplo, para filtrar *spam* de e-mails utilizando determinadas palavras chave. Entretanto, estes métodos apresentavam grandes desvantagens, tais como a necessidade de um entendimento profundo de como o sistema em questão funciona e o fato de que cada decisão lógica dependeria de uma regra totalmente especificada, de forma que com uma pequena mudança no comportamento o código já poderia deixar de funcionar (MÜLLER; GUIDO, 2017).

O grande sucesso dos algoritmos de Inteligência Artificial, se dá pela sua capacidade de encontrar padrões sem a necessidade de programação explícita, isto é, o algoritmo é capaz de aprender por conta própria, sendo esta habilidade conhecida como Aprendizagem de Máquina (do Inglês *Machine Learning* - ML) (GOODFELLOW; BENGIO; COURVILLE, 2016). A forma pela qual um algoritmo é capaz de realizar tal feito é a extração de conhecimento através de dados. Machine Learning está em uma intersecção entre diversas áreas, tais como estatística, inteligência artificial, ciência da computação, análise preditiva e aprendizado estatístico. Esta revolução permitiu que tarefas consideradas não solucionadas até pouco tempo, como a detecção de rostos que até 2001 não possuía uma abordagem eficiente, pudesse hoje ser tratada como uma tecnologia a qual estamos habituados (MÜLLER; GUIDO, 2017). Para nós humanos, aprender algo diz respeito a experienciar situações e ser capaz de inferir a razão de um determinado comportamento ou mesmo realizar previsões com base em nossas experiências passadas. Já para uma máquina a aprendizagem vem através dos dados, de forma que um algoritmo tenta automatizar o processo de ganhar experiência, para que seja possível realizar uma previsão com base em suas conclusões (SCHULD; PETRUCCIONE, 2018).

Neste capítulo, apresentaremos uma visão ampla sobre os conceitos fundamentais que estruturam o aprendizado de máquina, começando com uma contextualização histórica que destaca os principais marcos e avanços que moldaram essa área (seção 2.1). Em seguida, abordaremos os conceitos centrais do aprendizado de máquina (seção 2.2), com destaque para as diferenças entre aprendizado supervisionado e não supervisionado (subseção 2.2.1), além

de discutir as tarefas de regressão e classificação (subseção 2.2.2). Exploraremos também a importância das funções custo na avaliação e otimização de modelos (subseção 2.2.3), assim como os desafios de *overfitting* (sobreajuste) e *underfitting* (subajuste) e seus impactos no desempenho dos modelos (subseção 2.2.4).

Para a avaliação dos modelos, discutiremos diversas métricas, com foco nas específicas para classificação (subseção 2.2.5.1) e regressão (subseção 2.2.5.2), além da importância da validação cruzada (subseção 2.2.5.3) e dos desafios na avaliação de modelos (subseção 2.2.5.4). Abordaremos a otimização de hiperparâmetros (seção 2.2.6) e a relevância do pré-processamento de dados na preparação dos conjuntos de dados para o treinamento dos modelos (seção 2.2.7). Em seguida, apresentaremos exemplos de algoritmos clássicos (seção 2.3), como *k*-Nearest Neighbors (subseção 2.3.1), Perceptrons Multi Camadas (subseção 2.3.2), Árvores de Decisão (subseção 2.3.3), Máquinas de Vetores de Suporte (subseção 2.3.4) e Análise do Componente Principal (subseção 2.3.5), destacando suas aplicações e características.

No final, exploraremos as Redes Adversariais Generativas (GANs) (seção 2.4), com foco nas Conditional Tabular GANs (CTGANs) (subseção 2.4.1), e apresentaremos os avanços em inteligência artificial explicável (Explainable AI) (seção 2.5), com ênfase nas *Shapley Additive Explanations* (subseção 2.5.1), que auxiliam na interpretabilidade dos modelos. Por fim, discutiremos as tendências futuras (seção 2.6), que prometem moldar o campo do aprendizado de máquina. Este capítulo servirá como base para os conceitos e técnicas que sustentam a metodologia aplicada nesta tese.

## 2.1 Contextualização Histórica

Nos últimos anos, o campo de Aprendizado de Máquina (AM) tem emergido como uma área de pesquisa amplamente reconhecida e tem despertado considerável interesse devido a uma série de avanços significativos. O interesse crescente pelo AM é, em parte, uma consequência direta dos avanços tecnológicos na área de hardware. Processadores de alto desempenho, em particular as GPUs, proporcionaram uma capacidade de processamento paralelo sem precedentes, possibilitando a aceleração de algoritmos de AM complexos, como redes neurais profundas. Essa capacidade computacional expandida viabilizou o treinamento de modelos mais robustos e a exploração de arquiteturas cada vez mais profundas, resultando em melhorias substanciais no desempenho em tarefas diversas, como visão computacional, processamento de linguagem natural e reconhecimento de padrões. Adicionalmente, o vasto acervo de dados disponíveis atualmente tem desempenhado um papel fundamental no sucesso do AM. A proliferação de dispositivos conectados à Internet e sensores em uma variedade de campos, juntamente com o armazenamento eficiente de informações, tem gerado conjuntos de dados de dimensões sem precedentes. Essa abundância de dados permite o treinamento de modelos mais precisos e a extração de insights mais profundos, contribuindo para a ampla aplicabilidade do AM em

cenários do mundo real, como medicina, finanças, indústria e ciência (FRADKOV, 2020). Para compreender plenamente a ascensão atual do AM e seu potencial futuro, é fundamental traçar as origens dessa disciplina.

Ao longo dos tempos, desde a antiguidade, quando Homero mencionou as “tripés” mecânicas a serviço dos deuses, máquinas imaginárias de assistência mecânica têm sido parte integrante da nossa cultura. Filósofos usaram a ideia de máquinas inteligentes como uma ferramenta literária para nos ajudar a definir a essência da humanidade. Por exemplo, René Descartes com o “homem mecânico” e Gottfried Wilhelm Leibniz que vislumbrou a possibilidade de usar dispositivos mecânicos para resolver disputas através de regras lógicas. Autores de ficção científica utilizaram a perspectiva de máquinas inteligentes para explorar a fantasia de seres não humanos inteligentes, nos incitando a refletir sobre nossas próprias características humanas. No entanto, somente nas últimas décadas, a comunidade de IA conseguiu construir máquinas experimentais que colocam à prova teorias sobre os mecanismos do pensamento e do comportamento inteligente, efetivamente demonstrando conceitos que anteriormente existiam apenas como possibilidades teóricas (BUCHANAN, 2005).

Os primeiros avanços significativos na área do Aprendizado de Máquina são atribuídos a Frank Rosenblatt, cuja notável contribuição consistiu no desenvolvimento do modelo de perceptron. Esse modelo foi concebido com base em analogias ao sistema nervoso humano, demonstrando a capacidade de converter sinais analógicos em discretos, mediante a aplicação de um limiar (*threshold*) apropriado (ROSENBLATT, 1957; ROSENBLATT, 1958; ROSENBLATT, 1960). A inovação introduzida por Rosenblatt marcou o ponto de partida no desenvolvimento das Redes Neurais Artificiais (RNAs), representando, efetivamente, o pioneirismo na elaboração dos primeiros estudos matemáticos e práticos nesse campo. O ano de 1962 testemunhou outro marco relevante na história do Aprendizado de Máquina, quando Seymour Novikoff formulou um teorema fundamental que estabelece condições essenciais para a convergência de aprendizado de algoritmos perceptron (NOVIKOFF, 1962). Esse teorema, elaborado por Novikoff, forneceu uma base teórica sólida e critérios bem definidos para a convergência dos perceptrons, tornando-se um pilar conceitual e funcional para a compreensão do processo de treinamento de RNAs.

A coexistência desses dois eventos históricos, com a contribuição pioneira de Rosenblatt na criação do modelo perceptron e a subsequente formalização teórica de Novikoff, desempenhou um papel fundamental no estabelecimento das bases sólidas do Aprendizado de Máquina e nas primeiras incursões na compreensão e aplicação de RNAs. Essas realizações iniciais abriram caminho para o desenvolvimento e a evolução posteriores do campo, levando à diversificação de algoritmos de aprendizado e arquiteturas de RNAs, bem como ao surgimento de técnicas mais avançadas no domínio do Aprendizado de Máquina. No começo da década de 1960, diversos grupos estavam envolvidos no projeto e teste de sistemas de reconhecimento de padrões. Em 1963-1964, foram propostas declarações gerais sobre o problema de reconhecimento de

padrões por Abramson et al. (ABRAMSON; BRAVERMAN; SEBESTYEN, 1963) e Widrow (WIDROW, 1964). Nesta época foram desenvolvidas abordagens tanto determinísticas quanto probabilísticas para o problema.

Em 1969, M. Minsky e S. Papert publicaram um livro (MINSKY; PAPERT, 1969) que estabeleceu algumas limitações importantes em relação ao que os perceptrons poderiam realizar. Como por exemplo, o fato de que os perceptrons não eram capazes de representar certas funções lógicas, como XOR ou NXOR. Essas descobertas tiveram um impacto significativo, levando a um declínio nas pesquisas nessa área que persistiu até o início da década de 1980. Como resultado desse declínio, o financiamento para pesquisas em Inteligência Artificial (IA) também foi reduzido consideravelmente, o que deu origem a um período que é agora chamado de "primeiro inverno da IA" (FRADKOV, 2020).

Na década de 1970-1980, estudos sobre as redes neurais multicamadas avançaram. Em 1980, Kunihiko Fukushima propôs o neocognitron, uma rede neural convolucional hierárquica multicamada (FUKUSHIMA, 1980). Cita-se também, nesta época o algoritmo de aprendizado *backpropagation*, inventado em meados dos anos 1980 por Rumelhart e outros (RUMELHART; HINTON; WILLIAMS, 1986), teve um grande impacto, embora suas ideias iniciais datem dos anos 1960. Entretanto, o *backpropagation* possui limitações, especialmente quando as classes não podem ser separadas linearmente. Apesar da esperança gerada por essas inovações, resultados aquém das expectativas levaram a outra redução de investimentos em aprendizado de máquina, conhecido como o segundo inverno da IA. No entanto, o interesse em redes neurais para aprendizado de máquina cresceu nos anos 1990, culminando com o impactante artigo de Cortes e Vapnik (1995), que introduziu as *support-vector networks* para casos não separáveis (CORTES; VAPNIK, 1995). Essas se provaram eficazes e tiveram um grande impacto na academia, inspirando pesquisas em diversas áreas (FRADKOV, 2020).

A primeira década do século atual foi um marco na história da Aprendizagem de Máquina, impulsionado pelas duas tendências do surgimento do Big Data e a redução dos custos de computação paralela e memória, exemplificada pelo MapReduce e pelo Hadoop do Google, que permitiu o processamento distribuído de dados em grande escala. Mais recentemente, chegando aos dias atuais, o desenvolvimento do aprendizado profundo, originado do perceptron, tem ganhado muito destaque. Na década passada, o campo das Redes Neurais Profundas (*Deep Neural Networks - DNN*) experimentou um avanço notável. Geoffrey Hinton (Britânico que agora vive no Canadá), junto com Ian LeCun e Yehoshua Benjo, conhecidos como LBH, desempenharam um papel essencial na popularização das DNN, publicando extensivamente sobre a área, ganhando reconhecimento na comunidade científica. Em 2018, o trio recebeu o Prêmio Turing e o Canadá se tornou um epicentro de Machine Learning, hospedando conferências de destaque como NeurIPS. Essas realizações marcaram uma era de avanços significativos na área (FRADKOV, 2020).

## 2.2 Conceitos Fundamentais

Nesta seção, exploraremos os conceitos fundamentais que formam a base do aprendizado de máquina e forneceremos uma visão geral dos conceitos necessários para a compreensão completa da metodologia de trabalho adotada, assim como dos resultados obtidos nesta tese.

### 2.2.1 Supervisionado X Não Supervisionado

A maioria dos algoritmos de ML pode ser dividida em duas categorias: aprendizagem supervisionada e não supervisionada (GOODFELLOW; BENGIO; COURVILLE, 2016) há também a chamada aprendizagem semi-supervisionada e por reforço (RAY, 2019), entretanto, não entraremos em detalhes nestes dois últimos casos, visto que não serão utilizados em nossas aplicações. Por supervisionado, entende-se observar e dirigir a execução de uma tarefa, projeto ou atividade. Isso é feito carregando o algoritmo com conhecimento para que possamos prever instâncias futuras, esse conhecimento é dado na forma de dados rotulados nos quais o modelo deve ser treinado (CUNNINGHAM; CORD; DELANY, 2008). Existem dois tipos de técnicas de aprendizado supervisionado, são elas: classificação e regressão (DELUA, 2021). Os algoritmos de classificação produzem valores categóricos, ou seja, indicam pertencer a um determinado grupo, enquanto os regressores tentam estimar quantitativamente um valor de saída, com base nos dados fornecidos. A ideia central é fazer do modelo dos algoritmos uma função que mapeie as entradas para as saídas desejadas, por meio dos exemplos dados (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Os algoritmos geralmente começam com o conjunto de dados rotulados, divididos em dois grupos: treinamento e teste, normalmente os dados de teste estão entre 10% e 30 % da amostra. Em seguida, os dados de treinamento são utilizados para modelar uma função, e a partir disso, os dados de teste são utilizados para medir sua eficácia e possíveis ações para melhorar o modelo a ser realizado. A partir deste ponto, novos dados podem ser introduzidos na função para fazer as previsões desejadas (MAIMON; ROKACH, 2005).

Para o caso de modelos não supervisionados, a ideia central deste tipo de algoritmo é, a partir de um conjunto de dados não rotulados, tentar encontrar padrões e semelhanças entre características, permitindo assim novas possibilidades de interpretação e caracterização do campo amostral (ALPAYDIN, 2004). Em comparação com a aprendizagem supervisionada, a não supervisionada possui menos modelos e menos métodos de avaliação, que são responsáveis por garantir a precisão do resultado. A redução de dimensão, estimativa de densidade e agrupamento são técnicas de aprendizado de máquina não supervisionadas amplamente utilizadas (DELUA, 2021).

### 2.2.2 Regressão e Classificação

Ao descrevermos uma variável, podemos representá-la como quantitativa ou qualitativa. Variáveis quantitativas são aquelas que assumem valores numéricos, como idade, altura, e

temperatura. Por outro lado, variáveis qualitativas representam diferentes classes ou categorias. Exemplos de variáveis qualitativas incluem tipo de carro (sedan, SUV, caminhonete), preferência de pagamento (cartão de crédito, dinheiro, débito) ou resultado de uma pesquisa de satisfação (satisfeito, neutro, insatisfeito). Em termos de problemas de análise de dados, nos referimos geralmente a problemas que envolvem uma resposta quantitativa como problemas de regressão, já os de resposta qualitativa são geralmente chamados de problemas de classificação (WITTEN; JAMES, 2013).

A regressão é uma técnica fundamental no campo do aprendizado de máquina que visa modelar e analisar a relação entre variáveis. Ela desempenha um papel crucial na previsão e no entendimento de tendências e padrões em dados. Em termos simples, a regressão busca encontrar a melhor maneira de relacionar uma variável de saída (ou resposta) a uma ou mais variáveis de entrada (ou preditores). Um exemplo clássico de aplicação da regressão é prever o preço de uma casa com base em várias características, como área, número de quartos, localização, etc. A regressão nos permite criar um modelo matemático que descreve como as características estão relacionadas ao preço da casa, permitindo-nos fazer previsões precisas. A regressão linear é um dos métodos mais simples e amplamente utilizados na análise de dados. Ela descreve a relação entre uma variável de saída (dependente) e uma única variável de entrada (independente) por meio de uma equação de reta. A equação da reta de regressão assume a forma  $y = mx + b$ , onde  $y$  representa a variável de saída,  $x$  é a variável de entrada,  $m$  é a inclinação da reta (coeficiente angular) e  $b$  é a interceptação no eixo  $y$  (coeficiente linear). Enquanto a regressão linear é adequada para modelar relacionamentos lineares simples, a regressão polinomial expande essa capacidade para capturar relacionamentos não-lineares entre variáveis. Nesse tipo de regressão, utilizamos equações polinomiais para descrever a relação entre as variáveis. Por exemplo, uma regressão polinomial de segunda ordem pode ter uma equação como  $y = ax^2 + bx + c$ . Isso é especialmente útil quando os dados exibem padrões curvilíneos ou não-lineares, como crescimento exponencial ou diminuição de taxa variável. A regressão polinomial nos permite ajustar curvas mais complexas aos dados, melhorando a precisão das previsões (MONTGOMERY; PECK; VINING, 2021).

Já a classificação é uma técnica dedicada a atribuir rótulos ou categorias a dados de entrada com base em suas características. A habilidade de um modelo de aprendizado de máquina em classificar corretamente novos exemplos é fundamental para o seu sucesso, tornando a classificação uma das tarefas centrais neste domínio. Suas aplicações são vastas, estendendo-se a áreas como reconhecimento de padrões (HART; STORK; DUDA, 2000), processamento de linguagem natural (YOUNG et al., 2018), diagnóstico médico (RAJKOMAR; DEAN; KOHANE, 2019), detecção de fraudes (AWOYEMI; ADETUNMBI; OLUWADARE, 2017) e muito mais. Em um cenário de classificação binária, o sistema deve fazer uma escolha entre duas categorias mutuamente exclusivas, enquanto a classificação multiclasse lida com três ou mais categorias possíveis. Para realizar a classificação, os modelos de aprendizado de máquina analisam as características dos dados de entrada e aplicam algoritmos para tomar

decisões. Como exemplo, um sistema para identificar dígitos escritos a mão é um problema de classificação (BISHOP; NASRABADI, 2006). A precisão da classificação é um fator crítico na avaliação do desempenho desses modelos, e várias técnicas e algoritmos são empregados para atingir resultados precisos e confiáveis.

### 2.2.3 Funções Custo

No domínio da Aprendizagem de Máquina, a escolha adequada de uma função custo desempenha um papel crítico no treinamento e avaliação de modelos. As funções de custo, também conhecidas como funções de perda ou funções de erro (HASTIE et al., 2009), são responsáveis por quantificar o quão bem um modelo se ajusta aos dados de treinamento e, conseqüentemente, orientam o processo de otimização. As funções de custo têm o objetivo fundamental de quantificar o desempenho do modelo, atribuindo um valor numérico que reflete o quão distante ou próximo as previsões do modelo estão dos valores reais dos dados de treinamento. Quanto menor o valor da função de custo, melhor o ajuste do modelo aos dados. Além de medir o desempenho, as funções de custo orientam o processo de otimização do modelo. Durante o treinamento, o objetivo é minimizar a função de custo, ajustando os parâmetros do modelo. Isso é alcançado por meio de técnicas de otimização, como o gradiente descendente, que ajustam iterativamente os parâmetros para reduzir a função de custo.

A escolha da função de custo a ser utilizada depende do tipo de problema de aprendizado que está sendo abordado. Em problemas de regressão, onde a saída é uma variável contínua, as duas funções de custo mais comuns são: o Erro Quadrático Médio (MSE) e o Erro Absoluto Médio (MAE). O Erro Quadrático Médio é uma função de custo amplamente empregada em problemas de regressão (GÉRON, 2022). Sua equação é dada por:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (1)$$

onde  $n$  representa o número de exemplos no conjunto de dados.  $y_i$  é o valor real da saída para o exemplo  $i$ .  $\hat{y}_i$  é a previsão feita pelo modelo para o exemplo  $i$ . O MSE mede o quadrado da diferença entre as previsões do modelo e os valores reais. Quanto menor o valor do MSE, mais próximas estão as previsões do modelo dos valores reais.

O Erro Absoluto Médio é outra função de custo adequada para problemas de regressão. Sua equação é representada por:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (2)$$

onde os termos têm os mesmos significados que no MSE. O MAE mede a média das diferenças absolutas entre as previsões do modelo e os valores reais.

Além disso, em problemas de classificação, a Entropia Cruzada (Cross-Entropy) é uma função de custo comumente utilizada. Geralmente associada ao quão bem um modelo de classificação se ajusta aos dados de treinamento (GÉRON, 2022). Ela é particularmente relevante em problemas de classificação multiclasse. A equação da entropia cruzada é definida como:

$$J(\Theta) = -\frac{1}{m} \sum_{i=1}^m \sum_{k=1}^K y_k^{(i)} \log(\hat{p}_k^{(i)}) \quad (3)$$

onde  $J(\Theta)$ : É a função de custo, que mede a discrepância entre as previsões do modelo e os rótulos reais dos dados de treinamento.  $\Theta$ : É o vetor de parâmetros do modelo. Ele contém os pesos e os vieses que o modelo ajusta durante o treinamento para minimizar a função de custo.  $m$ : O número total de exemplos de treinamento no conjunto de dados.  $K$  é o número de classes.  $y_k^{(i)}$  é o rótulo real (ou alvo) do exemplo de treinamento  $i$  para a classe  $k$ . É uma variável binária que indica se o exemplo pertence ou não à classe  $k$ .  $\hat{p}_k^{(i)}$ : Esta é a probabilidade prevista pelo modelo de que o exemplo  $i$  pertença à classe  $k$ . O objetivo principal da entropia cruzada é medir o quão bem as probabilidades previstas pelo modelo se alinham com os rótulos reais dos dados. Quanto menor o valor da entropia cruzada, melhor o ajuste do modelo aos dados.

A função de custo é mais do que apenas uma métrica de avaliação, é a bússola que guia o aprendizado de máquina em direção ao seu destino: modelos mais precisos e generalizáveis. A escolha sábia da função de custo não só influencia a capacidade do modelo de se ajustar aos dados, mas também determina sua capacidade de fazer previsões confiáveis em situações do mundo real. Por fim, destaca-se que é crucial dominar a arte de escolher e implementar a função de custo adequada, pois isso não apenas impulsiona o progresso da aprendizagem de máquina, mas também nos aproxima de soluções inteligentes para desafios complexos do mundo real.

#### 2.2.4 *Overfitting* e *Underfitting*

Quando começamos a nos aprofundar no domínio do aprendizado de máquina, dois problemas muito comuns podem aparecer e prejudicar o desempenho do modelo. *Overfitting* (Superajuste) e *underfitting* (subajuste) são dois problemas comuns no treinamento de modelos de aprendizado de máquina e representam extremos opostos no desempenho do modelo.

O *overfitting* ocorre quando um modelo é muito complexo em relação aos dados de treinamento. Isso significa que o modelo se ajusta tão bem aos dados de treinamento que captura não apenas os padrões reais, mas também o ruído e as flutuações aleatórias nos dados. O resultado do *overfitting* é que o modelo tem um desempenho excepcional nos dados de treinamento, mas um desempenho ruim em dados de teste ou em novos dados que não foram vistos durante o treinamento, ou seja, possui uma capacidade pobre de generalização. Já o

*underfitting* ocorre quando um modelo é muito simples em relação à complexidade dos dados de treinamento. Isso significa que o modelo não consegue capturar adequadamente os padrões presentes nos dados. O resultado do *underfitting* é que o modelo não consegue ajustar nem mesmo os dados de treinamento adequadamente, levando a um desempenho fraco tanto nos dados de treinamento quanto nos dados de teste. No *underfitting* modelo tem alto viés e baixa variância, sendo que viés (ou *bias*) se refere a uma assunção forte e a variância se refere a dependência dos dados de treino (KOEHRSEN, 2018).

Reduzir tanto o *underfitting* quanto o *overfitting* é essencial para desenvolver modelos de aprendizado de máquina que generalizem bem para novos dados. Para combater o *underfitting*, é fundamental aumentar a complexidade do modelo. Isso pode envolver a escolha de algoritmos mais complexos, o aumento do número de recursos ou a realização de um pré-processamento mais completo dos dados. Além disso, ajustar os hiperparâmetros, como a taxa de aprendizado, pode ajudar a tornar o modelo mais flexível. Por outro lado, para mitigar o *overfitting*, é necessário simplificar o modelo. Isso pode ser alcançado reduzindo a complexidade do algoritmo, diminuindo o número de recursos ou aplicando técnicas de regularização, como L1 ou L2, para penalizar parâmetros excessivamente complexos (RUPPERT; CARROLL, 2000). Além disso, também é possível melhorar o desempenho do modelo aumentando o tamanho do conjunto de treinamento, se possível, para fornecer ao modelo uma visão mais abrangente dos dados. Encontrar o equilíbrio adequado entre complexidade e simplicidade é uma parte fundamental do ajuste de modelos de aprendizado de máquina. Isso requer experimentação e ajuste fino, acompanhados de validação cruzada e monitoramento constante do desempenho do modelo (KOEHRSEN, 2018).

## 2.2.5 Métricas de Avaliação e Validação

A avaliação de modelos em ML desempenha um papel crucial no desenvolvimento de soluções de aprendizado de máquina confiáveis e precisas. Ela envolve a aplicação de métricas e técnicas específicas para medir o desempenho de um modelo em relação aos dados de teste ou validação. Uma avaliação adequada permite que os cientistas de dados e engenheiros de ML entendam como seus modelos estão funcionando, identifiquem problemas e façam ajustes para melhorar o desempenho.

A divisão de dados desempenha um papel crucial no processo de treinamento, avaliação e validação de modelos de machine learning. É uma prática fundamental que envolve a separação do conjunto de dados disponível em duas ou mais partes distintas: geralmente, um conjunto de treinamento e um conjunto de teste. Essa divisão é essencial para garantir uma avaliação imparcial e realista do desempenho do modelo. O conjunto de treinamento é usado para alimentar o modelo durante a fase de treinamento, permitindo que ele aprenda a relação entre as variáveis de entrada e a variável de saída. É nesse conjunto que o modelo ajusta seus parâmetros e tenta capturar padrões e informações relevantes dos dados. Por outro lado,

o conjunto de teste é reservado para avaliar o desempenho do modelo após o treinamento. Ele contém dados não vistos pelo modelo durante o treinamento e, portanto, atua como um indicador de quão bem o modelo generaliza para novos exemplos. A divisão de dados é geralmente feita de forma aleatória, mas também pode envolver estratégias mais avançadas, como validação cruzada, que divide os dados em múltiplos conjuntos de treinamento e teste para uma avaliação mais robusta. É importante manter uma proporção equilibrada entre os conjuntos de treinamento e teste, garantindo que haja dados suficientes para treinar o modelo adequadamente e avaliar sua capacidade de generalização (MÜLLER; GUIDO, 2016). Outra prática comum é a divisão dos dados em três conjuntos: treinamento, validação e teste. O conjunto de validação é usado durante o treinamento para ajustar hiperparâmetros e otimizar o modelo, enquanto o conjunto de teste permanece isolado até a avaliação final. Isso ajuda a evitar o *overfitting*, onde o modelo se adapta demais aos dados de treinamento e não generaliza bem para novos dados.

A seguir iremos abordar em mais detalhes algumas métricas de avaliação específicas para a classificação e regressão.

#### 2.2.5.1 Métricas na Classificação

Existem várias métricas de classificação amplamente utilizadas, cada uma com seu propósito e interpretação específicos. A seguir serão abordadas algumas das mais comuns.

**Acurácia (*accuracy*):** Uma das métricas mais simples e comuns é a acurácia (*accuracy*), que mede a proporção de exemplos classificados corretamente em relação ao total de exemplos, como visto a seguir.

$$\text{Acurácia} = \frac{\text{Número de previsões corretas}}{\text{Número total de previsões}} \quad (4)$$

Embora seja uma métrica fácil de entender, a acurácia pode ser enganosa em casos de conjuntos de dados desbalanceados, onde uma classe é muito mais prevalente do que outra. Outra métrica importante é a matriz de confusão (*confusion matrix*), que fornece uma visão mais detalhada do desempenho do modelo. Ela divide as previsões em quatro categorias: verdadeiros positivos, verdadeiros negativos, falsos positivos e falsos negativos (BROWNLEE, 2014). Com base nesses valores, várias métricas podem ser calculadas, incluindo:

**Precisão (*precision*):** mede a proporção de exemplos positivos previstos corretamente em relação ao total de exemplos previstos como positivos. É útil quando o foco está na minimização de falsos positivos. Sua equação é dada por:

$$\text{Precisão} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Positivos}} \quad (5)$$

**Revocação (*recall*):** mede a proporção de exemplos positivos previstos corretamente em relação ao total de exemplos reais positivos. É útil quando o foco está na minimização de

falsos negativos. Sua equação é dada por:

$$\text{Revocação} = \frac{\text{Verdadeiros Positivos}}{\text{Verdadeiros Positivos} + \text{Falsos Negativos}} \quad (6)$$

*F1-Score*: É a média harmônica entre precisão e revocação e é especialmente importante quando há uma relação de troca (*trade-off*) entre essas duas métricas. O *trade-off* ocorre porque, ao tentar maximizar a precisão, o modelo pode se tornar mais conservador e prever menos exemplos como positivos, resultando em um aumento nos falsos negativos (reduzindo a revocação). Por outro lado, ao tentar maximizar a revocação, o modelo pode prever mais exemplos como positivos, mas isso pode levar a um aumento nos falsos positivos (reduzindo a precisão). O *F1-Score* busca um equilíbrio entre esses extremos, sendo ideal quando se deseja considerar tanto os falsos positivos quanto os falsos negativos. Sua equação é dada por:

$$\text{F1-Score} = \frac{2 \cdot \text{Precisão} \cdot \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (7)$$

#### 2.2.5.2 Métricas na Regressão

Assim como para a classificação, na regressão diversas métricas podem ser utilizadas para se avaliar o desempenho do modelo. Entre as mais utilizadas, se destacam o Erro Quadrático Médio e Erro Absoluto Médio, que já foram abordados na subseção 2.2.3, quando abordamos as funções custo. Em seguida, serão abordadas outras métricas muito populares para a avaliação de desempenho na regressão.

O coeficiente de determinação, representado como  $R^2$  (R-squared), é uma métrica que varia de 0 a 1 e quantifica a proporção da variabilidade na variável de saída que é explicada pelo modelo. Um valor de  $R^2$  próximo de 1 indica que o modelo é capaz de explicar a maior parte da variabilidade dos dados, enquanto um valor próximo de 0 sugere que o modelo não está capturando bem os padrões nos dados (CHICCO; WARRENS; JURMAN, 2021). Sua equação é dada por:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (8)$$

Além dessas métricas, existem outras métricas de regressão mais avançadas, como o Erro Quadrático Médio Normalizado (NMSE - Normalized Mean Squared Error), que normaliza o MSE pelo valor médio dos dados de teste. Isso permite comparar o desempenho de modelos em diferentes escalas de dados. Sua equação é dada por:

$$NMSE = \frac{1}{n} \sum_{i=1}^n \left( \frac{y_i - \hat{y}_i}{\bar{y}} \right)^2 \quad (9)$$

É importante escolher a métrica de regressão mais apropriada com base na natureza do problema e nas características dos dados. Por exemplo, se os valores discrepantes são problemáticos, o MAE pode ser preferível ao MSE. Além disso, interpretar as métricas em contexto é fundamental para entender o impacto prático do desempenho do modelo de regressão.

### 2.2.5.3 Validação Cruzada

A validação cruzada é uma técnica essencial em aprendizado de máquina para avaliar o desempenho de modelos de forma robusta e confiável. Trata-se de uma técnica da família dos métodos de Monte Carlo. Ela é especialmente útil quando se trabalha com conjuntos de dados limitados e visa estimar como um modelo treinado em um conjunto de dados específico se comportaria ao fazer previsões em dados não vistos. O conceito principal por trás da validação cruzada é dividir o conjunto de dados em partes (ou "*folds*") para treinar e avaliar o modelo múltiplas vezes. A forma mais comum de validação cruzada é a validação cruzada *k-fold*, onde os dados são divididos em *k* partes aproximadamente iguais. O modelo é treinado *k* vezes, cada vez usando *k* - 1 partes para treinamento e a parte restante para avaliação. Esse processo é repetido *k* vezes, de forma que cada parte do conjunto de dados seja usada como conjunto de teste exatamente uma vez. As métricas de desempenho são então calculadas a partir das *k* iterações, e geralmente é calculada a média ou a mediana dessas métricas para obter uma estimativa robusta do desempenho do modelo (BERRAR et al., 2019).

A validação cruzada ajuda a evitar problemas como *overfitting* e *underfitting*. Além da validação cruzada *k-fold*, existem outras variantes, como a validação cruzada estratificada, que mantém a proporção de classes em cada *fold*, e a validação cruzada leave-one-out (LOOCV), que usa um único exemplo como conjunto de teste em cada iteração. A escolha do número de *folds* *k* pode variar dependendo do tamanho do conjunto de dados e do objetivo da avaliação. Em geral, valores comuns para *k* incluem 5 e 10. A validação cruzada fornece uma estimativa mais robusta do desempenho do modelo, permitindo uma avaliação mais confiável de sua capacidade de generalização para novos dados. É uma prática recomendada ao desenvolver e ajustar modelos de aprendizado de máquina (BERRAR et al., 2019).

### 2.2.5.4 Desafios na Avaliação

A avaliação de modelos de aprendizado de máquina é uma etapa crítica, mas também repleta de desafios. Alguns destes incluem por exemplo o *overfitting* e *underfitting* já citados. Geralmente, estes estão associados ao próprio banco de dados que será utilizado pelo algoritmo afim de performar as operações.

Em problemas de classificação, pode haver um desequilíbrio significativo entre as classes. Por exemplo, em um problema de detecção de fraudes, a maioria das transações pode ser legítima, tornando as transações fraudulentas uma classe minoritária. Isso pode levar a métricas enganosas, como a acurácia, que podem não refletir adequadamente o desempenho do modelo. Métricas como precisão, recall e F1-score são mais informativas nesses casos (BROWNLEE, 2015). Já em problemas que envolvem dados de séries temporais, como previsão de preços de ações ou previsão meteorológica, a validação tradicional pode não ser apropriada. A dependência temporal dos dados requer estratégias especiais, como validação cruzada temporal ou caminhada no tempo (BROWNLEE, 2016).

Em muitos casos, o tamanho do conjunto de dados disponível pode ser limitado. Isso pode levar a problemas de generalização, onde o modelo pode não ser capaz de capturar padrões significativos devido à falta de dados. Aumentar o tamanho do conjunto de dados sempre que possível é uma abordagem recomendada, porém um dos fatores fundamentais é a representatividade dos dados sobre a distribuição original (ALTHNIAN et al., 2021). Outro fator a ser considerado é que os dados usados para treinar modelos podem conter vieses, refletindo preconceitos existentes ou desigualdades na sociedade. Modelos treinados com esses dados podem perpetuar esses preconceitos. Identificar e mitigar o viés nos dados é um desafio ético e técnico (LEE; RESNICK; BARTON, 2019). Por outro lado, já não mais falando dos dados, os modelos de aprendizado de máquina são complexos, como as redes neurais profundas e podem ser difíceis de interpretar (GIORDANO et al., 2021). Entender como um modelo toma decisões é importante, especialmente em aplicações críticas, como assistência médica e direção autônoma. Lidar com esses desafios requer uma abordagem cuidadosa e muitas vezes a combinação de várias técnicas e estratégias de avaliação. Além disso, é importante estar ciente desses desafios ao desenvolver e implantar modelos de aprendizado de máquina para garantir que eles sejam eficazes e éticos em seu uso.

### 2.2.6 Hiperparâmetros e Otimização

Hiperparâmetros e otimização são aspectos críticos no processo de desenvolvimento de modelos de aprendizado de máquina. Hiperparâmetros são configurações ajustáveis do modelo que não são aprendidas diretamente dos dados, mas desempenham um papel fundamental em sua eficácia. Esses hiperparâmetros incluem, por exemplo, a taxa de aprendizado, o número de camadas ocultas em uma rede neural, o número de árvores em um algoritmo de floresta aleatória e muitos outros. A escolha correta dos hiperparâmetros pode ter um impacto significativo no desempenho do modelo (CLAESEN; MOOR, 2015).

A otimização de hiperparâmetros é o processo de encontrar a combinação ideal de hiperparâmetros que resulta no melhor desempenho do modelo. Isso geralmente envolve a realização de experimentos com diferentes valores de hiperparâmetros e a avaliação do desempenho do modelo em um conjunto de validação. Técnicas como busca em grade e busca aleatória são comumente usadas para automatizar esse processo. Além disso, a otimização de hiperparâmetros também pode envolver a utilização de técnicas mais avançadas, como otimização bayesiana, que busca encontrar os melhores hiperparâmetros de forma mais eficiente do que a busca em grade. Um conjunto bem ajustado de hiperparâmetros pode fazer a diferença entre um modelo que não atende às expectativas e um modelo altamente preciso e eficaz. Portanto, é uma prática recomendada investir tempo e esforço na otimização cuidadosa dos hiperparâmetros ao desenvolver modelos de aprendizado de máquina (GÉRON, 2022).

## 2.2.7 Pré-Processamento de Dados

O pré-processamento de dados é uma etapa crucial no fluxo de trabalho de aprendizado de máquina, destinada a preparar e limpar os dados brutos para que possam ser utilizados de maneira eficaz por algoritmos de aprendizado. Essa etapa envolve várias técnicas e tarefas essenciais para garantir a qualidade e a adequação dos dados. a seguir, abordaremos em detalhe alguns pontos importantes no pré-processamento, indicados por (GARCÍA et al., 2016).

1. **Limpeza de Dados:** Isso envolve a identificação e o tratamento de dados ausentes, inconsistentes ou incorretos. Isso pode incluir preenchimento de valores ausentes, remoção de outliers (valores extremos) e tratamento de erros nos dados.
2. **Normalização e Padronização:** Para garantir que diferentes características tenham a mesma escala e não dominem umas sobre as outras, é comum aplicar técnicas de normalização (escala para um intervalo específico) e padronização (média zero e desvio padrão um) aos dados.
3. **Codificação de Variáveis Categóricas:** Muitos algoritmos de aprendizado de máquina não podem lidar diretamente com variáveis categóricas (por exemplo, cores ou categorias de produtos). Portanto, é necessário converter essas variáveis em um formato numérico, usando técnicas como codificação *one-hot*.
4. **Seleção de Características:** Nem todas as características podem ser relevantes para a tarefa de aprendizado. A seleção de características envolve a escolha das características mais informativas e a eliminação das menos relevantes para melhorar a eficiência e a eficácia do modelo.
5. **Engenharia de Características:** Às vezes, é útil criar novas características com base nas existentes para melhorar o desempenho do modelo. Isso pode envolver a combinação de características, criação de características de interação ou transformações específicas.
6. **Balanceamento de Classes:** Quando se lida com problemas de classificação em que as classes são desequilibradas (uma classe tem muito mais exemplos do que a outra), é importante aplicar técnicas de balanceamento, como *oversampling* (aumentar a classe minoritária) ou *undersampling* (diminuir a classe majoritária).
7. **Tratamento de Dados Temporais:** Em dados de séries temporais, como previsão do tempo ou análise financeira, técnicas específicas de pré-processamento, como diferenciação ou suavização, podem ser necessárias.

O pré-processamento de dados desempenha um papel fundamental na construção de modelos de aprendizado de máquina eficazes e confiáveis. Uma abordagem cuidadosa e bem planejada para essa etapa pode ter um impacto significativo no desempenho e na capacidade

de generalização do modelo. Portanto, é essencial investir tempo e esforço na preparação adequada dos dados antes de prosseguir para o treinamento e a avaliação do modelo.

## 2.3 Exemplos de Algoritmos

Nesta seção, exploraremos alguns dos algoritmos fundamentais em aprendizado de máquina que desempenham um papel crucial na construção de modelos e na resolução de problemas complexos. Cada um desses algoritmos representa uma abordagem única para a aprendizagem de padrões e modelagem de dados. Abordaremos brevemente cinco algoritmos-chave que abrangem uma variedade de técnicas e aplicações. O primeiro é o  $k$ -Nearest Neighbors ( $k$ -NN): Um método de classificação e regressão simples baseado na proximidade dos pontos de dados. O  $k$ -NN é eficaz em problemas nos quais a similaridade entre exemplos é um indicador importante. Em seguida, abordaremos as *Multi-Layer Perceptron* (MLP): Uma forma de rede neural artificial composta por várias camadas de neurônios interconectados. O MLP é utilizado para tarefas de aprendizado profundo e é capaz de modelar relacionamentos complexos entre variáveis. A terceira classe de algoritmo são as árvores de decisão, estruturas de árvore que dividem um conjunto de dados em segmentos menores com base nas características mais relevantes. Árvores de decisão são amplamente usadas em classificação e regressão. Em seguida abordaremos as máquinas de Vetores de Suporte (SVM): Algoritmos de aprendizado supervisionado que mapeiam os dados em um espaço de alta dimensão e procuram encontrar um hiperplano de separação ótimo entre as classes. E por fim, encerramos a seção com a Análise do Componente Principal (PCA): Uma técnica de redução de dimensionalidade que identifica as direções de maior variância nos dados originais. PCA é comumente usado para simplificar conjuntos de dados complexos. Vamos a seguir analisar cada um desses algoritmos em mais detalhes.

### 2.3.1 $k$ -Nearest Neighbors

O princípio básico do  $k$ -NN é a presunção de que objetos semelhantes estão próximos uns dos outros, estes são chamados de vizinhos. Normalmente o algoritmo é baseado na distância euclidiana entre a amostra de teste e os dados de treinamento (PETERSON, 2009). Para atribuir uma classe a um novo dado, o algoritmo observa os  $k$  (número arbitrário) vizinhos mais próximos, identificando a que pertence a maioria (MOHAMMED; KHAN; BASHIER, 2016). O  $k$ -NN tem as vantagens de ser uma técnica simples e de fácil implementação, sendo um esquema de classificação flexível e com uma taxa de erro no máximo duas vezes a taxa de erro de Bayes, sendo o melhor algoritmo em certas ocasiões. Tem a desvantagem de ser caro computacionalmente, à medida que os dados de treinamento aumentam, além de ser suscetível a ruídos, que podem diminuir a precisão do algoritmo. O  $k$ -NN é utilizado nas mais diversas áreas do conhecimento, desde sistemas de recomendação, a diagnósticos médicos, no mercado

financeiro, reconhecimento de vídeo, entre outros (RAY, 2019). Em termos gerais, a etapa principal do  $k$ -NN envolve:

1. Calcular a distância entre o ponto de teste e todos os pontos de treinamento. Neste caso, a distância euclidiana é a medida de distância mais comum e é calculada como a distância em linha reta entre dois pontos no espaço de características. Dado  $P = (p_1, p_2, \dots, p_n)$  como o ponto de teste e  $Q = (q_1, q_2, \dots, q_n)$  como um ponto de treinamento, a distância euclidiana entre esses dois pontos é dada por:

$$d(P, Q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \quad (10)$$

Outra medida de distância comum é a distância de Manhattan (Manhattan Distance ou L1-Norm), que é calculada como a soma das diferenças absolutas entre as coordenadas dos pontos:

$$d(P, Q) = \sum_{i=1}^n |p_i - q_i| \quad (11)$$

Além dessas duas medidas, outras medidas de distância, como a distância de Minkowski, também podem ser usadas, variando o valor do parâmetro  $p$  para ajustar a sensibilidade à diferença entre as coordenadas. Sua expressão é dada por:

$$d(P, Q) = \left( \sum_{i=1}^n |p_i - q_i|^p \right)^{\frac{1}{p}} \quad (12)$$

A escolha da medida de distância depende do problema específico e da natureza dos dados. A distância euclidiana é mais comumente usada, mas a distância de Manhattan pode ser preferível em casos onde as coordenadas têm unidades diferentes ou onde é importante considerar o movimento “em grade”.

2. Selecionar os  $k$  pontos de treinamento mais próximos com base nessas distâncias.
3. Atribuir ao ponto de teste a classe mais comum entre os  $k$  vizinhos mais próximos.
4. (Opcional, em variantes como  $k$ -means ou  $k$ -NN ponderado): Atualizar o centróide das classes (ou ajustar os pesos das contribuições dos vizinhos). Nos casos em que o algoritmo de  $k$ -NN é adaptado para gerar clusters (como no  $k$ -means) ou quando as previsões ponderam as contribuições de cada vizinho, o centróide de cada classe pode ser recalculado após cada atribuição, conforme a média das coordenadas dos pontos pertencentes a essa classe. A atualização do centróide de uma classe  $C_j$  com  $N$  pontos é dada por:

$$\text{Centróide}(C_j) = \frac{1}{N} \sum_{i=1}^N X^{(i)} \quad (13)$$

onde  $X^{(i)}$  representa os pontos atribuídos à classe  $C_j$ . Isso permite que o modelo se adapte dinamicamente e melhore a representatividade das classes.

O número de vizinhos  $K$  é um hiperparâmetro que precisa ser definido antes do treinamento do modelo. Um valor pequeno de  $K$  resulta em um modelo mais flexível que pode reconhecer padrões complexos nos dados, mas isso pode levar a possíveis problemas de superajuste. Por outro lado, um valor grande de  $K$  pode resultar em um modelo mais estável e robusto, mas pode falhar em capturar padrões locais nos dados. Normalmente, o algoritmo é iterado para encontrar o valor ótimo de  $K$ . Ao longo desta tese, este algoritmo será utilizado em aplicações práticas, tanto na sua versão classificador (para classificar fases quânticas) quanto regressor (afim de prever o fenômeno de sincronização quântica).

### 2.3.2 Perceptrons Multi Camadas

*Multi-Layer Perceptron* (MLP), ou "Perceptrons Multi-Camadas" em tradução livre, são uma forma de rede neural artificial composta por várias unidades de processamento (GORI; SCARSELLI, 1998). O funcionamento dessas unidades parte do princípio básico do perceptron, mas em vez de ter uma única camada de unidades de processamento, o MLP organiza essas unidades em camadas: uma camada de entrada, uma ou mais camadas ocultas, e uma camada de saída. O que diferencia o MLP dos perceptrons simples é a presença dessas camadas intermediárias (camadas ocultas), que permitem ao modelo aprender relações não-lineares entre as entradas e saídas.

Cada unidade de uma camada oculta recebe sinais de entrada, que são ponderados pelos pesos associados a eles. Em seguida, é feita uma soma ponderada desses sinais e, ao invés de simplesmente aplicar uma regra binária de ativação (como no perceptron simples), o MLP utiliza funções de ativação não lineares, como a função sigmoide, ReLU (*Rectified Linear Unit*), ou tangente hiperbólica (*tanh*). Essas funções de ativação introduzem a não linearidade necessária para que o modelo possa aprender padrões complexos.

Suponhamos que tenhamos " $p$ " sinais de entrada, representados por  $X_1, X_2, \dots, X_p$ , pesos correspondentes  $w_1, w_2, \dots, w_p$ , e um limiar (*threshold*)  $t$ . Nesse modelo, o "nível de atividade"  $a$  de uma unidade de processamento é calculado como:

$$a = w_1X_1 + w_2X_2 + \dots + w_pX_p + b \quad (14)$$

onde  $b$  é o *bias*, um termo adicional que permite que a função de ativação seja deslocada horizontalmente, facilitando o aprendizado de funções complexas.

A saída  $z$  da unidade após a aplicação da função de ativação  $\sigma(a)$  é:

$$z = \sigma(a) \quad (15)$$

A função de ativação  $\sigma(a)$  pode ser, por exemplo:

- Sigmoide:  $\sigma(a) = \frac{1}{1+e^{-a}}$

- ReLU:  $\sigma(a) = \max(0, a)$
- Tangente hiperbólica:  $\sigma(a) = \tanh(a)$

O MLP aprende ajustando os pesos  $w$  e os *biases*  $b$  durante o processo de treinamento, por meio de um algoritmo chamado *backpropagation* (retropropagação), que é combinado com um otimizador, como o gradiente descendente. Durante a fase de treinamento, o erro entre a saída prevista e a saída real é propagado para trás através da rede, camada por camada. O objetivo é minimizar uma função de custo  $J(\Theta)$ , ajustando os pesos para reduzir a discrepância entre as previsões do modelo e os dados reais. As principais etapas do funcionamento de uma rede MLP incluem:

1. Propagação para frente (*forward pass*): Os dados de entrada passam pela rede, camada por camada, e as ativações são calculadas em cada unidade até chegar à camada de saída, onde a previsão final é feita.
2. Cálculo do erro: O erro é calculado comparando a previsão do modelo com o rótulo real. Uma função de custo comum, como o erro quadrático médio (MSE) ou a entropia cruzada, é utilizada para medir a discrepância.
3. Retropropagação do erro (*backpropagation*): O erro calculado na camada de saída é propagado para trás, ajustando os pesos de cada camada com base no gradiente da função de custo em relação aos pesos.
4. Atualização dos pesos: Os pesos são atualizados usando um método de otimização, como o gradiente descendente, o gradiente descendente estocástico (SGD) ou otimizadores mais avançados, como Adam. A atualização dos pesos  $w$  é dada por:

$$w^{(t+1)} = w^{(t)} - \eta \frac{\partial J}{\partial w} \quad (16)$$

onde  $\eta$  é a taxa de aprendizado e  $\frac{\partial J}{\partial w}$  é o gradiente da função de custo em relação aos pesos.

5. Iteração: O processo é repetido até que o erro seja minimizado ou até que um critério de parada seja atingido.

Além do MLP, existem outros tipos de redes neurais, cada uma com sua arquitetura específica e aplicabilidade. Redes neurais convolucionais (CNNs) são amplamente utilizadas em tarefas de visão computacional, enquanto redes neurais recorrentes (RNNs) são adequadas para dados sequenciais, como linguagem natural e séries temporais. Redes neurais profundas (Deep Neural Networks ou DNNs) são compostas por múltiplas camadas ocultas de neurônios e são capazes de aprender representações complexas dos dados (ALZUBAIDI et al., 2021). As redes neurais têm demonstrado desempenho notável em uma ampla variedade de tarefas de

aprendizado de máquina, incluindo reconhecimento de padrões, processamento de linguagem natural, visão computacional e muito mais. Seu poder reside na capacidade de aprender representações abstratas dos dados, permitindo a resolução de problemas complexos (SARKER, 2021).

### 2.3.3 Árvores de Decisão

Árvores de Decisão são um dos algoritmos mais versáteis e amplamente utilizados em aprendizado de máquina, pois são eficazes tanto para problemas de classificação quanto para problemas de regressão. Esses algoritmos são chamados de “árvores” porque têm uma estrutura hierárquica semelhante a uma árvore, com um nó raiz, nós internos e folhas. Cada nó representa uma característica do conjunto de dados e uma decisão com base nessa característica. As ramificações a partir de um nó levam a outros nós, representando os diferentes caminhos de decisão. O primeiro nó, no topo da árvore, é chamado de nó raiz e contém a característica que é o melhor preditor para separar os dados. Os nós que não são folhas são chamados de nós internos. Eles representam testes de decisão e direcionam o fluxo de acordo com o valor da característica testada. Os nós folha são os terminais da árvore e representam as saídas ou rótulos. Por exemplo, em uma árvore de classificação, as folhas podem representar as classes às quais os exemplos pertencem. Para decidir como dividir os dados em cada nó, os algoritmos de árvore de decisão utilizam critérios como erro quadrático médio, ganho de informação ou índice de Gini, dependendo se é uma árvore de decisão para classificação ou regressão (KINGSFORD; SALZBERG, 2008). Para calcular o ganho de informação:

A entropia de um conjunto de dados  $S$  com relação a uma variável-alvo  $Y$  é dada por:

$$H(S) = - \sum_{y \in Y} p(y) \log_2(p(y)) \quad (17)$$

onde  $H(S)$  é a entropia de  $S$ .  $y$  é uma classe possível em  $Y$ .  $p(y)$  é a proporção de exemplos em  $S$  que pertencem à classe  $y$ .

O Ganho de Informação é calculado como a diferença entre a entropia do conjunto de dados pai (antes da divisão) e a entropia ponderada dos filhos (após a divisão):

$$\text{Information Gain} = H(S_{\text{pai}}) - \sum_{v \in \text{filhos}} \frac{|S_v|}{|S_{\text{pai}}|} H(S_v) \quad (18)$$

onde  $H(S_{\text{pai}})$  é a entropia do conjunto de dados pai.  $|S_v|$  é o número de exemplos no filho  $v$ .  $|S_{\text{pai}}|$  é o número de exemplos no conjunto de dados pai e  $H(S_v)$  é a entropia do filho  $v$ .

O Índice de Gini é calculado para os filhos da mesma maneira que para o Ganho de Informação, usando a distribuição de classe nos filhos. Para calcular o índice de Gini:

$$\text{Gini}(S) = 1 - \sum_{y \in Y} (p(y))^2 \quad (19)$$

onde  $Gini(S)$  é o índice de Gini de  $S$ .  $y$  é uma classe possível em  $Y$  e  $p(y)$  é a proporção de exemplos em  $S$  que pertencem à classe  $y$ .

Árvores de decisão têm uma tendência a se ajustar demais aos dados de treinamento, criando ramificações muito profundas e complexas. O *overfitting* pode ser evitado por meio de técnicas como poda da árvore ou limitação da profundidade máxima. Uma grande vantagem das árvores de decisão é a interpretabilidade. As regras de decisão ao longo dos caminhos da árvore podem ser facilmente entendidas e visualizadas. As árvores de decisão podem lidar tanto com variáveis categóricas quanto com numéricas, escolhendo automaticamente a melhor maneira de dividir os dados com base na natureza das características (PEDREGOSA et al., 2011). Para melhorar o desempenho e reduzir o *overfitting*, as árvores de decisão frequentemente são usadas em conjuntos, como o *Random Forest* (RF). Como pode se pressupor, uma RF possui um grande número de Árvores de Decisão. Cada árvore, de acordo com o seus parâmetros lógicos, vota em uma designação para um determinado *input*, o *output* com a maior quantidade de votos é a classificação final do algoritmo. A precisão é baseada no grande número de árvores “votando” (BREIMAN, 2001a). Sua interpretabilidade, versatilidade e capacidade de lidar com diferentes tipos de dados as tornam uma escolha popular em aprendizado de máquina.

*Random Forest* (RF) constitui uma metodologia de aprendizado de conjunto que funciona gerando inúmeras árvores de decisão durante a fase de treinamento e, subsequentemente, amalgamando suas previsões para derivar a decisão final (BREIMAN, 2001b). Essa técnica é baseada no princípio de “bagging” (*bootstrap aggregating*) (BREIMAN, 1996), em que cada árvore é construída utilizando um subconjunto aleatório dos dados com reposição, resultando em uma coleção heterogênea de modelos. Além disso, no processo de crescimento das árvores, o RF seleciona aleatoriamente um subconjunto de características em cada nó para a divisão, introduzindo variabilidade adicional entre as árvores. A aleatoriedade na seleção das amostras e das características é fundamental para aumentar a diversidade das árvores no conjunto, pois reduz a correlação entre elas, melhorando a generalização do modelo. A seleção aleatória de características é governada por um hiperparâmetro importante, denominado  $m_{try}$ , que representa o número de características a serem consideradas para a divisão de cada nó. Um valor padrão comumente utilizado é  $\sqrt{d}$ , onde  $d$  é o número total de características. Isso evita que o modelo dependa excessivamente de uma única característica forte, promovendo a robustez do algoritmo.

Após o treinamento, o RF combina as previsões de todas as árvores de decisão individuais. No caso de um problema de classificação, a decisão final é obtida por meio de uma votação majoritária entre as árvores. Em problemas de regressão, a saída final é a média das previsões de todas as árvores. Essa agregação de resultados ajuda a suavizar as previsões e a reduzir o risco de *overfitting*, uma vez que a média ou maioria das previsões tende a ser mais estável do que as previsões individuais. O RF também oferece um método interno de validação chamado *out-of-bag* (OOB). Como cada árvore é treinada com um subconjunto aleatório dos dados,

algumas amostras não são utilizadas durante o treinamento e servem como amostras de teste internas. A estimativa OOB fornece uma medida de desempenho confiável sem a necessidade de separar explicitamente um conjunto de validação. Devido à sua adaptabilidade e facilidade de uso, o RF tem sido implementado com sucesso em diversas áreas, incluindo bioinformática (DÍAZ-URIARTE; ANDRÉS, 2006), finanças (BOU-HAMAD, 2017) e sensoriamento remoto (PAL, 2005). Sua eficácia é especialmente notável quando o número de características é grande ou quando os dados apresentam ruídos e outliers, pois o RF é capaz de isolar automaticamente essas características menos informativas e priorizar as mais relevantes.

### 2.3.4 Máquinas de Vetores de Suporte

Máquinas de Vetores de Suporte (SVM, do inglês “Support Vector Machines”) são uma categoria poderosa de algoritmos de aprendizado de máquina que se destacam tanto em tarefas de classificação quanto de regressão. O principal objetivo das SVMs é encontrar um hiperplano de separação no espaço de características que maximize a margem entre as classes. Esse hiperplano é escolhido de forma que os pontos de dados mais próximos a ele, conhecidos como vetores de suporte, tenham a maior distância possível até o hiperplano (GANDHI, 2018). Isso não apenas resulta em um modelo de classificação robusto, mas também permite que as SVMs lidem eficazmente com conjuntos de dados complexos e casos de separação de classes não-lineares.

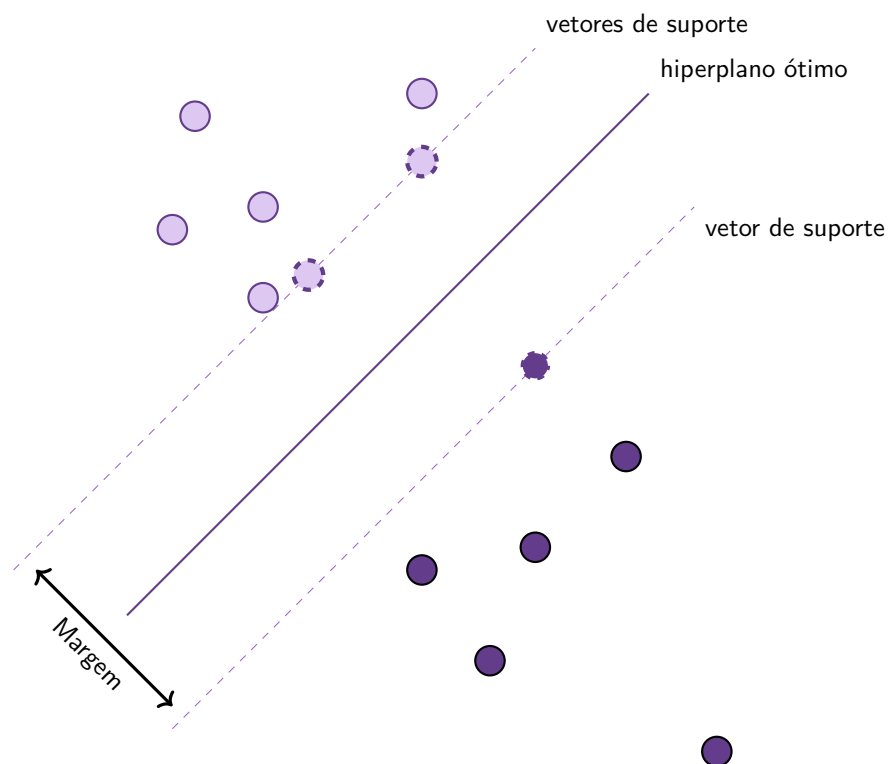


Figura 1 – Ilustração do funcionamento de uma Support Vector Machine (SVM). Autoria própria. Inspirado em (ALCALDE, 2017).

As SVMs são especialmente conhecidas por sua capacidade de separar classes em pro-

blemas de classificação binária, mas também podem ser estendidas para tarefas de classificação multiclasse usando vários hiperplanos. Uma característica distintiva das SVMs é o “kernel trick”, que possibilita o manuseio de dados não-lineares ao mapeá-los para um espaço dimensional superior, onde a separação linear é possível. Alguns dos kernels populares incluem o kernel linear, polinomial e radial (RBF) (XU; ZOMER; BRERETON, 2006). Definimos o RBF como:

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right) \quad (20)$$

onde  $x$  e  $x'$  são os vetores de entrada, e  $\sigma$  é um parâmetro que controla a largura do kernel. O uso de kernels permite que as SVMs lidem eficazmente com dados não-lineares, tornando-as extremamente versáteis em uma variedade de aplicações.

Além disso, as SVMs oferecem um mecanismo de controle para o equilíbrio entre a maximização da margem e a minimização do erro de classificação através do parâmetro de regularização  $C$ . Em uma SVM, o objetivo é encontrar o hiperplano de separação que maximize a margem entre as classes, mas, em situações onde a separabilidade perfeita não é possível, o parâmetro  $C$  permite ajustar a sensibilidade do modelo a erros de classificação. A função objetivo da SVM inclui o termo  $\frac{1}{2}\|\mathbf{w}\|^2 + C \sum_i \xi_i$  onde  $\|\mathbf{w}\|$  representa a norma do vetor de pesos, que está associada à largura da margem, e  $\xi_i$  são as variáveis de erro associadas a cada exemplo. O valor de  $C$  controla o peso do termo de penalização dos erros  $\xi_i$ ; quando  $C$  é pequeno, permite-se uma margem mais larga ao custo de algumas classificações incorretas. Em contraste, valores altos de  $C$  restringem a margem e forçam o modelo a classificar cada ponto corretamente, potencialmente resultando em *overfitting*. Essa versatilidade na escolha do parâmetro  $C$  torna as SVMs uma ferramenta poderosa para trabalhar com dados de alta dimensionalidade, proporcionando uma boa capacidade de lidar com dados não-lineares, interpretabilidade e controle sobre a regularização.

A seguir, o funcionamento do algoritmo pode ser resumido nos seguintes passos:

1. Representação dos dados: Os dados de entrada são organizados em um espaço de características, onde cada ponto representa um exemplo do conjunto de dados, rotulado conforme sua classe.
2. Busca pelo hiperplano de separação: O algoritmo procura identificar um *hiperplano* que separe as classes da melhor forma possível. A melhor separação é aquela que *maximiza a margem*, ou seja, a distância entre o hiperplano e os pontos mais próximos de cada classe.
3. Identificação dos vetores de suporte: Os pontos mais próximos ao hiperplano de separação são chamados de *vetores de suporte*. Eles são fundamentais para a definição da margem e do próprio hiperplano — sem eles, a posição ótima da fronteira não seria determinada.

4. Cálculo da margem e da função objetivo: A SVM busca *maximizar a margem* enquanto *minimiza os erros de classificação*. Isso é feito resolvendo uma função de otimização que inclui:

$$\frac{1}{2} \|\mathbf{w}\|^2 + C \sum_i \xi_i \quad (21)$$

5. Aplicação do kernel (quando necessário): Quando os dados não são linearmente separáveis, aplica-se o *truque do kernel* para mapear os dados para um espaço de maior dimensionalidade, onde uma separação linear seja possível. Exemplos de kernels incluem: *linear*, *polinomial* e *radial (RBF)*.
6. Treinamento da SVM: A partir da otimização e aplicação (ou não) de kernels, o modelo é treinado para encontrar o *hiperplano ótimo* que melhor separa as classes no conjunto de treinamento.
7. Classificação de novos dados: Após o treinamento, a SVM pode ser usada para *classificar novos pontos* de dados, verificando de que lado do hiperplano eles se encontram.

### 2.3.5 Análise do Componente Principal

A Análise do Componente Principal (PCA, do inglês "*Principal Component Analysis*"), é uma técnica fundamental de redução de dimensionalidade amplamente utilizada no campo do aprendizado de máquina e análise de dados. Seu principal objetivo é simplificar conjuntos de dados complexos, preservando informações essenciais e reduzindo a dimensionalidade. Isso facilita a visualização e interpretação dos dados. O PCA busca encontrar novos eixos, chamados componentes principais, nos quais os dados apresentam a maior variância. Esses componentes principais são ortogonais entre si, o que simplifica a interpretação e a análise dos dados (BRO; SMILDE, 2014).

Matematicamente, dado um conjunto de dados com  $n$  amostras e  $p$  variáveis, representado por uma matriz  $X$  de dimensão  $n \times p$ , o PCA realiza uma transformação linear dos dados para encontrar uma nova base onde os dados são projetados. A primeira etapa consiste em centralizar os dados subtraindo a média de cada variável de cada entrada de  $X$ , de forma que cada coluna da matriz tenha média zero. Em seguida, calcula-se a matriz de covariância dos dados centralizados, dada por

$$\mathbf{C} = \frac{1}{n-1} X^T X. \quad (22)$$

Para encontrar os componentes principais, resolvem-se os autovalores  $\lambda_i$  e autovetores  $\mathbf{v}_i$  da matriz de covariância  $\mathbf{C}$  através da equação:

$$\mathbf{C}\mathbf{v}_i = \lambda_i\mathbf{v}_i, \quad (23)$$

onde  $\lambda_i$  representa a variância explicada pelo  $i$ -ésimo componente principal e  $\mathbf{v}_i$  é o vetor direcional correspondente. Esses autovetores  $\mathbf{v}_i$  formam a nova base, e os dados originais são projetados nesses componentes principais. Para reduzir a dimensionalidade, seleciona-se apenas os componentes principais associados aos maiores valores de  $\lambda_i$ , que retêm a maior variância dos dados originais. Embora reduza a dimensionalidade, o PCA tenta preservar a maior quantidade possível de variância nos dados originais, mantendo as informações mais significativas (RINGNÉR, 2008). Além disso, o PCA é uma ferramenta valiosa na análise exploratória de dados e pode ser usado como um passo inicial na análise de dados antes de aplicar outros algoritmos de aprendizado de máquina.

## 2.4 Redes Adversariais Generativas (GANs)

Redes Adversariais Generativas (GANs) são frameworks de aprendizado de máquina para gerar dados sintéticos semelhantes a conjuntos de dados reais. Introduzidas por Goodfellow et al. em 2014 (GOODFELLOW et al., 2014), as GANs consistem em um gerador  $G$  e um discriminador  $D$  treinados conjuntamente por meio de um processo adversarial. O gerador mapeia um vetor de ruído latente  $\vec{z}$  de uma distribuição anterior  $p_z(\vec{z})$  para o espaço de dados, com o objetivo de produzir dados indistinguíveis dos dados reais. O discriminador  $D$  é um classificador binário que distingue entre dados reais  $\vec{x}$  da distribuição real  $p_{\text{data}}(\vec{x})$  e dados sintéticos gerados por  $G$ . Ele fornece uma probabilidade  $D(\vec{x})$  de que os dados sejam reais. O objetivo de treinamento das GANs é um jogo de minimax entre  $G$  e  $D$ , definido pela função de valor  $V(D, G)$  (GOODFELLOW et al., 2014):

$$\min_G \max_D V(D, G) = \mathbb{E}_{\vec{x} \sim p_{\text{data}}(\vec{x})} [\log D(\vec{x})] + \mathbb{E}_{\vec{z} \sim p_z(\vec{z})} [\log (1 - D(G(\vec{z})))] . \quad (24)$$

O gerador visa maximizar a probabilidade de o discriminador classificar incorretamente suas saídas como reais. Enquanto isso, o discriminador busca identificar corretamente amostras reais e falsas, minimizando a verossimilhança negativa das classificações corretas. As GANs são treinadas iterativamente, com ambas as redes melhorando continuamente: o gerador em produzir dados realistas e o discriminador em distinguir entre dados reais e sintéticos.

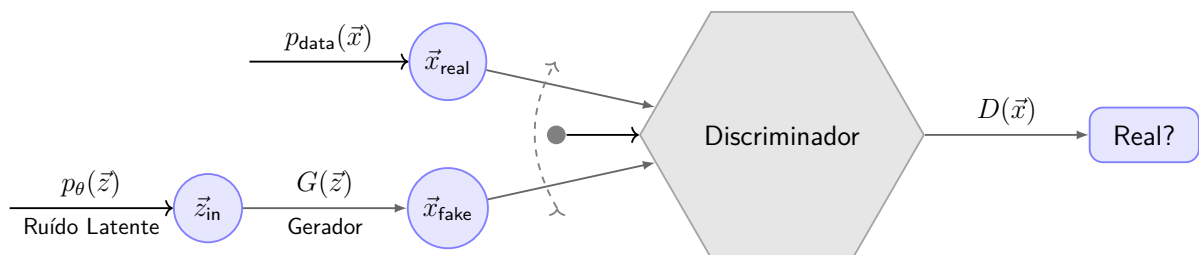


Figura 2 – Estrutura básica de uma Rede Adversarial Generativa (GAN). O gerador cria dados sintéticos, que são avaliados pelo discriminador. O feedback do discriminador melhora o desempenho do gerador em iterações subsequentes.

Dados sintéticos de GANs são comumente utilizados para aumento de dados em processamento de imagens (STÖCKL, 2022; SHORTEN; KHOSHGOFTAAR, 2019). Modelos como o DALL-E da OpenAI utilizam imagens sintéticas em seus conjuntos de treinamento (RAMESH et al., 2021), embora o DALL-E seja um modelo baseado em *transformers*, e não uma GAN.

Nesta tese, iremos mostrar na seção 5.2 que exploramos se a adição de amostras geradas por GAN ao nosso conjunto de treinamento de fases quânticas poderia melhorar o desempenho preditivo dos nossos modelos de aprendizado de máquina, especialmente com conjuntos de dados limitados. Ao aumentar com amostras sintéticas, visamos aprimorar a generalização e a precisão do modelo na classificação de fases quânticas, abordando a questão de pequenos conjuntos de dados ao fornecer dados adicionais e reduzir a dependência de grandes conjuntos de dados reais.

### 2.4.1 Conditional Tabular GANs (CTGANs)

*Conditional Tabular GANs* (CTGANs) são um tipo de Rede Adversarial Generativa (GAN) projetada para dados tabulares sintéticos (XU et al., 2019). Enquanto as GANs tradicionais se destacam com dados de imagem ou áudio, elas enfrentam dificuldades com dados tabulares devido a tipos de dados mistos, distribuições desequilibradas e à necessidade de preservar relações entre características (WANG et al., 2024). As CTGANs abordam esses desafios condicionando em valores específicos de características, possibilitando a geração de dados sintéticos que espelham a distribuição original. Isso é crucial para características categóricas, garantindo a fidelidade estatística. A arquitetura da CTGAN, assim como as GANs tradicionais, inclui um gerador e um discriminador, mas adiciona normalização específica por modo e treinamento por amostragem para lidar com as complexidades dos dados tabulares (XU et al., 2019).

## 2.5 Explainable AI

A Inteligência Artificial Explicável (XAI, do inglês *Explainable Artificial Intelligence*) refere-se à capacidade de sistemas de inteligência artificial (IA) fornecerem explicações compreensíveis para suas decisões ou previsões (ALI et al., 2023). Em uma era em que a IA está cada vez mais integrada em diversos aspectos da sociedade, a necessidade de transparência e responsabilidade nos sistemas de IA tornou-se primordial. A XAI busca resolver a natureza de “caixa preta” de muitos modelos de IA, que tomam decisões sem explicações claras, sem fornecer insights sobre como essas decisões são tomadas (HOFFMAN et al., 2019). A necessidade de XAI surge da complexidade dos modelos de IA modernos, especialmente redes neurais profundas e algoritmos complexos de aprendizado de máquina. Embora esses modelos muitas vezes alcancem altos níveis de desempenho, eles carecem de transparência, o que dificulta a

confiança dos usuários em suas decisões, especialmente em domínios críticos, como a saúde (SARASWAT et al., 2022) e finanças (MISHEVA et al., 2021).

Os princípios-chave da XAI incluem interpretabilidade, transparência, justificabilidade e responsabilidade. A interpretabilidade envolve garantir que os modelos de IA sejam projetados de forma que os humanos possam entender como as entradas são transformadas em saídas. A transparência refere-se a tornar os processos de tomada de decisão dos modelos de IA transparentes, permitindo que os usuários compreendam as razões por trás de uma decisão ou previsão específica. A justificabilidade envolve fornecer justificativas claras e concisas para as decisões da IA, explicando como as evidências foram usadas para chegar a uma conclusão. Por fim, a responsabilidade exige que os desenvolvedores e operadores de sistemas de IA garantam que esses sistemas sejam explicáveis e compreensíveis, e que possam ser responsabilizados pelas decisões tomadas por esses sistemas (Barredo Arrieta et al., 2020). Diversas técnicas foram desenvolvidas para melhorar a explicabilidade dos sistemas de IA, incluindo métodos de interpretabilidade de modelos, como visualização da importância de características (CASALICCHIO; MOLNAR; BISCHL, 2019), análise de sensibilidade (STEIN et al., 2022) e geração de explicações pós-hoc (MADSEN; REDDY; CHANDAR, 2022). No entanto, persistem desafios na criação de técnicas de explicabilidade escaláveis e eficazes, equilibrando a relação entre desempenho e interpretabilidade (BELL et al., 2022), e garantindo justiça e imparcialidade nas explicações geradas pela IA.

Em nossas análises, utilizamos o algoritmo *Shapley Additive Explanations* (SHAP). O SHAP é uma técnica baseada na teoria dos jogos cooperativos que atribui a cada característica de uma previsão um valor de importância, representando sua contribuição para a saída do modelo. a seguir descrevemos em mais detalhes o método.

### 2.5.1 *Shapley Additive Explanations*

*Shapley Additive Explanations* é uma abordagem no campo da Inteligência Artificial Explicável que fornece uma estrutura unificada para interpretar a saída de modelos de aprendizado de máquina (LUNDBERG; LEE, 2017). Baseado na teoria dos jogos cooperativos, o SHAP atribui a cada característica em uma previsão um valor de importância que representa sua contribuição para a saída do modelo (MERRICK; TALY, 2020). A ideia central do SHAP é decompor a previsão do modelo em contribuições de características individuais, permitindo uma compreensão mais granular de como cada característica influencia o resultado final. A importância do SHAP reside em sua capacidade de fornecer explicações locais e globais para as previsões do modelo. As explicações locais focam em explicar a previsão de uma única instância, atribuindo a contribuição de cada característica àquela previsão específica. Já as explicações globais fornecem uma visão geral da importância das características em todo o conjunto de dados, permitindo que os usuários entendam quais características têm o maior impacto no comportamento geral do modelo (LUNDBERG et al., 2020).

Em termos práticos, o SHAP produz visualizações como gráficos de resumo SHAP, gráficos de dependência SHAP e gráficos de valores individuais SHAP para ajudar os usuários a interpretar as previsões dos modelos (LUNDBERG; ERION; LEE, 2018). Essas visualizações fornecem insights sobre como cada característica afeta a saída do modelo e podem auxiliar na depuração de modelos, compreensão do comportamento do modelo e aumento da confiança nos sistemas de IA. Apesar de sua eficácia, o SHAP não está isento de limitações. Pode ser computacionalmente caro, especialmente para grandes conjuntos de dados e modelos complexos, e o SHAP não aborda diretamente o problema da justiça (HICKEY; STEFANO; VASILEIOU, 2021). Além disso, interpretar a saída do SHAP requer algum entendimento do modelo subjacente de aprendizado de máquina e pode não fornecer explicações intuitivas para usuários não técnicos.

O SHAP quantifica quanto cada característica adiciona ou subtrai do valor previsto, fornecendo uma atribuição justa de importância entre as características. No SHAP, essa atribuição é representada como um modelo de atribuição aditiva de características, expresso pela seguinte equação linear (LUNDBERG; LEE, 2017):

$$g(z') = \phi_0 + \sum_{j=1}^M \phi_j z'_j \quad (25)$$

onde  $g(z')$  representa o modelo de explicação, que aproxima a saída do modelo original de forma interpretável e transparente com base nas contribuições individuais das características. Nesta equação,  $\phi_0$  é a previsão base do modelo, ou seja, a saída esperada na ausência de qualquer característica; em outras palavras,  $\phi_0$  é o valor médio de todas as previsões no conjunto de dados. O parâmetro  $M$  denota o número total de características no modelo, e cada  $\phi_j$  representa o valor de Shapley para a característica  $j$ , indicando sua contribuição específica para a previsão. Finalmente,  $z'_j$  é um indicador binário para a característica  $j$ , com 1 indicando sua presença e 0 indicando sua ausência.

O modelo de explicação  $g(z')$  é projetado para fornecer uma aproximação linear do comportamento do modelo de aprendizado para uma entrada específica. Essencialmente, a ideia é substituir o vetor original de características  $x$  por uma sequência binária,  $x'$ , que representa a presença ou ausência de cada característica. Essa entrada simplificada  $x'$  é então mapeada de volta ao espaço de entrada original por meio de uma função  $h_x(x')$ , específica para cada entrada  $x$ . Na prática, quando  $x'$  é uma sequência completa de uns (indicando que todas as características estão presentes), o mapeamento assegura que  $f(x) = f(h_x(x'))$ , onde  $f(x)$  representa a função preditiva do modelo original. O modelo de explicação  $g(z')$  é construído de forma que  $g(z') \approx f(h_x(z'))$  sempre que  $z' \approx x'$ , sendo  $z'$  uma sequência binária que representa o subconjunto de características selecionadas para a aproximação.

Muitos métodos explicáveis seguem o modelo de atribuição aditiva de características, no qual a previsão é decomposta em contribuições lineares de cada característica (RIBEIRO; SINGH; GUESTRIN, 2016; SHRIKUMAR; GREENSIDE; KUNDAJE, 2017; BACH et al., 2015),

com a principal diferença estando na forma como a contribuição de cada característica,  $\phi_j$ , é calculada. A inovação do SHAP reside, de fato, na maneira rigorosa e unificada como ele calcula os valores  $\phi_j$  para cada característica. Esses valores, chamados de valores de Shapley, são calculados para garantir uma atribuição justa e única de importância para cada característica, obedecendo a três propriedades desejáveis: precisão local, ausência de influência e consistência. Formalmente, os valores de Shapley  $\phi_j$  são calculados como a média das contribuições marginais de uma característica em todos os subconjuntos possíveis de características. Assim, dado um conjunto de características, o valor de Shapley de cada característica reflete a contribuição média da característica para o resultado em diferentes “coalizões” ou subconjuntos de características.

O valor clássico de Shapley é definido com base na teoria dos jogos cooperativos e é dado pela equação (LUNDBERG; LEE, 2017):

$$\phi_j = \sum_{S \subseteq N \setminus \{j\}} \frac{|S|!(|N| - |S| - 1)!}{|N|!} (v(S \cup \{j\}) - v(S)) \quad (26)$$

onde  $S$  representa um subconjunto de características sem a característica  $j$ ,  $N$  é o conjunto total de características,  $v(S)$  é o valor da função (ou previsão do modelo) para o conjunto  $S$ , e os fatores combinatórios ponderam as contribuições de cada subconjunto, garantindo que todas as combinações sejam consideradas. Apesar da solução formal fornecida pela Eq. (26), calcular os valores de Shapley exatos avaliando todas as combinações possíveis de características é computacionalmente caro. Assim, uma aproximação prática, introduzida em (LUNDBERG; LEE, 2017), é o método Kernel SHAP, que utilizamos em nossa análise. O Kernel SHAP combina ideias do método LIME (RIBEIRO; SINGH; GUESTRIN, 2016) com os valores clássicos de Shapley, proporcionando uma aproximação eficiente que evita a necessidade de avaliar todas as permutações de subconjuntos de características. Esse método ajuda a identificar a influência de características individuais nas previsões do modelo, permitindo compreender a importância relativa de cada característica.

## 2.6 Tendências Futuras

O campo do aprendizado de máquina está em constante evolução, e várias tendências emocionantes estão moldando seu futuro. Isso inclui o aprofundamento do aprendizado profundo para uma ampla variedade de tarefas, como visão computacional, processamento de linguagem natural e robótica (WIRIYATHAMMABHUM et al., 2016). O aprendizado auto-supervisionado (*Self-Supervised Learning* - SSL) também está ganhando destaque, possibilitando que modelos aprendam com dados não rotulados, o que é crucial em cenários com escassez de dados rotulados (RANI et al., 2023). Além disso, a interpretabilidade continuará sendo uma prioridade (BUSUIOC, 2021). O foco na ética e regulamentação da IA deve aumentar, com diretrizes para garantir o uso responsável dessa tecnologia, e a integração da IA em setores diversos, como saúde e agricultura (SMITH, 2018), promete trazer avanços significativos. Por fim, vale

mencionar a computação quântica como grande promessa para revolucionar o aprendizado de máquina, através do aprendizado quântico (BIAMONTE et al., 2017). Neste século a humanidade tem o desafio de utilizar essas tendências para moldar o futuro da IA de forma a causar um impacto positivo na sociedade.

# 3 Informação Quântica

Este capítulo explora os fundamentos e conceitos essenciais Em Ciência da Informação Quântica (em inglês, *Quantum Information Science - QIS*) necessários para a compreensão das aplicações práticas que serão desenvolvidas nos capítulos seguintes da tese. Iremos iniciar com uma breve introdução histórica que destaca os marcos na evolução do entendimento da informação quântica. Exploraremos em seguida o conceito de superposição, onde partículas quânticas podem existir em múltiplos estados simultaneamente, e o qubit, a unidade básica de informação quântica. Além disso, apresentaremos a Esfera de Bloch e as bases que fornecem uma representação geométrica útil para um qubit. Abordaremos também operações em um qubit, o operador densidade, observáveis e medidas, incluindo estados puros e estados mistos, população, coerência e o Hamiltoniano. Finalmente, examinaremos sistemas quânticos abertos e a evolução temporal desses sistemas, incluindo a importante Equação Mestra de Lindblad.

## 3.1 Uma Breve Introdução Histórica

A história da informação quântica é uma narrativa interdisciplinar que entrelaça os campos da teoria da informação, da computação e do desenvolvimento da física, com foco especial na teoria quântica. Neste capítulo, exploraremos marcos históricos cronologicamente, destacando eventos significativos em diversas disciplinas que contribuíram para viabilizar essa teoria de forma abrangente.

No início do século XX, os físicos depararam-se com desafios significativos na explicação de fenômenos observados experimentalmente. Um exemplo clássico é a não confirmação experimental da existência da "catástrofe do ultravioleta" (EISBERG; RESNICK, 1985). Esta era uma contradição aparente entre as previsões da física clássica e os resultados experimentais relacionados à distribuição espectral da radiação emitida por um corpo negro (EISBERG; RESNICK, 1985). De acordo com a teoria clássica da física, o espectro de radiação de um corpo negro deveria ser contínuo e aumentar infinitamente em intensidade à medida que a frequência da radiação se aproximasse do violeta na faixa do espectro eletromagnético. No entanto, experimentos precisos revelaram que a intensidade da radiação diminuía abruptamente para frequências mais altas, o que contradizia as expectativas (EISBERG; RESNICK, 1985). Essa discrepância levou os físicos, especialmente a partir do início dos anos 1920, a desenvolver uma nova teoria que posteriormente seria conhecida como teoria quântica. Essa teoria não apenas resolveu a "catástrofe do ultravioleta", explicando corretamente a distribuição espectral do corpo negro, mas também se tornou uma estrutura fundamental em diversas áreas da ciência, incluindo o entendimento de supercondutores, a fusão nuclear nas estrelas, a estrutura do DNA e muitos outros fenômenos (NIELSEN; CHUANG, 2000).

Já a ciência da computação, deve muito ao trabalho inovador de Alan Turing em 1936, que introduziu o conceito de um computador programável por meio do seu modelo de máquina, que posteriormente ficou conhecido em sua homenagem como Máquina de Turing (TURING et al., 1936). Ele demonstrou a existência de uma Máquina de Turing Universal capaz de simular qualquer outra Máquina de Turing, afirmando que essa abrangia a essência da computação algorítmica, uma noção conhecida como a tese de Church-Turing, que equipara a capacidade física de executar algoritmos em um dispositivo com o conceito matemático de uma Máquina de Turing Universal. Essa tese lançou as bases para o desenvolvimento da ciência da computação. Pouco depois do trabalho de Turing, surgiram os computadores eletrônicos, e John von Neumann desenvolveu um modelo prático para reunir os componentes necessários para alcançar as capacidades de uma Máquina de Turing Universal. A invenção do transistor por John Bardeen, Walter Brattain e Will Shockley em 1947 marcou um momento crucial, impulsionando o desenvolvimento do hardware de computadores. Ao longo do tempo, o hardware de computadores experimentou um crescimento exponencial, observado na lei de Moore em 1965 (NIELSEN; CHUANG, 2000). A Lei de Moore, proposta por Gordon Moore em 1965 (MOORE, 1965), observou que o número de transistores em um chip dobraria aproximadamente a cada dois anos, impulsionando o aumento da capacidade de processamento. Essa previsão tornou-se um princípio orientador para a evolução da tecnologia de semicondutores.

Na década de 1960 e início dos anos 1970, desafios surgiram para a tese Church-Turing, como a computação analógica, que inicialmente parecia contornar essa tese, mas foi posteriormente limitada pela presença de ruído. Além disso, a descoberta de algoritmos probabilísticos, como o teste de primaridade de Solovay-Strassen (HIDARY, 2021), levantou dúvidas sobre a tese, sugerindo que existem problemas que podem ser eficientemente resolvidos por algoritmos probabilísticos, mas não por máquinas de Turing determinísticas. Isso motivou a consideração de uma modificação da tese, permitindo máquinas de Turing probabilísticas. Por fim, a busca por uma base sólida para a tese levou cientistas como David Deutsch ao desenvolvimento de modelos de computação baseados em princípios da mecânica quântica (NIELSEN; CHUANG, 2000). A primeira noção de um computador quântico geralmente é atribuída a Richard Feynman, entretanto, alguns cientistas, tais como Paul Benioff em um artigo de 1979 já apresentavam bases teóricas sob as quais a computação quântica poderia evoluir. Em 1980, o russo Yuri Manin também apresentou algumas ideias centrais em seu livro, que foi traduzido muitos anos após. Apenas em 1981 Feynman em uma aula intitulada "Simulando a Física com Computadores", argumentou que a natureza é quântica, logo, se queremos simulá-la, devemos fazer quanticamente (na época não era claro para Feynman, ou para a comunidade em geral como se construir tal máquina) (HIDARY, 2021).

Na década de 1980 a criptografia quântica, incluindo o protocolo BB84 proposto por Charles Bennett e Gilles Brassard (BENNETT; BRASSARD, 2014), começou a ganhar destaque nesse período, estabelecendo as bases para a segurança da informação quântica. Já na década de 1990, houve um grande aumento de interesse na área, sendo um marco crucial a descoberta

do algoritmo de Shor por Peter Shor em 1994 (SHOR, 1994; SHOR, 1999). Este demonstrou a superioridade da computação quântica ao resolver eficientemente um problema desafiador para computadores clássicos: a fatoração de números inteiros grandes em seus números primos constituintes. A importância desse feito reside no fato de que a fatoração de números inteiros grandes desempenha um papel fundamental na quebra de criptografia de chave pública, como o RSA. Em 1995, Lov Grover apresentou mais evidências do poder dos computadores quânticos ao demonstrar que outro problema importante, a busca em um espaço não estruturado, também poderia ser acelerado usando um computador quântico (GROVER, 1996). Embora o algoritmo de Grover não tenha proporcionado um *speed-up* (aumento na velocidade, calculada como o número de passos computacionais realizados para resolver o problema) tão espetacular quanto o algoritmo de Shor (que era exponencial, frente ao *speed-up* quadrático de Grover), a ampla aplicabilidade de metodologias baseadas em busca gerou interesse em seu algoritmo (NIELSEN; CHUANG, 2000). Em conjunto essas descobertas causaram um grande interesse em pesquisa, “aquecendo” a área de informação e computação quântica.

No século XXI, a informação quântica continuou a se expandir, com avanços notáveis em áreas como criptografia quântica, comunicação quântica segura e computação quântica. Atualmente, os computadores quânticos já são uma realidade, e é possível acessá-los e realizar computações por meio de acesso remoto a dispositivos de diferentes empresas, como a IBM (SANTOS, 2016). No entanto, ainda nos encontramos em uma era na qual os maiores computadores quânticos possuem apenas algumas dezenas ou centenas de *qubits* (unidade de informação básica da computação quântica), e estes ainda são muito suscetíveis ao ruído, o que caracteriza essa era como a NISQ (*Noise Intermediate-Scale Quantum*), ou seja, computadores de escala média e ruidosos (PRESKILL, 2018). A história da informação quântica é uma narrativa contínua de descobertas científicas e inovações tecnológicas que prometem revolucionar a forma como processamos, transmitimos e protegemos informações. No entanto, ainda há muito trabalho a ser feito até que os computadores quânticos estejam à altura de suas promessas e tenham um impacto significativo no dia a dia da sociedade.

## 3.2 Conceitos Fundamentais

Nesta seção, serão explorados conceitos fundamentais da informação quântica, proporcionando a compreensão dos princípios subjacentes à mecânica quântica e à manipulação de informações em sistemas quânticos. Iniciamos com a notação de Dirac, que nos permite expor conceitos quânticos de forma eficiente, seguida por uma discussão sobre o conceito de superposição, que permite que um sistema quântico coexista em múltiplos estados simultaneamente. A seguir, examina-se o qubit, que é a unidade básica de informação quântica, e as operações aplicáveis a ele. O tópico seguinte é o operador densidade, que descreve a densidade de probabilidade do estado quântico. Além disso, são abordadas medidas em sistemas quânticos, englobando estados puros, estados mistos e os conceitos de população e coerência. Em seguida,

é explorado o Hamiltoniano, que descreve a energia total de um sistema quântico e desempenha um papel crucial na evolução temporal dos estados quânticos. Por fim, consideraremos sistemas quânticos abertos, nos quais o sistema interage com o ambiente, levando à perda de coerência e dissipação. Nesse contexto, abordaremos a equação mestra de Lindblad, uma ferramenta importante, que nos ajuda a descrever a dinâmica desses sistemas. Esses conceitos fundamentais são essenciais para o entendimento dos problemas a serem resolvidos pelos algoritmos de aprendizado de máquina presentes nesta tese.

### 3.2.1 Notação de Dirac

A notação de Dirac é uma notação matemática amplamente utilizada na física quântica para descrever estados e operadores quânticos de maneira concisa. Foi desenvolvida pelo físico teórico Paul Dirac e é especialmente útil na representação de estados de superposição e operadores lineares (DIRAC, 1939). A notação de Dirac é baseada em bras e kets, os primeiros são análogos a vetores coluna, enquanto os segundos são seus conjugados transpostos (EZHOV; VENTURA, 2000). Em notação de Dirac, um bra é representado como  $\langle\psi|$ , onde  $\psi$  é o nome do estado ou vetor. Por exemplo, um bra que representa o estado  $\psi_1$  seria denotado como  $\langle\psi_1|$ . Por outro lado, um ket é representado como  $|\phi\rangle$ , onde  $\phi$  é o nome do estado ou vetor. Por exemplo, um ket que representa o estado  $\phi_2$  seria denotado como  $|\phi_2\rangle$ . A seguir serão indicadas algumas operações importantes e em alguns casos, a sua representação na notação de Dirac.

A notação  $z^*$  representa o conjugado complexo do número complexo  $z$ , ou seja,  $(1 + i)^* = (1 - i)$ . O produto interno entre dois vetores é denotado como  $\langle\phi|\psi\rangle$ . A notação  $|\phi\rangle \otimes |\psi\rangle$  representa o produto tensorial entre  $|\phi\rangle$  e  $|\psi\rangle$ , enquanto  $|\phi\rangle|\psi\rangle$  é uma forma abreviada de representar o mesmo produto tensorial. Para matrizes,  $A^*$  indica o conjugado complexo de  $A$ ,  $A^T$  é a transposição de  $A$ , e  $A^\dagger$  representa o transposto complexo conjugado de  $A$ , (para uma matriz ou operador  $A$  ser hermitiano, sua matriz transposta ( $A^T$ ) deve ser igual à conjugada complexa da matriz original, e isso deve ocorrer para todos os elementos da matriz). Finalmente, o produto interno entre um vetor  $|\phi\rangle$  e o operador  $A|\psi\rangle$  é dado por  $\langle\phi|A|\psi\rangle$ , o qual pode também ser expresso como o produto interno entre  $A^\dagger|\phi\rangle$  e  $|\psi\rangle$ . Essas notações são fundamentais na descrição matemática e operações em sistemas quânticos (NIELSEN; CHUANG, 2000).

### 3.2.2 Superposição

A superposição é um conceito fundamental na mecânica quântica e surge a partir das propriedades intrínsecas das partículas subatômicas, como elétrons, fótons e átomos. Mais especificamente, ela tem sua origem na dualidade onda-partícula das partículas subatômicas, que podem se comportar tanto como partículas pontuais quanto como ondas propagantes (ARNOLD et al., 1999). Em vez de seguir trajetórias claramente definidas, essas partículas são descritas por

funções de onda que contêm informações sobre todas as possíveis configurações do sistema. O princípio da superposição estabelece que, até o momento da medida, o estado quântico de uma partícula é uma combinação linear de todos os estados possíveis, representada por coeficientes complexos chamados amplitudes de probabilidade. Quando uma medida é efetuada, a função de onda colapsa para um estado específico, determinado pelas probabilidades associadas. A superposição é uma característica intrínseca das partículas subatômicas e é evidenciada experimentalmente por fenômenos como o experimento de fenda dupla e a interferência quântica (EISBERG; RESNICK, 1985).

Matematicamente, a soma da norma ao quadrado dos coeficientes complexos  $\alpha_i$  deve satisfazer

$$\sum_i |\alpha_i|^2 = 1 \quad (27)$$

de modo que  $|\alpha_i|^2$  represente a probabilidade de encontrar o sistema quântico no estado de base  $|\phi_i\rangle$ .

A superposição pode então ser expressa da seguinte forma (ARRUDA, 2011):

$$|\psi\rangle = \sum_i \alpha_i |\phi_i\rangle \quad (28)$$

onde  $|\psi\rangle$  é o estado quântico superposto,  $\alpha_i$  são coeficientes complexos que determinam a amplitude e a fase de cada componente, e  $|\phi_i\rangle$  são os estados de base do sistema quântico, correspondendo aos possíveis resultados de uma medida.

### 3.2.3 O Qubit

O qubit, abreviação de "quantum bit", é a unidade fundamental de informação quântica e desempenha um papel central na computação quântica. O conceito surgiu como uma contrapartida clássica da unidade fundamental da ciência da informação, o *bit*. *Bits* clássicos podem estar em estados 0 ou 1, podendo ser imaginados como uma lâmpada acesa ou apagada (ARRUDA, 2011). A mecânica quântica introduziu a ideia de que partículas subatômicas podem existir em estados de superposição, ou seja, em combinações lineares de múltiplos estados simultaneamente. Logo, a origem do qubit está relacionada com a busca por um equivalente que poderia ser usado para processamento de informações em sistemas quânticos.

Um qubit pode ser representado matematicamente como um vetor em um espaço de Hilbert complexo (espaço vetorial complexo, munido de produto escalar) de duas dimensões ( $\mathbb{C}^2$ ), o que significa que ele pode existir em uma combinação linear de dois estados básicos, geralmente referidos como  $|0\rangle$  e  $|1\rangle$ . Essa representação é dada pela seguinte equação (NIELSEN; CHUANG, 2000):

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \quad (29)$$

onde  $|\psi\rangle$  é o estado do qubit.  $\alpha$  e  $\beta$  são coeficientes complexos que determinam a amplitude e a fase de cada componente da superposição.  $|0\rangle$  e  $|1\rangle$  são os estados de base do qubit, representando os estados 0 e 1, respectivamente.

Os qubits, podem ainda ser representados de forma eficaz por meio de matrizes coluna. Essa representação é especialmente útil para expressar estados quânticos e operadores em sistemas de qubits. Um qubit no espaço de Hilbert complexo  $\mathbb{C}^2$  é comumente representado como uma matriz coluna, onde cada elemento da matriz representa a probabilidade de encontrar o qubit em um determinado estado (ARRUDA, 2011). Por exemplo, um qubit no estado  $|0\rangle$  pode ser representado como:

$$|0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad (30)$$

enquanto um qubit no estado  $|1\rangle$  é representado como:

$$|1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (31)$$

Além disso, a representação de estados de superposição, como  $\alpha|0\rangle + \beta|1\rangle$ , também pode ser expressa como uma matriz coluna:

$$\alpha|0\rangle + \beta|1\rangle = \begin{bmatrix} \alpha \\ \beta \end{bmatrix}. \quad (32)$$

Essa notação de matriz coluna facilita o cálculo de transformações de estados e operações em qubits, tornando-a uma ferramenta valiosa na representação e manipulação de informações quânticas. Assim como na superposição, a norma dos ao quadrado coeficientes  $\alpha$  e  $\beta$  deve ser igual a 1 para garantir que a probabilidade total de encontrar o qubit em um estado específico seja 1 (NIELSEN; CHUANG, 2000):

$$|\alpha|^2 + |\beta|^2 = 1 \quad (33)$$

Essa propriedade do qubit permite que ele represente informações de forma muito mais rica e complexa do que um bit clássico, sendo um componente essencial para explorar a computação quântica.

### 3.2.4 Esfera de Bloch e bases para um Qubit

A esfera de Bloch é uma representação gráfica fundamental na teoria da informação quântica. Essa esfera unitária representa todos os (infinitos) estados possíveis de um qubit (NIELSEN; CHUANG, 2000), onde os pontos na superfície da esfera correspondem a estados puros. Um estado puro é um estado quântico se possui uma certeza completa sobre as propriedades do sistema, ou seja, ele pode ser descrito por um único vetor de estado. Por outro lado, os estados mistos são representados por pontos dentro da esfera de Bloch, ou seja,

no seu interior. Eles correspondem a sistemas quânticos nos quais há presença de incertezas ou misturas probabilísticas. Esses estados são representados por matrizes densidade, que são combinações de estados puros ponderados por coeficientes probabilísticos (LANDI, 2018). A esfera de Bloch utiliza três eixos principais,  $X$ ,  $Y$ , e  $Z$ , para representar estados-chave, como superposições e fases complexas. Isso facilita a visualização e a compreensão das transformações quânticas e do emaranhamento.

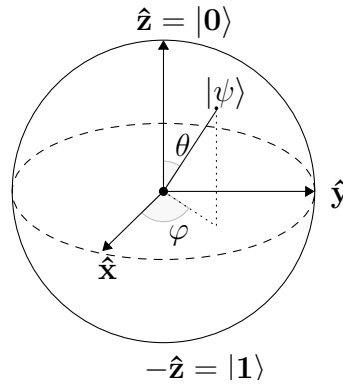


Figura 3 – Representação gráfica da esfera de Bloch

O estado de um qubit  $|\psi\rangle$  pode ser expresso como uma combinação linear de estados de base  $|0\rangle$  e  $|1\rangle$ , onde  $\theta$ ,  $\phi$ , e  $\gamma$  são números reais (NIELSEN; CHUANG, 2000):

$$|\psi\rangle = e^{i\gamma} \left( \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle \right). \quad (34)$$

O fator  $e^{i\gamma}$  pode ser ignorado, pois não tem efeitos observáveis. Portanto, podemos efetivamente escrever o estado como:

$$|\psi\rangle = \cos \left( \frac{\theta}{2} \right) |0\rangle + e^{i\phi} \sin \left( \frac{\theta}{2} \right) |1\rangle. \quad (35)$$

Os números  $\theta$  e  $\phi$  definem um ponto na Esfera de Bloch. Os estados  $|+\rangle$  e  $|-\rangle$  são superposições equilibradas correspondem aos autovetores de  $\sigma_x$ , ou seja, à base de medida ao longo do eixo  $X$  na esfera de Bloch:

$$|+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad (36)$$

$$|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (37)$$

Os estados  $|R\rangle$  e  $|L\rangle$  correspondem aos autovetores de  $\sigma_y$ , ou seja, à base de medida ao longo do eixo  $Y$  na esfera de Bloch.

$$|R\rangle = \frac{1}{\sqrt{2}}(|0\rangle + i|1\rangle), \quad (38)$$

$$|L\rangle = \frac{1}{\sqrt{2}}(|0\rangle - i|1\rangle). \quad (39)$$

Em conjunto, esses estados são essenciais para entender a manipulação e a representação geométrica dos qubits em informação quântica, permitindo a visualização de rotações e transformações nos estados quânticos em relação aos eixos na esfera de Bloch.

### 3.2.5 Operações em um Qubit

Operações em um qubit são transformações aplicadas a estados quânticos. Três operadores fundamentais em um qubit são as portas de Pauli:  $X$ ,  $Y$ , e  $Z$  (LANDI, 2018). Cada um deles realiza uma operação específica nos estados de um qubit, e podem ser representadas por rotações na esfera de Bloch, em torno de seus respectivos eixos. As representações matriciais dessas portas são as seguintes:

- Porta de Pauli- $X$  ( $X$ ):

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad (40)$$

- Porta de Pauli- $Y$  ( $Y$ ):

$$Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad (41)$$

- Porta de Pauli- $Z$  ( $Z$ ):

$$Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (42)$$

Essas operações são aplicadas a um qubit por meio de multiplicação de matriz, resultando em uma transformação do estado do qubit. Por exemplo, a aplicação da porta  $X$  em um qubit no estado  $|0\rangle$  resultará em  $|1\rangle$ , como demonstrado a seguir, multiplicando a matriz  $X$  pelo vetor  $|0\rangle$ :

$$X|0\rangle = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 + 1 \cdot 0 \\ 1 \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |1\rangle. \quad (43)$$

Enquanto a aplicação da porta  $Y$  criará uma combinação complexa dos estados  $|0\rangle$  e  $|1\rangle$  com uma fase imaginária. Como demonstrado a seguir:

$$Y|0\rangle = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 - i \cdot 0 \\ i \cdot 1 + 0 \cdot 0 \end{bmatrix} = \begin{bmatrix} 0 \\ i \end{bmatrix} = i|1\rangle. \quad (44)$$

As portas de Pauli são elementos-chave na manipulação de informações quânticas e desempenham um papel fundamental em algoritmos quânticos e em circuitos quânticos. Além destas, existem outras operações importantes que podem ser aplicadas a um qubit. A porta

Hadamard (H) é amplamente utilizada na criação de superposições e desempenha um papel fundamental na computação quântica. Sua representação matricial é dada por:

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \quad (45)$$

A aplicação da porta Hadamard em um qubit muda-o de uma base de  $|0\rangle$  e  $|1\rangle$  para uma base de  $|+\rangle$  e  $|-\rangle$ , onde  $|+\rangle$  e  $|-\rangle$  são estados superpostos igualmente distribuídos. Matematicamente, a transformação é definida como:

$$H|0\rangle = |+\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad (46)$$

$$H|1\rangle = |-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle). \quad (47)$$

Estas operações em conjunto são cruciais para o desenvolvimento da teoria de informação quântica, assim como dos capítulos práticos desta tese.

### 3.2.6 Operações com dois Qubits

As operações com dois qubits desempenham um papel fundamental no campo da informação quântica. Elas são essenciais para a construção de algoritmos quânticos e a resolução de problemas complexos. Para o entendimento destas operações, é necessário inicialmente compreender o conceito de produto tensorial, para que possamos nos aprofundar em operações específicas possíveis, assim como posteriormente no fenômeno de entrelaçamento (ou emaranhamento).

#### 3.2.6.1 Produto Tensorial

O produto tensorial é uma forma de criar espaços vetoriais maiores combinando espaços vetoriais (NIELSEN; CHUANG, 2000). O produto tensorial é representado pelo símbolo  $\otimes$ , que, por exemplo, combina dois qubits em um sistema quântico composto. Se temos dois qubits  $|a\rangle$  e  $|b\rangle$ , seu produto tensor é dado por:

$$|a\rangle \otimes |b\rangle = |a\rangle|b\rangle = |ab\rangle. \quad (48)$$

Isso resulta em um espaço de Hilbert maior, onde as combinações de estados possíveis crescem exponencialmente. Em notação matricial, podemos exemplificar o conceito para o caso de querermos combinar dois qubits, o primeiro no estado  $|0\rangle$  e o segundo no estado  $|1\rangle$ :

$$|0\rangle \otimes |1\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ 0 \cdot \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \quad (49)$$

### 3.2.6.2 Portas de Dois Qubits

As portas quânticas de dois qubits são operadores que atuam em dois qubits simultaneamente (as portas introduzem interações entre os qubits). Um exemplo importante é a porta CNOT (Controlled-NOT), que executa uma operação NOT no segundo qubit se o primeiro qubit estiver no estado  $|1\rangle$ . A representação da matriz da porta CNOT é:

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (50)$$

Outra porta crucial é a porta SWAP, que troca os estados de dois qubits. A matriz SWAP é definida como:

$$\text{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (51)$$

Além disso, temos a porta Controlled-Z (CZ), que é usada para introduzir uma fase de  $-1$  no estado  $|11\rangle$  enquanto deixa os outros estados inalterados. A matriz da porta CZ é dada por:

$$\text{CZ} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \quad (52)$$

Várias outras portas quânticas podem ser consultadas em (NIELSEN; CHUANG, 2000), aqui nos ateremos a aquelas utilizadas nas aplicações práticas.

### 3.2.6.3 Emaranhamento

O emaranhamento quântico é uma propriedade não clássica essencial de sistemas compostos por dois ou mais qubits. Enquanto um estado separável de dois qubits pode ser escrito como produto tensorial de estados individuais,

$$|\psi_{\text{sep}}\rangle = |\alpha\rangle \otimes |\beta\rangle,$$

um estado emaranhado não admite essa decomposição. Em outras palavras, não existem vetores  $|\alpha\rangle$  e  $|\beta\rangle$  que cumpram essa igualdade (LANDI, 2018). Essa inseparabilidade implica correlações quânticas que não têm análogo clássico — medições realizadas em um dos qubits afetam instantaneamente as estatísticas do outro, mesmo que eles estejam espacialmente separados.

Os estados de Bell constituem um exemplo paradigmático de emaranhamento máximo. São quatro estados ortonormais em  $\mathbb{C}^2 \otimes \mathbb{C}^2$ :

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad (53)$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad (54)$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad (55)$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \quad (56)$$

Nesses estados, as medidas individuais em cada qubit são completamente aleatórias, mas as correlações conjugadas são perfeitas: por exemplo, em  $|\Phi^+\rangle$  observar “0” no primeiro qubit garante “0” no segundo, e o mesmo vale para “1”.

Esse emaranhamento máximo serve como recurso fundamental em diversos protocolos de informação quântica. No teletransporte quântico, um par de qubits emaranhados permite transferir fielmente um estado desconhecido de um qubit origem para um qubit destino usando apenas operações locais e comunicação clássica (BENNETT et al., 1993). Na codificação superdensa, o emissor envia duas mensagens clássicas ao aplicar portas locais sobre seu qubit em um par de Bell compartilhado, dobrando assim a capacidade de transmissão de informação clássica (BENNETT; WIESNER, 1992). Em esquemas de distribuição de chaves quânticas (QKD), como o protocolo Ekert 91, o uso de estados emaranhados garante segurança contra interceptação, já que qualquer tentativa de medir os qubits altera as correlações e é imediatamente detectada (EKERT, 1991).

Operacionalmente, gera-se emaranhamento puro aplicando uma porta de Hadamard seguida de uma porta CNOT (NIELSEN; CHUANG, 2000):

$$|00\rangle \xrightarrow{H \otimes I} \frac{|0\rangle + |1\rangle}{\sqrt{2}} \otimes |0\rangle \xrightarrow{\text{CNOT}} \frac{|00\rangle + |11\rangle}{\sqrt{2}} = |\Phi^+\rangle.$$

Entender e controlar o emaranhamento — assim como quantificá-lo através de medidas é crucial para a construção de algoritmos quânticos robustos e para o desenvolvimento de arquiteturas de computação e comunicação quânticas.

### 3.2.7 Operador Densidade

O operador de densidade ( $\rho$ ), também conhecido como matriz de densidade, fornece uma representação mais abrangente (quando comparado aos kets  $|\psi\rangle$ ) de um sistema quântico. Ao incorporar probabilidades quânticas e clássicas, este operador fornece uma visão mais completa do estado (SCHULD; PETRUCCIONE, 2018). O conceito de matrizes de densidade também está intrinsecamente ligado ao emaranhamento, uma propriedade quântica fundamental. Mesmo na ausência de incertezas clássicas, existem situações em que é necessário recorrer a

matrizes de densidade para descrever sistemas quânticos com precisão. Conforme destacado por (LANDI, 2018), o operador densidade é considerado o conceito mais importante na teoria quântica. É definido matematicamente como:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad (57)$$

onde  $\rho$  é o operador densidade.  $p_i$  são as probabilidades de o sistema estar em cada estado  $|\psi_i\rangle$ .  $|\psi_i\rangle$  são os estados do sistema e  $|\psi_i\rangle\langle\psi_i|$  é o operador projeção no estado  $|\psi_i\rangle$ .

A matriz densidade  $\rho$  é representada por uma matriz Hermitiana. Sua forma geral em um espaço de dimensão  $n$  é dada por:

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} & \cdots & \rho_{1n} \\ \rho_{21} & \rho_{22} & \cdots & \rho_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{nn} \end{pmatrix}, \quad (58)$$

onde  $\rho_{ij}$  são os elementos da matriz densidade. Os elementos da diagonal principal  $\rho_{ii}$  são chamados populações, pois cada  $\rho_{ii} \geq 0$  representa a probabilidade de encontrar o sistema no estado puro  $|i\rangle$ . Por construção,  $\sum_{i=1}^n \rho_{ii} = 1$ . Já os elementos fora da diagonal principal,  $\rho_{ij}$  com  $i \neq j$ , são denominados coerências. Esses números complexos codificam as fases relativas e as correlações quânticas entre os estados  $|i\rangle$  e  $|j\rangle$ . A presença de coerências ( $\rho_{ij} \neq 0$ ) distingue comportamentos quânticos, como interferência, de misturas clássicas. A matriz densidade é uma ferramenta poderosa para descrever sistemas quânticos em situações onde não se conhece completamente o estado puro do sistema, tornando-se essencial na teoria da informação quântica e na descrição de sistemas abertos e misturas estatísticas.

### 3.2.8 Medidas e Observáveis

Em mecânica quântica, a medição de um sistema faz com que a função de onda colapse (LANDI, 2018). Podemos modelar isso através da chamada de medida projetiva, dada por:

$$p_i = |\langle i|\psi\rangle|^2, \quad (59)$$

onde  $p_i$  é a probabilidade de encontrar o sistema no estado  $|i\rangle$  após a medida.  $\langle i|$  é um dos possíveis estados do sistema após a medida.  $|\psi\rangle$  é o estado do sistema antes da medida.  $\langle i|\psi\rangle$  é o produto interno (ou amplitude de probabilidade) entre o estado  $|i\rangle$  e o estado  $|\psi\rangle$ .

O valor de  $\langle i|\psi\rangle$  representa a amplitude de probabilidade de encontrar o sistema no estado  $|i\rangle$ . Ao elevar ao quadrado o módulo desta amplitude, obtemos a probabilidade real associada. Importante notar que a soma das probabilidades para todos os possíveis estados  $|i\rangle$  é igual a 1, desde que o estado  $|\psi\rangle$  esteja devidamente normalizado, como indicado em (LANDI, 2018):

$$\sum_i p_i = \sum_i \langle \psi | i \rangle \langle i | \psi \rangle = \langle \psi | \psi \rangle = 1. \quad (60)$$

Em informação quântica, observáveis são grandezas físicas associadas a propriedades mensuráveis de um sistema quântico. Eles são representados por operadores hermitianos. Um observável quântico  $A$  é representado pelo operador  $A$  e suas propriedades podem ser descritas através de valores esperados em estados quânticos. A média de um observável  $A$  em um estado  $\rho$  é dada por (COOLEN et al., 2022):

$$\langle A \rangle = \text{Tr}(\rho A), \quad (61)$$

onde  $\langle A \rangle$  é o valor esperado de  $A$ .  $\text{Tr}$  é a função de traço, que calcula a soma dos elementos diagonais de uma matriz.  $\rho$  é o operador densidade que descreve o estado quântico e  $A$  é o operador associado ao observável.

Os observáveis representam quantidades físicas como energia, momento angular, spin, posição, e muitos outros em sistemas quânticos. A teoria dos observáveis é fundamental na descrição e na previsão de resultados de medidas em sistemas quânticos, desempenhando um papel crucial na interpretação e na exploração da mecânica quântica (BUSCH; LAHTI, 2008).

Suponha que tenhamos um operador hermitiano  $A$  com autovetores  $|\lambda_i\rangle$  e autovalores  $\lambda_i$ . Quando realizamos uma medida na base  $|\lambda_i\rangle$ , identificamos o operador  $A$  com um dos autovalores  $\lambda_i$  com uma probabilidade  $p_i$ . Isso ilustra o conceito de medida de observáveis, onde um observável (sendo hermitiano) pode assumir uma gama de valores representados por seus autovalores  $\lambda_i$ , cada um ocorrendo com uma probabilidade  $p_i = |\langle \lambda_i | \psi \rangle|^2$ . É importante notar que qualquer conjunto de estados base  $|i\rangle$  pode ser associado a um observável, o que estabelece uma equivalência entre medir na base e medir um observável. Seguindo essa abordagem, podemos definir o valor esperado do operador  $A$  (LANDI, 2018). No entanto, para fazer isso, tratamos isso como uma média ponderada. Portanto, o valor esperado é dado por:

$$\langle A \rangle := \sum_i \lambda_i p_i. \quad (62)$$

Esta expressão pode também ser escrita de forma alternativa como:

$$\langle A \rangle := \langle \psi | A | \psi \rangle, \quad (63)$$

onde  $\langle A \rangle$  é o valor esperado de  $A$ .  $\lambda_i$  são os autovalores do operador  $A$  e  $p_i$  são as probabilidades de encontrar o sistema em um autoestado correspondente  $|i\rangle$ .

Portanto, o valor esperado do operador pode ser interpretado como o produto escalar entre o operador  $A$  e o estado  $|\psi\rangle$ . Para facilitar ainda mais essa compreensão, usamos o operador de projeção  $P_i = |i\rangle\langle i|$ . A probabilidade de encontrar o sistema no estado  $|i\rangle$  é então

expressa como:

$$p_i = \langle \psi | P_i | \psi \rangle. \quad (64)$$

Essencialmente, essa formulação nos diz que a probabilidade pode ser vista como o valor esperado do operador de projeção no estado  $|\psi\rangle$  (LANDI, 2018).

### 3.2.8.1 População e Coerência

Na teoria da informação quântica, “população” e “coerência” são conceitos fundamentais que descrevem como a informação está distribuída nos estados quânticos de um sistema. Eles são especialmente relevantes ao lidar com estados mistos, que são sistemas quânticos descritos por operadores densidade  $\rho$  em vez de estados puros. A população de um estado quântico representa a probabilidade de encontrar o sistema em um estado específico dentro de um estado misto  $\rho$ . Matematicamente, a população  $P_i$  de um estado  $|\psi_i\rangle$  em um estado misto é dada por:

$$P_i = \langle \psi_i | \rho | \psi_i \rangle, \quad (65)$$

onde  $P_i$  é a população do estado  $|\psi_i\rangle$ .  $|\psi_i\rangle$  é um estado quântico específico.  $\rho$  é o operador densidade que descreve o estado misto.

A população mede a contribuição de um estado quântico específico para a descrição do sistema quântico geral (KORZEKWA et al., 2018). É uma medida de probabilidade e indica a chance de encontrar o sistema em um estado particular quando uma medida é realizada.

A coerência está relacionada à fase relativa entre diferentes estados quânticos dentro de um estado misto. Ela descreve a parte “não-diagonal” da matriz de densidade  $\rho$ , que é responsável por interferências quânticas (KORZEKWA et al., 2018). Matematicamente, a coerência  $\rho_{ij}$  entre os estados  $|\psi_i\rangle$  e  $|\psi_j\rangle$  é dada por:

$$\rho_{ij} = \langle \psi_i | \rho | \psi_j \rangle. \quad (66)$$

A coerência reflete a preservação das relações de fase entre os componentes de um estado quântico misto. Quando a coerência é presente, os estados quânticos individuais ainda mantêm informações sobre a fase relativa entre si, permitindo interferências quânticas. Em resumo, população e coerência são conceitos cruciais em informação quântica que descrevem como a informação está distribuída em estados quânticos mistos. A população mede a probabilidade de encontrar o sistema em um estado específico, enquanto a coerência reflete a preservação das relações de fase entre diferentes estados no sistema. Ambos desempenham papéis significativos na descrição e manipulação de sistemas quânticos complexos.

### 3.2.8.2 Operação Traço Parcial

Na informação quântica, a operação de traço parcial é uma ferramenta que desempenha um papel essencial na análise e descrição de sistemas quânticos. Ela permite que foquemos

apenas nos graus de liberdade de interesse em um sistema composto, ignorando os graus de liberdade do ambiente ou do sistema total. A operação de traço parcial é frequentemente representada por  $\text{Tr}_E$ , onde  $E$  se refere aos graus de liberdade que desejamos “traçar” ou ignorar. Suponha que temos um sistema quântico composto  $S + E$  com operador densidade total  $\rho_{SE}$ , onde  $S$  é o sistema de interesse e  $E$  é o ambiente. A operação de traço parcial  $\text{Tr}_E$  é definida como (NIELSEN; CHUANG, 2000):

$$\rho_S = \text{Tr}_E(\rho_{SE}). \quad (67)$$

Aqui,  $\rho_S$  é o operador densidade reduzido do sistema  $S$ , obtido após traçar o ambiente. Isso significa que obtemos uma descrição simplificada do sistema  $S$  isolando-o de  $E$ , mantendo apenas as informações relevantes.

### 3.2.9 Hamiltoniano

O Hamiltoniano é um conceito fundamental em informação quântica e na mecânica quântica em geral. Ele descreve a energia total de um sistema quântico e permite entendermos a evolução temporal dos estados quânticos ao longo do tempo. Determinar o Hamiltoniano necessário para descrever um sistema físico específico é um problema desafiador que, em grande parte, requer informações experimentais. O Hamiltoniano, sendo um operador Hermitiano, possui uma decomposição espectral (NIELSEN; CHUANG, 2000):

$$H = \sum E|E\rangle\langle E|, \quad (68)$$

onde Autovalores  $E$  representam as energias associadas ao sistema e  $|E\rangle$  são os estados próprios do sistema relacionados a essas energias. O estado de menor energia é chamado de estado fundamental e os estados estacionários são chamados assim porque sua única mudança no tempo é adquirir um fator numérico global (NIELSEN; CHUANG, 2000),

$$|E\rangle \rightarrow \exp(-iEt/\hbar)|E\rangle. \quad (69)$$

O hamiltoniano geral de um único qubit, por exemplo, seria dado por  $H = \omega X$ . Onde  $\omega$  representa a frequência angular ou energia associada ao sistema e  $X$  é o operador de Pauli  $\sigma_x$ .

O operador Hamiltoniano desempenha um papel central na mecânica quântica, pois governa a dinâmica dos sistemas quânticos. A equação de Schrödinger, uma das equações fundamentais da mecânica quântica, descreve como os estados quânticos evoluem no tempo sob a influência do Hamiltoniano (ARRUDA, 2011):

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H} |\psi(t)\rangle, \quad (70)$$

onde  $i$  é a unidade imaginária.  $\hbar$  é a constante reduzida de Planck.  $|\psi(t)\rangle$  é o estado quântico no tempo  $t$ .

A solução desta equação determina como os estados quânticos se desenvolvem no tempo, permitindo a previsão de resultados de medidas em sistemas quânticos. O Hamiltoniano é especialmente importante em sistemas quânticos, como átomos, moléculas e partículas subatômicas, onde desempenha um papel fundamental na descrição de transições de energia e na evolução dos estados quânticos. Em resumo, o Hamiltoniano é uma peça central na mecânica quântica e desempenha um papel crucial na compreensão e na manipulação de sistemas quânticos.

### 3.3 Sistemas Quânticos Abertos

O estudo dos sistemas quânticos abertos busca compreender como os sistemas quânticos interagem com ambientes. Nesta seção, exploraremos os princípios fundamentais da evolução temporal de sistemas quânticos abertos e discutiremos a equação mestra de Lindblad, que é uma ferramenta poderosa para descrever essa evolução.

#### 3.3.1 Evolução Temporal

Na teoria quântica, existem dois tipos de sistemas físicos: sistemas fechados e sistemas abertos. Os sistemas quânticos fechados são ideais, perfeitamente isolados do ambiente, enquanto os sistemas reais estão sujeitos a interações com o meio externo. Essas interações indesejadas são inevitáveis e tornam todos os sistemas quânticos sistemas abertos. Sistemas quânticos abertos não seguem uma evolução temporal unitária, ao contrário dos sistemas fechados. A dinâmica desses sistemas abertos é afetada por interações com o ambiente, resultando em uma evolução não unitária, o que significa que sua evolução não pode ser representada de forma simples e determinística (ARRUDA, 2011). A evolução temporal de sistemas quânticos abertos é um tópico central na física quântica e na teoria da informação quântica. Ela lida com a dinâmica de sistemas quânticos que interagem com seus ambientes externos, resultando em mudanças nos estados quânticos ao longo do tempo (KHALIL; RICHERT, 2021). Uma forma de descrever essa evolução, é usando a equação mestra, que é uma equação diferencial que modela a dinâmica dos sistemas quânticos abertos.

Um sistema aberto,  $S$ , interage com seu ambiente, formando um sistema combinado  $S + E$ . Embora o sistema combinado seja considerado fechado e evolua unitariamente, o sistema  $S$  passa por uma evolução não unitária devido à interação com o ambiente. A descrição completa da dinâmica do sistema combinado é complexa, geralmente envolvendo infinitos graus de liberdade no meio ambiente. Portanto, concentramos nossa análise no sistema reduzido,  $S$ , utilizando um operador densidade  $\rho_s$  obtido por traço parcial sobre o meio ambiente (ARRUDA, 2011).

A evolução temporal do operador densidade reduzido  $\rho_s(t)$  é determinada pela seguinte

equação:

$$\frac{d}{dt}\rho_s(t) = -\frac{i}{\hbar}\text{Tr}_e\{[H_{\text{int}}, \rho(t)]\}, \quad (71)$$

onde  $H_{\text{int}}$  é o Hamiltoniano de interação,  $\rho(t)$  é o operador densidade do sistema combinado e  $\hbar$  é a constante de Planck reduzida. Esta equação descreve a evolução do sistema de interesse enquanto interage com seu ambiente e é fundamental na teoria de sistemas quânticos abertos (ARRUDA, 2011). A equação mestra é fundamental para entender a evolução temporal de sistemas quânticos abertos, especialmente em aplicações como a descrição de sistemas emaranhados, a modelagem de sistemas quânticos complexos e a análise de sistemas quânticos com decoerência. Ela fornece uma base sólida para a compreensão das transformações quânticas que ocorrem quando os sistemas interagem com seus ambientes, tornando-se uma ferramenta essencial na teoria da informação quântica e na física quântica em geral.

### 3.3.2 Equação Mestra de Lindblad

A equação mestra de Lindblad descreve a evolução temporal de sistemas quânticos em interação com ambientes que causam dissipação e decoerência. Ela é amplamente utilizada em áreas como informação quântica, física da matéria condensada e estudos de sistemas quânticos complexos. A forma geral da equação mestra de Lindblad é dada por (NIELSEN; CHUANG, 2000):

$$\frac{d}{dt}\rho = -i[H, \rho] + \sum_k \left( L_k \rho L_k^\dagger - \frac{1}{2} \{L_k^\dagger L_k, \rho\} \right), \quad (72)$$

onde  $\rho$  é o operador densidade do sistema quântico,  $H$  é o Hamiltoniano do sistema, representando a parte não dissipativa da dinâmica,  $L_k$  são os operadores de Lindblad que descrevem as interações com o ambiente e  $\{A, B\}$  denota o anticomutador entre os operadores  $A$  e  $B$ .

Este formalismo assume markovianidade, ou seja, trata a evolução como um processo sem memória: a taxa de variação de  $\rho$  em cada instante  $t$  depende apenas do estado  $\rho(t)$  e não de valores anteriores. A equação mestra de Lindblad é, portanto, a descrição geral de uma dinâmica quântica markoviana.

O primeiro termo,  $-i[H, \rho]$ , representa a evolução unitária do sistema sob o Hamiltoniano  $H$ , governando a parte não dissipativa da dinâmica. Os termos subsequentes,  $\sum_k \left( L_k \rho L_k^\dagger - \frac{1}{2} \{L_k^\dagger L_k, \rho\} \right)$ , modelam as interações dissipativas com o ambiente. Eles descrevem como os estados quânticos do sistema evoluem devido a processos de relaxamento e decoerência induzidos pelo ambiente. Os operadores de Lindblad  $L_k$  são escolhidos para representar as interações específicas com o ambiente, e sua escolha depende das características do sistema em estudo. Eles são fundamentais para descrever a dissipação e a perda de coerência que ocorrem em sistemas quânticos abertos.

### 3.4 Aplicações e Perspectivas Futuras

A informação quântica representa um campo científico de vanguarda que promete redefinir múltiplas áreas da ciência e tecnologia. Suas aplicações atuais já impressionam, com destaque para a computação quântica (NIELSEN; CHUANG, 2000), que é uma grande promessa para o futuro da computação, e a criptografia quântica, que garante comunicações seguras em um mundo cada vez mais digital (BENNETT; BRASSARD, 2014). Além disso, a simulação quântica permite a modelagem precisa de sistemas complexos (FEYNMAN, 2018), desde reações químicas até materiais avançados (CAO et al., 2019). Outras aplicações incluem avanços na descoberta de novos materiais (ASPURU-GUZIK; WALTHER, 2012), desenvolvimento de medicamentos personalizados e segurança reforçada (BENNETT; BRASSARD; EKERT, 1992) em comunicações globais. Com a contínua evolução da tecnologia quântica, novas aplicações e descobertas estão moldando o futuro da ciência e da tecnologia, apontando para uma era empolgante de inovações e avanços.

# 4 Quantum Machine Learning

Desde os anos 1990, há um interesse crescente na construção de computadores quânticos e na pesquisa de algoritmos quânticos. Isso levou ao desenvolvimento de linguagens e teorias para descrever esses dispositivos. No entanto, a construção de computadores quânticos é desafiadora devido à necessidade de controle preciso de variáveis microscópicas e correção de erros. Embora a implementação de muitos algoritmos quânticos ainda seja um desafio, há otimismo de que computadores quânticos de “larga escala” se tornarão disponíveis no futuro, com grande envolvimento da indústria da tecnologia. Além disso, dispositivos quânticos menores sem correção de erros estão sendo desenvolvidos, com ênfase na busca por algoritmos úteis que superem os algoritmos clássicos. A aprendizagem de máquina, que se concentra em fazer com que os computadores aprendam a partir de dados, é vista como uma aplicação promissora na computação quântica (SCHULD; PETRUCCIONE, 2018).

A aprendizagem de máquina, juntamente com a computação quântica, prometem desempenhar um papel significativo na sociedade. A combinação dessas disciplinas é explorada no campo emergente do aprendizado quântico, que busca aproveitar as sinergias entre eles. Em sua definição mais ampla, existem inúmeras abordagens que podem ser utilizadas com esse intuito, incluindo o uso de técnicas matemáticas da teoria quântica para desenvolver novos métodos no aprendizado de máquina e como o aprendizado de máquina pode ser aplicado na análise de dados provenientes de experimentos quânticos. Outro campo ambicioso é a aplicação do aprendizado de máquina com o auxílio de computadores quânticos, visando observar se computadores quânticos têm o potencial de acelerar a resolução de problemas, aprender com conjuntos menores de dados e lidar com níveis mais elevados de ruído (SCHULD; PETRUCCIONE, 2018).

Podemos dividir o QML em quatro grandes áreas (AI'MEUR; BRASSARD; GAMBS, 2006) e o diagrama QQ, QC, CQ, CC é frequentemente usado para categorizar estas áreas, indicando se elas envolvem ou não componentes quânticos (Q) e/ou clássicos (C) em diferentes partes do processo (WOSSNIG, 2021). A primeira letra se refere a origem dos dados utilizados, enquanto a segunda, ao dispositivo que processa a informação. O diagrama pode ser observado abaixo:

- QQ (Quantum to Quantum): Nessa categoria, tanto a entrada quanto o processamento dos dados são quânticos. Isso significa que os dados de entrada são quânticos e o algoritmo de aprendizado também opera em um ambiente quântico;
- QC (Quantum to Classical): Nesse caso, os dados de entrada são quânticos, mas o processamento é feito em um ambiente clássico.

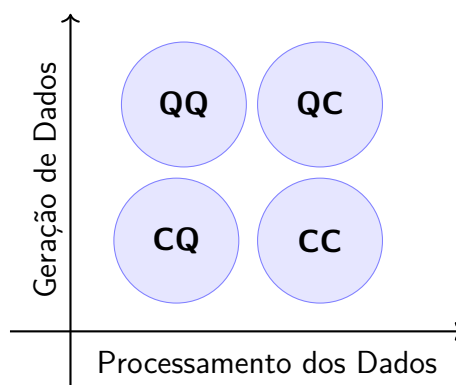


Figura 4 – Áreas de Estudo dentro de Aprendizado de Máquina Quântico. Do autor, com base em (SCHULD; PETRUCCIONE, 2018)

- CQ (Classical to Quantum): Aqui, os dados de entrada são clássicos, mas o processamento é feito em um ambiente quântico.
- CC (Classical to Classical): Nessa categoria, tanto os dados de entrada quanto o processamento são clássicos, sem envolver dados quânticos.

O diagrama QQ, QC, CQ, CC é útil para categorizar as diferentes abordagens e algoritmos no campo do aprendizado de máquina quântica, fornecendo uma maneira de entender onde a computação quântica desempenha um papel e onde a computação clássica é usada no processo de aprendizado (SCHULD; PETRUCCIONE, 2018).

Neste capítulo iremos abordar a área promissora do QML. Começaremos com uma contextualização histórica, seguida de uma breve apresentação sobre estratégias de codificação de informações, como Basis Encoding e Amplitude Encoding, que desempenham um papel fundamental na realização de algoritmos quânticos eficazes. Além disso, examinaremos o conceito de Quantum Speedup e suas limitações no contexto de Quantum Machine Learning. Prosseguiremos abordando tópicos essenciais, como Circuitos Quânticos Variacionais, Redes Neurais Quânticas e Métodos de Kernel, que são ferramentas cruciais para o desenvolvimento de algoritmos quânticos de aprendizado de máquina. Este capítulo também fornecerá insights sobre perspectivas futuras e tendências empolgantes no campo da Quantum Machine Learning.

## 4.1 Contextualização Histórica

O Quantum Machine Learning, é um campo de pesquisa novo que tem suas raízes na evolução da computação quântica e na tentativa de explorar as capacidades dos sistemas quânticos para resolver problemas complexos de maneira mais eficiente. Desde 1995, houveram tentativas esporádicas de unir as disciplinas de computação quântica e aprendizado de máquina, incluindo investigações de modelos quânticos de redes neurais (SCHULD; SINAYSKIY; PETRUCCIONE, 2014), no entanto, essas abordagens iniciais eram principalmente inspira-

das na biologia. No início dos anos 2000, houveram discussões sobre teoria de aprendizado estatístico em um contexto quântico além de alguns workshops sobre “Computação Quântica e Aprendizado”. Surgiram também publicações esporádicas sobre algoritmos quânticos de aprendizado de máquina, como a “memória quântica associativa” de Ventura e Martinez (VENTURA; MARTINEZ, 2000), e o algoritmo “QBoost” de Hartmut Neven (NEVEN et al., 2009). Por volta de 2013, o termo “aprendizado de máquina quântico” começou a ser usado, a partir de (LLOYD; MOHSENI; REBENTROST, 2013) e o interesse pelo campo aumentou significativamente. A partir deste ponto, vários workshops e conferências internacionais foram organizados, e muitos grupos de pesquisa iniciaram projetos e colaborações nessa área. A combinação de um mercado em crescimento de bilhões de dólares da computação quântica, com o já consolidado mercado da aprendizagem de máquina também despertou interesse na indústria (SCHULD; PETRUCCIONE, 2018).

Ainda existem muitos desafios a serem superados, incluindo a construção de hardware quântico mais robusto e escalável. No entanto, o QML oferece a promessa de revolucionar a computação e o aprendizado de máquina, abrindo portas para novas descobertas científicas e aplicações práticas.

## 4.2 *Quantum Speedup*

O conceito de “*quantum speedup*” refere-se à capacidade intrínseca dos algoritmos quânticos de superar seus equivalentes clássicos em termos de eficiência computacional. Isso significa que, para determinados problemas, algoritmos quânticos podem realizar o mesmo cálculo em menos etapas ou com menos recursos computacionais do que os algoritmos clássicos correspondentes. A existência de quantum speedup depende da perspectiva sob a qual se analisa: uma visão formal de ciência da computação exige provas matemáticas, enquanto uma perspectiva prática se baseia em evidências estatísticas obtidas com dispositivos de tamanho finito. A performance máxima de algoritmos clássicos muitas vezes não é conhecida, tornando o quantum speedup uma incógnita. As medidas ideais para caracterizar quantum speedup incluem complexidade de consulta (quantas vezes a fonte de informação é acessada) e complexidade de portas (quantas operações quânticas são necessárias para obter o resultado desejado). Essas medidas ajudam a determinar se um algoritmo quântico supera seu equivalente clássico, mas mapear essa idealização para a realidade e quantificar os recursos necessários continua sendo um desafio (BIAMONTE et al., 2017).

Na prática, algoritmos quânticos para aprendizado de máquina, como o Quantum Basic Linear Algebra Subroutines (qBLAS), demonstram quantum speedup ao superar os melhores algoritmos clássicos conhecidos em tarefas específicas. Esses algoritmos quânticos podem acelerar significativamente operações fundamentais, como transformadas de Fourier, busca de autovalores e autovetores, resolução de equações lineares, entre outros. Como resultado,

eles apresentam quantum speedup em uma variedade de algoritmos de análise de dados e aprendizado de máquina, incluindo álgebra linear, ajuste por mínimos quadrados, análise de componentes principais e programação linear. Além disso, processadores de informação quântica especializados, como os *quantum annealers*, mostram-se como grande promessa para aplicação em arquiteturas de aprendizado profundo. Embora a extensão desse potencial ainda não esteja completamente compreendida, existe um otimismo substancial de que os computadores quânticos têm a capacidade de identificar padrões em dados que permanecem inacessíveis aos computadores clássicos.(BIAMONTE et al., 2017).

### 4.3 Codificação da Informação

Especialmente quando queremos usar um computador quântico no aprendizado de máquina com dados clássicos (denominado caso CQ), é necessário criar um método para “carregar” dados de uma memória clássica para o computador quântico, chamado de preparação de estados. A estratégia de representar informações como estados quânticos orienta o *design* do algoritmo quântico e é determinante nos possíveis ganhos de velocidade. A eficiência, precisão e ruído desempenham um papel crucial na avaliação do desempenho, especialmente porque muitos algoritmos quânticos fornecem resultados probabilísticos, exigindo várias repetições, tornando essencial a interface entre a memória clássica e o dispositivo quântico (SCHULD; PETRUCCIONE, 2018). Três formas fundamentais de representação da informação são discutidas neste capítulo: a Codificação de Bases, que envolve a representação de informações em uma base específica; a Codificação de Amplitude, que se concentra na manipulação das amplitudes de probabilidade associadas a estados quânticos; e o Feature Map de Pauli, que utiliza operadores de Pauli para mapear dados clássicos em estados quânticos de maneira eficiente. À medida que continuamos nossa jornada na compreensão da computação quântica e do aprendizado de máquina quântico, essas técnicas se tornam elementos-chave na construção de algoritmos eficientes.

#### 4.3.1 *Basis Encoding*

A codificação de base é uma abordagem fundamental na computação quântica em que cada estado de base computacional de um sistema de  $n$ -qubits corresponde diretamente a uma sequência clássica de  $n$ -bits. Ou seja, apenas substituímos os bits pelos qubits. As amplitudes de cada estado de base representam a probabilidade de medir esse estado. O principal objetivo de um algoritmo quântico que utiliza a codificação de base é aumentar a probabilidade (o quadrado do módulo da amplitude) de medir o estado de base que codifica a solução desejada. Em essência, a computação quântica visa “marcar” a resposta correta com uma alta probabilidade de medição. Assim como na computação clássica, a codificação de base depende de representações binárias de números. Os estados quânticos representam números

reais na forma binária, e diferentes representações binárias (por exemplo, fixas ou de ponto flutuante) podem ser utilizadas conforme necessário (SCHULD; PETRUCCIONE, 2018).

A codificação de base permite a representação de elementos como estados de base de  $n$ -qubit em um espaço de Hilbert dimensional de  $2^n$ . Para um único ponto de dados  $x$  em um conjunto de dados  $X$ , a codificação de base o representa como um estado de base de  $n$ -qubit:

$$\mathcal{E}(x) = |x\rangle \quad (73)$$

Para até  $2^n$  elementos em  $X$ , a codificação de base os representa como uma superposição de estados de base:

$$\mathcal{E}(X) = \frac{1}{\sqrt{k}} \sum_x |x\rangle \quad (74)$$

onde  $k$  é o número de estados  $|x\rangle$  na superposição. O termo  $\frac{1}{\sqrt{k}}$  é um fator de normalização, necessário para garantir que o estado quântico resultante tenha uma norma igual a 1. O carregamento de um único ponto de dados tem um custo computacional que escala como  $O(n)$ , enquanto o carregamento de uma superposição de itens de dados escala como  $O(kn)$  (ZOUFAL, 2021).

Para preparar superposições de entradas, é utilizado um procedimento de preparação de estados. Isso envolve o carregamento de padrões em um registro quântico, a divisão do estado e a fusão em um ramo de memória. Esse processo permite a eficiente codificação de padrões binários em estados quânticos. A codificação de base permite que algoritmos quânticos operem em características binárias, tornando possível executar qualquer operação que possa ser realizada em um computador clássico. Teoricamente, é possível traduzir algoritmos clássicos de forma eficiente para algoritmos quânticos, embora essa tradução possa ter alguma sobrecarga. Ao trabalhar com modelos de aprendizado de máquina quântica, os resultados também podem ser codificados na base. A estimativa da probabilidade de resultados de medição e, portanto, da previsão, pode ser feita por meio de medições repetidas. O número de medições necessárias para estimativas confiáveis depende da margem de erro desejada e do nível de confiança (SCHULD; PETRUCCIONE, 2018).

### 4.3.2 Amplitude Encoding

Em codificação de amplitude, dados clássicos normalizados, representados como um conjunto  $X = \{x_0, \dots, x_{k-1}\}$  com  $\sum_{j=0}^{k-1} |x_j|^2 = 1$ , são mapeados nas amplitudes contínuas de um estado quântico de  $n$ -qubits. A codificação de amplitude refere-se a:

$$E(X) = \sum_{j=0}^{k-1} \sqrt{x_j} |j\rangle, \quad (75)$$

onde  $|j\rangle$  são estados de base do espaço de Hilbert subjacente. Esse método de codificação permite a representação de até  $2^n$  dados escalares com  $n$  qubits. A eficiência da codificação de

amplitude depende da capacidade de preparar estados quânticos arbitrários de forma eficiente e robusta. (ZOUFAL, 2021).

A codificação de amplitude em computação quântica envolve representar conjuntos de dados usando as amplitudes de um estado quântico. Isso exige a normalização dos vetores de dados clássicos, resultando em representações quânticas que residem em uma dimensão a menos. A preparação eficiente de estados quânticos de amplitude e a leitura das amplitudes podem ser desafiadoras, dependendo da estrutura dos dados. A não-linearidade nas amplitudes e a necessidade de normalização são considerações críticas. Existe também uma relação entre a qualidade da codificação e a eficiência do hardware quântico disponível. No geral, a codificação de amplitude é uma técnica poderosa, mas requer considerações cuidadosas em relação à preparação de estados quânticos e às limitações inerentes à descrição probabilística (SCHULD; PETRUCCIONE, 2018).

### 4.3.3 Pauli Feature Map

No contexto de aprendizado de máquina quântico, o *feature map* de Pauli desempenha um papel fundamental na codificação de dados clássicos em estados quânticos. Através dele, os dados clássicos são mapeados em um circuito parametrizado, onde os parâmetros correspondem aos dados de entrada e são transformados em estados quânticos  $|\phi(\vec{x})\rangle$ . Essa técnica é especialmente útil para explorar as correlações não-lineares e as interações entre as características dos dados, potencialmente capturando padrões complexos que seriam difíceis de identificar em métodos clássicos (HOSSAIN et al., 2021).

Neste trabalho, utilizamos o ZFeatureMap (QUANTUM, 2021), uma instância especializada do circuito de expansão de Pauli (*Pauli Expansion Circuit*), para realizar a codificação de dados clássicos  $\vec{x} \in \mathbb{R}^n$  em estados quânticos. O ZFeatureMap aplica rotações baseadas no operador de Pauli- $Z$ , uma escolha eficiente para capturar as relações entre as características de maneira simples, sem a introdução imediata de entrelaçamento entre os qubits. O ZFeatureMap é particularmente eficiente quando o objetivo é codificar os dados usando operações localizadas em qubits individuais. Além disso, ao repetir a aplicação do mapa, é possível capturar correlações de ordem superior entre as características dos dados, modelando assim padrões não-lineares complexos. Essas interações de características referem-se à habilidade de representar não apenas as características individuais, mas também como elas interagem de forma combinada para influenciar os resultados. O uso repetido desse mapa permite modelar essas interações de forma mais detalhada.

De maneira geral, um *feature map* de Pauli pode ser representado como (HAVLÍČEK et al., 2019a):

$$U_{\Phi(\vec{x})} = \exp\left(i \sum_{S \subseteq [n]} \phi_S(\vec{x}) \prod_{i \in S} P_i\right), \quad (76)$$

onde  $[n] = \{1, 2, \dots, n\}$  é o conjunto de índices dos qubits e cada  $S \subseteq [n]$  indica quais

qubits participam de um termo. Quando  $|S| = 1$  temos rotações locais, mas para  $|S| > 1$  surgem termos não locais que geram interações e potencial emaranhamento entre múltiplos qubits—portanto este mapa não é puramente local. Aqui  $P_i \in \{I, X, Y, Z\}$  atua no qubit  $i$ , e  $\phi_S(\vec{x})$  mapeia o vetor de características  $\vec{x}$  em fases quânticas.

No ZFeatureMap, somente termos com  $|S| = 1$  e  $P_i = Z$  são usados, resultando em rotações independentes no eixo  $Z$ . Mapas como o ZZFeatureMap incluem interações de dois qubits ( $|S| = 2$ ), aumentando a complexidade do encoding e introduzindo emaranhamento (JAVADI-ABHARI et al., 2024).

## 4.4 Algoritmos Quânticos Variacionais

Os Algoritmos Quânticos Variacionais (do Inglês *Variational Quantum Algorithms* - VQAs) são uma classe de algoritmos quânticos que desempenham um papel fundamental no campo da computação quântica. Estes recentemente têm emergido como a principal estratégia na tentativa de obter *speedups* na era NISQ. Eles são fundamentados na utilização de circuitos quânticos parametrizados, cujos parâmetros são representados por vetores, como  $\theta$ . Associados a esses circuitos está uma função de custo, denotada como  $C(\theta)$ , que constitui o alvo da otimização. A tarefa primordial dos VQAs é encontrar os valores ideais dos parâmetros  $\theta$  que minimizam essa função de custo. Isso é formulado da seguinte maneira:

$$\theta^* = \arg \min_{\theta} C(\theta), \quad (77)$$

onde  $\theta^*$  representa o conjunto ótimo de parâmetros a serem encontrados, e  $C(\theta)$  é a função de custo que avalia o desempenho do circuito parametrizado. Nesta equação, o operador  $\arg \min$  indica a busca pelo conjunto de parâmetros  $\theta$  que minimiza a função de custo  $C(\theta)$ . A solução ideal  $\theta^*$  corresponde aos parâmetros que resultam no valor mínimo de  $C(\theta)$  (CEREZO et al., 2021).

Um componente fundamental dos VQAs é a escolha de uma arquitetura de ansatz. O ansatz é o circuito quântico inicial com parâmetros ajustáveis. Diversas arquiteturas de ansatz comuns são utilizadas em aplicações quânticas. O “Hardware Efficient Ansatz” (KANDALA et al., 2017) é projetado para implementações eficientes em hardware quântico específico, respeitando sua conectividade e interações. O “Unitary Coupled Cluster Ansatz (UCC)” (ROMERO et al., 2018) é amplamente empregado em química quântica para calcular energias moleculares, construído com operadores de excitação simples e dupla. O “Quantum Alternating Operator Ansatz” (HADFIELD et al., 2019) é empregado em problemas de otimização combinatória, alternando operadores de problema e mistura em um circuito quântico. Por fim, o “Variational Hamiltonian Ansatz” (WECKER et al., 2015) é usado em simulação quântica e otimização, parametrizando um Hamiltoniano com termos ajustáveis.

Os VQAs dependem da colaboração com otimizadores clássicos para ajustar os parâmetros dos circuitos quânticos. Vários otimizadores são relevantes para essa tarefa. O “Gradiente

Descendente” (RUDER, 2016) emprega informações do gradiente da função de custo para iterativamente ajustar os parâmetros. O “Adam” (KINGMA; BA, 2014) é uma variação do gradiente descendente que se destaca por adaptar o tamanho dos passos da otimização. O “Simultaneous Perturbation Stochastic Approximation (SPSA)” (SPALL, 1992) é um método que evita o cálculo direto do gradiente, atualizando os parâmetros com base em diferenças finitas. Além disso, o “Meta-Learning” (FINN; ABBEEL; LEVINE, 2017) permite que o otimizador “aprenda a aprender” e adapte seus passos de atualização com base no histórico da otimização (CEREZO et al., 2021).

## 4.5 Redes Neurais Quânticas

Redes Neurais Quânticas (RNQ) representam uma abordagem inovadora na interseção entre a computação quântica e redes neurais artificiais. Elas baseiam-se em circuitos quânticos parametrizados, conhecidos como circuitos variacionais quânticos. Um circuito variacional quântico pode ser interpretado como uma camada linear em uma rede neural, na qual as conexões são representadas por matrizes unitárias e os pesos são os parâmetros do circuito (SCHULD; PETRUCCIONE, 2018). A representação de uma camada quântica em uma RNQ pode ser definida como:

$$U(\boldsymbol{\theta}) = G(\theta_L) \cdot \dots \cdot G(\theta_1) \quad (78)$$

onde  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_L)$  é o vetor de parâmetros do circuito, e  $G(\theta_i)$  representa uma porta quântica parametrizada na camada  $i$ . As RNQ frequentemente adotam uma abordagem híbrida, incorporando camadas quânticas em redes neurais profundas clássicas. Isso permite que aproveitem a capacidade de processamento quântico para tarefas específicas, mantendo a escalabilidade da computação clássica. Além disso, estratégias de decomposição de matrizes unitárias e a aplicação de operadores de controle em portas quânticas ajudam a limitar o número de parâmetros, garantindo que circuitos variacionais sejam capazes de realizar transformações complexas com um número de parâmetros eficiente em relação à dimensão de entrada. Portanto, as RNQ apresentam uma combinação única de poder computacional quântico e eficiência na representação de modelos. As Redes Neurais Quânticas têm o potencial de revolucionar a abordagem a tarefas de aprendizado de máquina complexas à medida que a tecnologia quântica continua a evoluir (SCHULD; PETRUCCIONE, 2018).

## 4.6 Métodos de Kernel

Os métodos de kernel baseiam-se no uso de *kernels*, que podem ser entendidos como medidas de distância no espaço de entrada. Esses kernels são usados para calcular distâncias entre entradas de treinamento e novas entradas para classificação. Os kernels correspondem a produtos internos de pontos de dados mapeados para um espaço de recursos de maior dimensão. O teorema de *representer* mostra que uma ampla classe de modelos treinados pode ser expressa

em termos de kernels (SCHULD, 2021). Além disso, o *kernel trick* permite a construção de modelos usando diferentes funções de kernel.

Um *kernel*  $\kappa : X \times X \rightarrow \mathbb{C}$  é uma função positiva semidefinida que mede a distância entre pontos de dados no espaço de entrada  $X$ . Ele satisfaz  $\kappa(x, x) \geq 0$  e  $\kappa(x, x') = \kappa(x', x)^*$ . Um *mapeamento de recursos*  $\phi : X \rightarrow \mathcal{C}X$  pode ser construído a partir de um kernel. Os vetores de recursos em  $\mathcal{C}X$  são funções complexas que mapeiam do espaço de entrada para  $\mathbb{C}$ . O produto interno de vetores de recursos pode ser definido como

$$f, g = \sum_{m=1}^M \sum_{m'=1}^{M'} \nu_m^* \mu_{m'} \kappa(x_m, x_{m'}). \quad (79)$$

Os mapeamentos de recursos fornecem uma maneira de mapear entradas para um espaço de recursos e o produto interno nesse espaço é definido em termos do kernel. Essa relação é uma ideia fundamental nos métodos de kernel, onde o kernel calcula o produto interno de entradas mapeadas. O espaço de recursos pode ser estendido para um espaço de Hilbert chamado *Espaço de Hilbert Reprodutor*. Da mesma forma, dada uma transformação  $\phi : X \rightarrow H$ , onde  $H$  é um espaço de Hilbert com vetores  $\phi(x)$ , um kernel pode ser definido como  $\kappa(x, x') = \langle \phi(x), \phi(x') \rangle$ . Esse kernel é positivo semidefinido, satisfazendo as propriedades necessárias (SCHULD; PETRUCCIONE, 2018).

Ou seja, embora o kernel em si não realize diretamente a transformação dos dados, ele permite computar o produto interno entre amostras como se elas tivessem sido mapeadas para um espaço de características de dimensão maior. Essa é uma das ideias centrais por trás do *kernel trick*: ao invés de transformar explicitamente os dados via um mapeamento não linear  $\phi : X \rightarrow H$ , basta definir um kernel  $\kappa(x, x') = \langle \phi(x), \phi(x') \rangle$ , evitando o custo computacional da transformação. Essa abordagem permite aplicar algoritmos lineares em espaços de alta (ou infinita) dimensão, onde as estruturas dos dados tornam-se linearmente separáveis, mesmo que originalmente não o fossem no espaço de entrada.

Em seguida, iremos definir um kernel quântico. A codificação de entrada mapeia  $x$  para um vetor em um espaço de Hilbert. O principal componente de um kernel quântico é interpretar o processo de codificação de uma entrada  $x$  em um estado quântico  $|\phi(x)\rangle$  como um *feature map* ou vetor de características. Se considerarmos o estado quântico  $|\phi(x)\rangle$  como um vetor de características na codificação de amplitude, também conhecido como um “estado de recurso quântico”, então o produto interno entre dois desses pode ser associado a um kernel quântico (SCHULD; PETRUCCIONE, 2018):

$$\kappa(x, x') = \langle \phi(x) | \phi(x') \rangle. \quad (80)$$

Qualquer dispositivo de computação quântica que seja capaz de efetuar o cálculo do produto interno em questão tem a capacidade de realizar estimativas do kernel. Tais estimativas, por sua vez, podem ser aplicadas em algoritmos de aprendizado de máquina baseados em métodos de kernel clássicos. Quando o resultado desse produto interno ultrapassa a capacidade

de simulação de um computador clássico, classificamos o kernel como intratável em termos clássicos. Atualmente, um desafio de relevância reside na identificação de kernels quânticos que não somente apresentem vantagens no contexto quântico, mas também demonstrem pertinência e utilidade em cenários de aprendizado de máquina. É fundamental destacar que a natureza intrínseca do kernel quântico é inteiramente definida pela estratégia de codificação de entrada selecionada, isto é, pelo circuito que realiza a preparação do estado de entrada  $|\phi(x)\rangle$ . A adoção de diferentes abordagens de codificação resulta na geração de distintos tipos de kernels (SCHULD; PETRUCCIONE, 2018).

#### 4.6.1 Máquina de Vetores de Suporte Quântica

A Máquina de Vetores de Suporte (SVM, do inglês Support Vector Machine) é um algoritmo de aprendizado supervisionado utilizado para tarefas de classificação e regressão (HEARST et al., 1998). Na classificação, a SVM funciona encontrando o hiperplano que melhor separa diferentes classes no espaço de características, maximizando a margem entre elas. Isso envolve buscar a maior distância possível entre o hiperplano separador e os pontos de dados mais próximos de cada classe, conhecidos como vetores de suporte (CORTES; VAPNIK, 1995; BURGES, 1998). Essa abordagem melhora a capacidade do modelo de generalizar de forma eficaz para novos dados não vistos.

Quando combinado com métodos de kernel, a SVM pode lidar efetivamente com dados que não são linearmente separáveis. Métodos de kernel mapeiam os dados de entrada para um espaço de maior dimensão onde a separação linear se torna viável. O kernel mais comumente utilizado é o kernel de base radial (RBF, do inglês Radial Basis Function), definido como:

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (81)$$

onde  $\gamma$  é um parâmetro do kernel que controla a influência de cada ponto de dados (BURGES, 1998). Essa técnica permite que a SVM crie fronteiras de decisão complexas adequadas para dados de alta dimensionalidade e problemas não lineares. Ao usar kernels, o problema original de separação linear no espaço de entrada é transformado em um problema de separação linear no espaço de características de maior dimensão, permitindo separações não lineares no espaço original. Uma QSVM (Máquina de Vetores de Suporte Quântica) estende os métodos clássicos de SVM ao aproveitar o *truque do kernel* quântico, onde os dados são mapeados para um espaço de Hilbert de maior dimensão por meio de estados quânticos, permitindo cálculos de kernel mais eficientes.

## 4.7 Perspectivas Futuras

Perspectivas futuras no campo do Quantum Machine Learning (QML) são promissoras, impulsionadas pelos avanços contínuos na tecnologia quântica. A proximidade da disponibilidade de pequenos computadores quânticos com 50 a 100 qubits e processadores quânticos

especializados, como simuladores e *annealers*, abre oportunidades em aprendizado de máquina e análise de dados. No entanto, desafios técnicos significativos persistem. A eficiência na leitura de dados é uma preocupação que pode se tornar um fator limitante. A obtenção de soluções completas em algoritmos quânticos é complexa, demandando estratégias que se concentrem em estatísticas resumidas. Além disso, a complexidade computacional exata desses algoritmos ainda não é bem compreendida. Estabelecer benchmarks sólidos para comparar o desempenho quântico com abordagens clássicas é um desafio essencial. Uma abordagem inovadora envolve a aplicação do QML a dados quânticos, aprimorando a caracterização e o controle dos computadores quânticos (BIAMONTE et al., 2017).

Por outro lado, pesquisadores defendem que a comparação direta entre os clássicos e quânticos e a busca desenfreada apenas pela vantagem quântica pode ser danosa para a academia. Neste caso, em vez de focar na busca por vantagens quânticas, a pesquisa futura deveria explorar questões mais fundamentais e considerar problemas complexos que possam ser abordados com computadores quânticos. Essa abordagem requer “coragem” para resistir às narrativas de “supremacia quântica” e termos chamativos. O futuro da pesquisa em aprendizado de máquina quântica pode se beneficiar ao priorizar questionamentos essenciais e explorar novos caminhos à medida que tecnologias quânticas avançam (SCHULD; KILLORAN, 2022).

Em resumo, as perspectivas futuras no campo do Quantum Machine Learning (QML) estão repletas de promessas e desafios técnicos significativos. À medida que a disponibilidade de hardware quântico se torna uma realidade próxima, a busca por soluções inovadoras e estratégias para enfrentar os obstáculos mencionados promete moldar o cenário do QML. Com inovação, podemos desvendar o potencial do QML, conduzindo-o em direção a futuras aplicações industriais em larga escala.

# 5 Classificando Fases Quânticas

No presente capítulo, investigamos a classificação de fases quânticas por meio de três abordagens. A primeira abordagem, apresentada na Seção 5.1, enfatiza a possibilidade de classificar fases utilizando algoritmos de ML com fragmentos do próprio modelo, em seguida escalando para o conceito de independência de modelo. Demonstra-se como um classificador treinado com dados de um modelo de Hamiltoniano pode ser utilizado para classificar fases de outro modelo distinto, evidenciando a generalização dos métodos aplicados. Em 5.2, abordamos a possibilidade de classificar tais fases em contextos nos quais a disponibilidade de dados é escassa, onde demonstramos que é possível classificar as fases em modelos distintos usando quantidades muito reduzidas de dados. Introduzimos o uso de redes generativas adversárias (GANs) para criação de dados sintéticos, essa abordagem permite ampliar e diversificar os conjuntos de dados, melhorando significativamente a eficiência e a precisão das classificações. Por fim, na Seção 5.3, fazemos uso de algoritmos de Aprendizado de Máquina Quânticos (QML) aliados a técnicas de seleção de atributos baseadas no método SHAP (Shapley Additive Explanations). Essa abordagem visa otimizar a escolha de features relevantes, maximizando a precisão dos classificadores e minimizando os recursos computacionais necessários. Cada uma dessas seções reflete contribuições individuais apresentadas em artigos publicados ou submetidos/ em produção, oferecendo uma visão abrangente e integrada das técnicas de classificação de fases quânticas.

## 5.1 Classificador Independente de Modelo

Esta seção de aplicação é baseada no artigo:

- Mahlow, F., Luiz, F.S., Malvezzi, A.L. et al. Model-independent quantum phases classifier. *Sci Rep* 13, 14411 (2023). <https://doi.org/10.1038/s41598-023-33301-0>

Matéria condensada é um ramo da física que estuda sistemas com muitas partículas em estados condensados (WEN, 2007). Uma das tarefas centrais e mais desafiadoras do campo é entender os diagramas de fase de sistemas quânticos de muitos corpos correlacionados (YANG et al., 2021). As fases contidas no paradigma de Landau geralmente requerem o conhecimento dos parâmetros de ordem, e aquelas além apresentam o obstáculo de encontrar os parâmetros adequados para que tais fases sejam devidamente representadas (SCHUCH; PÉREZ-GARCÍA; CIRAC, 2011; WEN, 1989; GINZBURG; LANDAU, 2009). Os sistemas quânticos apresentam um comportamento extremamente complexo em relação aos clássicos, exibindo fases com ordem topológica que não podem ser descritas pelo paradigma de quebra de simetria local de Landau (SCHUCH; PÉREZ-GARCÍA; CIRAC, 2011). Além disso, principalmente devido à complexidade

computacional no estudo de sistemas quânticos de muitos corpos e à falta de um *framework* de classificação (HASTINGS, 2013; YOSHIDA, 2011), classificar as fases e identificar suas respectivas transições é uma tarefa difícil, que depende de características únicas do sistema em questão. Em relação ao trabalho computacional, o aumento da complexidade do sistema, ou seja, do número de partículas estudadas, gera a conhecida "Maldição da Dimensionalidade", pois o espaço de Hilbert cresce exponencialmente com o tamanho do problema.

No que diz respeito à classificação das fases quânticas, muitas alternativas têm sido apontadas como promissoras (HASTINGS, 2013; SCHUCH; PÉREZ-GARCÍA; CIRAC, 2011; CHEN; GU; WEN, 2011) com especial ênfase no aprendizado de máquina ("*Machine Learning*" - ML), que vem ganhando destaque não apenas nesse quesito, mas em um amplo espectro de avanços científicos. Em (CANABARRO et al., 2019), estudou-se a possibilidade de fazer previsões usando ML no modelo Axial (ou Anisotrópico) Next-Nearest Neighbor Ising (ANNNI) Spin = 1/2, obtendo boas taxas de sucesso nessas tarefas. Aqui, damos um passo adiante, usando fragmentos aleatórios de parte do diagrama de fase hamiltoniano em estado fundamental para treinar algoritmos a fim de classificar o resto do diagrama de fase em si, estudando a relação entre quantidade de dados usados e precisão. Para esta tarefa, utilizaremos 4 algoritmos amplamente conhecidos na literatura, são eles: K-Nearest Neighbours ( $k$ -NN), Multi-Layer Perceptron (MPL), Random Forest (RF) e Gradient Boosting (GB). Em seguida, definiremos o conceito de Classificador de Fase Quântica Independente de Modelo. Usando o algoritmo K-Nearest Neighbours ( $k$ -NN), um dos mais simples e intuitivos algoritmos de classificação não probabilística (DUDANI, 1976), mostramos que as fases de um modelo específico podem ser determinadas a partir do conhecimento obtido de outros modelos distintos. Em outras palavras, o algoritmo é capaz de detectar as fases mesmo sem qualquer conhecimento prévio sobre o hamiltoniano, o que sugere que os ingredientes-chave de cada fase estão no estado quântico e não no modelo específico.

Este capítulo trata-se de uma versão estendida de (MAHLOW et al., 2023a), explorando os detalhes e análises mais amplas do que as contidas no artigo. O capítulo está organizado da seguinte forma. Inicialmente, em 5.1.1, é dada uma introdução teórica sobre o conceito de fases quânticas. Na seção 5.1.2 apresentaremos os três Hamiltonianos resolvidos na elaboração deste trabalho. Na seção 5.1.3, descrevemos o processo realizado para gerar os bancos de dados que serão utilizados. Em seguida, apresentaremos os resultados obtidos pelo treinamento dos modelos com elementos aleatórios, variando a quantidade de dados utilizados no treinamento. Em seguida, apresentaremos nossos resultados para a aprendizagem por transferência, ou seja, treinamento em um primeiro modelo e teste em um segundo desconhecido. Por fim, concluímos o trabalho, resumindo e discutindo o impacto dos resultados obtidos.

### 5.1.1 Transições de Fase

A física da matéria condensada é uma área fascinante, apresentando problemas na fronteira do conhecimento. É uma ciência na qual a teoria e práticas experimentais caminham em conjunto, permitindo que novos materiais sejam moldados de acordo com as necessidades da sociedade, trazendo, por consequência, grandes progressos tecnológicos. Algo essencial no estudo da matéria condensada é a detecção das transições de fase quânticas, assim como a representação do diagrama de fase de determinados materiais, a fim de observar propriedades importantes destes. Definições destas transições serão abordadas mais detalhadamente nas seções a seguir. (CONTINENTINO, 2006).

#### 5.1.1.1 Transições de Fase Clássicas

A presença de transições de fase é extremamente comum em nosso cotidiano. Ao se ferver a água para o café, observa-se parte dela evaporando, caracterizando uma transição da fase líquida para a gasosa. Ao esquecer e acidentalmente deixar uma bebida tempo demais no congelador, presencia-se uma transição da fase líquida para a sólida. Como pode-se observar, os dois exemplos citados, tratam-se de transições de fase térmicas, onde as características macroscópicas do sistema em questão, são alteradas quando a temperatura  $T$  se aproxima de uma determinada temperatura crítica  $T_c$ . (BRUM, 2015). Um outro exemplo de transição de fase, seria a desmagnetização de um fragmento de ferro. Este fragmento, inicialmente em fase ferromagnética, dada pelo ordenamento dos spins de seus átomos, ao ser submetido à um aumento significativo de temperatura (até aproximadamente 1043 K), deixa de exibir este comportamento, transitando para a fase paramagnética. *"Formalmente, uma transição de fase é a alteração de propriedades termodinâmicas de um material causada pela variação de um parâmetro externo, como por exemplo temperatura, pressão ou campo magnético"*. (OLIVEIRA, 2014).

Mesmo ambas se tratando de transições de fase clássicas, podemos caracteriza-las em ordens diferentes de transição, de acordo com suas características. O conceito é explicado por Oliveira a seguir.

Quanto à fenomenologia, as transições de primeira ordem são aquelas onde há coexistência de fases na temperatura crítica, sendo necessário uma quantidade finita de calor latente para completar a transição. Este é o território ao qual pertence a transição do gelo. Nas de segunda ordem, não se tem esta coexistência. Na verdade, tais fases tornam-se indistinguíveis, e a transição se dá de forma contínua, não havendo calor latente. Aqui se encontra a desmagnetização do ferro. (OLIVEIRA, 2014).

Sabendo que as transições mencionadas ocorreram devido à mudanças na temperatura dos elementos citados, somos levados à seguinte questão: supondo que diminuíssemos a temperatura até um limite em que esta tenderia a zero, ainda seria possível observar alguma transição de fase?

### 5.1.1.2 Transições de Fase Quânticas

Enquanto a temperatura tendesse à zero ( $T \rightarrow 0$ ), as flutuações térmicas tenderiam à zero e, portanto, não ocorrerão mais transições de fase térmicas. Entretanto, ao nos aproximarmos de  $0K$ , a mecânica quântica, que governa o mundo microscópico passa a exercer um papel crucial, onde, o Princípio da Incerteza de Heisenberg, indica a ainda existente presença de flutuações, mas agora de ordem quântica, que por consequência, podem levar à transições de fase. Portanto, pode-se dizer que há um “movimento” à baixíssimas temperaturas e este, permite que outros parâmetros, como a sua densidade de elétrons, a pressão aplicada à um certo cristal ou a força de um campo magnético possam gerar transições de fase. Estas transições de fase podem ocasionar mudanças significativas no estado macroscópico de determinados materiais. (SACHDEV, 2011).

Embora possa parecer que o problema seja específico, do domínio da física de temperaturas muito baixas (afinal,  $T = 0K$  não é possível, rigorosamente), resultados experimentais e teóricos mostraram a existência de pontos críticos quânticos [...], que quando existentes, afetam o comportamento do sistemas condensando a temperaturas finitas. (BRUM, 2015).

Como indicado, as transições de fase quânticas podem ter consequências importantes do ponto de vista prático, de forma que a identificação de características e padrões presentes em seus diagramas de fase se torne um aspecto relevante a ser estudado.

## 5.1.2 Os Três Modelos

Nesta seção, apresentamos os três Hamiltonianos estudados, descrevendo as fases quânticas contidas em cada diagrama de fase e suas principais características: Cadeias XXZ com Anisotropia do Tipo de Íon Único Uniaxial, Cadeias XXZ Alternadas Ligadas e o modelo Biquadrático Bilinear. Em seguida, resumimos as fases contidas nos diagramas, destacando as fases distintas e comuns entre os modelos, que foram posteriormente analisadas por algoritmos de Aprendizado de Máquina.

O primeiro modelo estudado foi o Hamiltoniano das Cadeias XXZ com Anisotropia do Tipo de Íon Único Uniaxial, descrito por:

$$\mathcal{H}_1 = \sum_{l=1}^N [J(S_l^x S_{l+1}^x + S_l^y S_{l+1}^y) + J_z S_l^z S_{l+1}^z] + D \sum_{l=1}^N S_l^{z2} \quad (82)$$

Neste modelo,  $S_l$  é um operador de spin-1,  $D$  representa a anisotropia de íon único uniaxial e  $N$  é o número total de spins na cadeia. Foi definido  $J = 1$  para corrigir a escala de energia. O diagrama de fases do estado fundamental inclui as fases: Haldane, Large-D, duas fases XY, Ferromagnética e uma fase Néel. Entre essas fases, ocorrem diversos tipos de transições (CHEN; HIDA; SANCTUARY, 2003).

O segundo modelo é o das Cadeias XXZ Alternadas Ligadas, descrito pelo Hamiltoniano:

$$\mathcal{H}_2 = \sum_{l=1}^N [1 - \delta(-1)^l] [S_l^x S_{l+1}^x + S_l^y S_{l+1}^y + \Delta S_l^z S_{l+1}^z] \quad (83)$$

Neste caso,  $\Delta$  é a intensidade da anisotropia do tipo Ising, associada à interação spin-órbita em materiais magnéticos, e  $\delta$  representa a alternância de ligações que descreve a dimerização. O diagrama de fases do modelo inclui as fases Ferromagnética, XY1, Néel, Haldane e Dimerizada (KITAZAWA; NOMURA; OKAMOTO, 1996).

O terceiro modelo é o Biquadrático Bilinear, descrito pelo Hamiltoniano:

$$\mathcal{H}_3 = \sum_{l=1}^N [\cos \theta (S_l \cdot S_{l+1}) + \sin \theta (S_l \cdot S_{l+1})^2] \quad (84)$$

Neste modelo,  $\theta \in [0, 2\pi)$  é o parâmetro que quantifica o acoplamento entre spins vizinhos mais próximos. O diagrama de fases é composto pelas fases Haldane, Trimerizada, Ferromagnética e Dimerizada (LÄUCHLI; SCHMID; TREBST, 2006).

Para resumir e ilustrar as fases contidas nos diagramas dos três Hamiltonianos, apresentamos a Tabela 1. Como se observa, todas as fases de  $\mathcal{H}_2$  estão contidas nas fases combinadas de  $\mathcal{H}_1$  e  $\mathcal{H}_3$ , sendo o único modelo com essa propriedade. A combinação de  $\mathcal{H}_1$  e  $\mathcal{H}_2$  abrange três das quatro fases de  $\mathcal{H}_3$ , enquanto a união de  $\mathcal{H}_2$  e  $\mathcal{H}_3$  inclui quatro das seis fases de  $\mathcal{H}_1$ . No total, cinco fases são compartilhadas por pelo menos dois modelos (Haldane, Néel, Ferromagnética, XY1 e Dímero) e três fases são exclusivas de um único modelo (Large-D, XY2 e Trimer). Os números entre parênteses indicam a quantidade total de dados utilizada em cada fase para os respectivos Hamiltonianos.

Quantum Phase	$\mathcal{H}_1$	$\mathcal{H}_2$	$\mathcal{H}_3$
Haldane	●(302)	●(467)	●(500)
Néel	●(2043)	●(1265)	
Ferromagnetic	●(2146)	●(800)	●(750)
Large-D	●(1665)		
XY1	●(214)	●(906)	
XY2	●(30)		
Dimer		●(2882)	●(500)
Trimer			●(250)

Tabela 1 – Fases contidas nos diagramas correspondentes aos três Hamiltonianos analisados. Os círculos preenchidos indicam que a fase quântica está presente no modelo. Os números sobrescritos nos círculos representam a quantidade de dados calculados para cada fase de cada modelo.

### 5.1.3 Resultados

Nesta seção, abordaremos os resultados obtidos nas tentativas de classificar os diagramas de fase Hamiltonianos descritos na seção 5.1.2. Para tal, iniciaremos em 5.1.3.1 com uma descrição detalhada da metodologia utilizada para gerar os bancos de dados utilizados, a fim de treinar os algoritmos de aprendizado de máquina. Em 5.1.3.2, apresentamos a abordagem utilizando fragmentos aleatórios do próprio modelo como um conjunto de treinamento, realizando então uma análise da relação entre a quantidade de dados e a precisão obtida. Por fim, em 5.1.3.3 expomos os resultados obtidos quanto às tentativas de classificação das fases em um segundo modelo, utilizando dados apenas dos outros dois, ou seja, sem conhecer o hamiltoniano em questão.

#### 5.1.3.1 Os Bancos de Dados

A ideia central por trás dos algoritmos de ML é definir, por meio dos exemplos dados, uma função que mapeia as entradas para as saídas desejadas (HASTIE; TIBSHIRANI; FRIEDMAN, 2009). Os exemplos, em nosso caso, são um grande conjunto de dados criado com base nos hamiltonianos dados pela Eq. (82), Eq. (83) e Eq. (84). Para gerar o conjunto de dados de correlação, consideramos milhares de valores diferentes para os parâmetros dos Hamiltonianos  $\mathcal{H}_1$ ,  $\mathcal{H}_2$  e  $\mathcal{H}_3$ . Para  $\mathcal{H}_1$ , Eq. 82, variamos os parâmetros  $J_z$  e  $D$  no intervalo  $[-4, 4]$  com um passo de 0.1, o que gera um conjunto de dados com 6400 pontos. Para o Hamiltoniano  $\mathcal{H}_2$ , Eq. 83, variamos os parâmetros,  $\Delta$  em um intervalo  $[0, 1]$ , e os parâmetros  $\delta$  em um intervalo  $[-1.5, 2.5]$ , com tamanhos de passo de 0.05 e 0.0125, respectivamente, o que gera um conjunto de dados com 6320 pontos. Finalmente, para o Hamiltoniano  $\mathcal{H}_3$ , Eq. 84, definimos o intervalo do parâmetro  $\theta$  no intervalo  $[0, 2\pi]$ , de forma igualmente espaçada resultando em um conjunto de dados com 4520 pontos. Os resultados apresentados foram obtidos por meio da diagonalização exata dos Hamiltonianos. Os rótulos das fases para os Hamiltonianos  $\mathcal{H}_1$ ,  $\mathcal{H}_2$  e  $\mathcal{H}_3$  são obtidos da literatura (CHEN; HIDA; SANCTUARY, 2003), (KITAZAWA; NOMURA; OKAMOTO, 1996) e (LÄUCHLI; SCHMID; TREBST, 2006), respectivamente.

Embora possível, não é desejável usar todas as variáveis que definem o estado fundamental como entrada para o algoritmo de ML. Na verdade, se fosse esse o caso, o número de entradas (também chamadas de *features*) é da ordem de  $3^N$  para uma cadeia de  $N$  spins-1. Assim, dada esta enorme quantidade de recursos, surge uma questão natural: para cada estado fundamental, como obter um conjunto menor de entradas que definem sua fase quântica? Considerando os hamiltonianos dados pela Eq. (82), Eq. (83) e Eq. (84), é natural pensar que correlações entre vizinhos podem fornecer informações relevantes sobre cada fase quântica. Nesse sentido, em vez de usar variáveis  $3^N$  como características, tomamos como entrada os valores esperados dos seguintes observáveis:  $\langle \sigma_i^x, \sigma_j^x \rangle$ ,  $\langle \sigma_i^y, \sigma_j^y \rangle$ ,  $\langle \sigma_i^z, \sigma_j^z \rangle$  com  $j > i$  e  $i = [1, N - 1]$ , onde  $N = 12$  é o número de spins e  $\langle \sigma_i^k, \sigma_j^k \rangle = \langle \lambda_0 | \sigma_i^k, \sigma_j^k | \lambda_0 \rangle$ , para  $k = \{x, y, z\}$ , são o valor esperado da correlação para o estado de menor energia do hamiltoniano (CANABARRO et al.,

2019). Além disso, uma vez que a cadeia é fechada e, conseqüentemente, as propriedades da cadeia são cíclicas, qualquer correlação não redundante entre dois spins é obtida para  $\langle S_1^k S_i^k \rangle$  com  $i = [1, N/2 + 1]$ . A escolha desses recursos é justificada porque, enquanto os correlatores de dois pontos identificam ordens locais simples, os correlatores de strings capturam ordens não locais, o que é importante para a generalização do método.

### 5.1.3.2 Método de Generalização

Os bancos de dados usados nesta etapa são fragmentos aleatórios dos diagramas de fase do modelo a ser testado. Eles foram obtidos usando o pacote `train_test_split` da biblioteca `sklearn.model_selection` em Python (PEDREGOSA et al., 2011), então o procedimento foi reiterado com pequenas variações na razão entre o treinamento e os dados de teste necessários. Nosso objetivo, portanto, é observar não apenas a habilidade dos algoritmos em classificar as fases quânticas estudadas, mas também observar como as quantidades de dados utilizadas influenciam o desempenho dos classificadores.

Para o diagrama de fase de  $\mathcal{H}_1$ , indicado na Fig. 5, 10% do diagrama foi usado, o que corresponde a 640 dados rotulados, de um total de 6400. Como pode ser visto, os quatro algoritmos discutidos até agora foram usados para realizar as tentativas de classificação das fases.

Os algoritmos avaliados demonstraram boa capacidade de descrever o diagrama de fases com elevada precisão, mesmo utilizando um conjunto de dados reduzido. Entre os modelos testados, o Random Forest (RF) obteve o melhor desempenho, com uma precisão final de 98.83%, seguido pelo Multi-Layer Perceptron (MLP), que alcançou 98.23%, e pelo K-Nearest Neighbors (KNN), com 98.03%. O Support Vector Machine (SVM) apresentou o menor desempenho, com uma precisão de 96.14%, evidenciando dificuldades em classificar a fase de Haldane. Embora os resultados possam ser melhorados por meio de uma seleção mais criteriosa dos dados ou refinamento dos hiperparâmetros dos modelos, os algoritmos foram intencionalmente aplicados com as configurações padrão da biblioteca Scikit-learn (PEDREGOSA et al., 2011), de forma a manter os métodos o mais simples possível. Esses resultados são encorajadores e indicam que é possível generalizar a descrição dos diagramas de fases utilizando técnicas de aprendizado de máquina, mesmo em cenários com dados limitados.

A seguir, analisamos a classificação do diagrama de fases de  $\mathcal{H}_2$ . Para o experimento representado na Fig. 6, foram utilizados 632 dados selecionados aleatoriamente, correspondendo a 10% do total de 6320 pontos disponíveis no diagrama de fases. Entre os algoritmos avaliados, o Random Forest (RF) apresentou o melhor desempenho, com uma precisão final de 99.20%, seguido pelo K-Nearest Neighbors (KNN), que obteve 98.48%, e pelo Multi-Layer Perceptron (MLP), com 96.69%. O Support Vector Machine (SVM) novamente teve o menor desempenho, alcançando 88.56%, evidenciando dificuldades significativas na identificação da fase de Haldane. Ainda assim, todos os algoritmos demonstraram excelente capacidade de previsão geral, mesmo

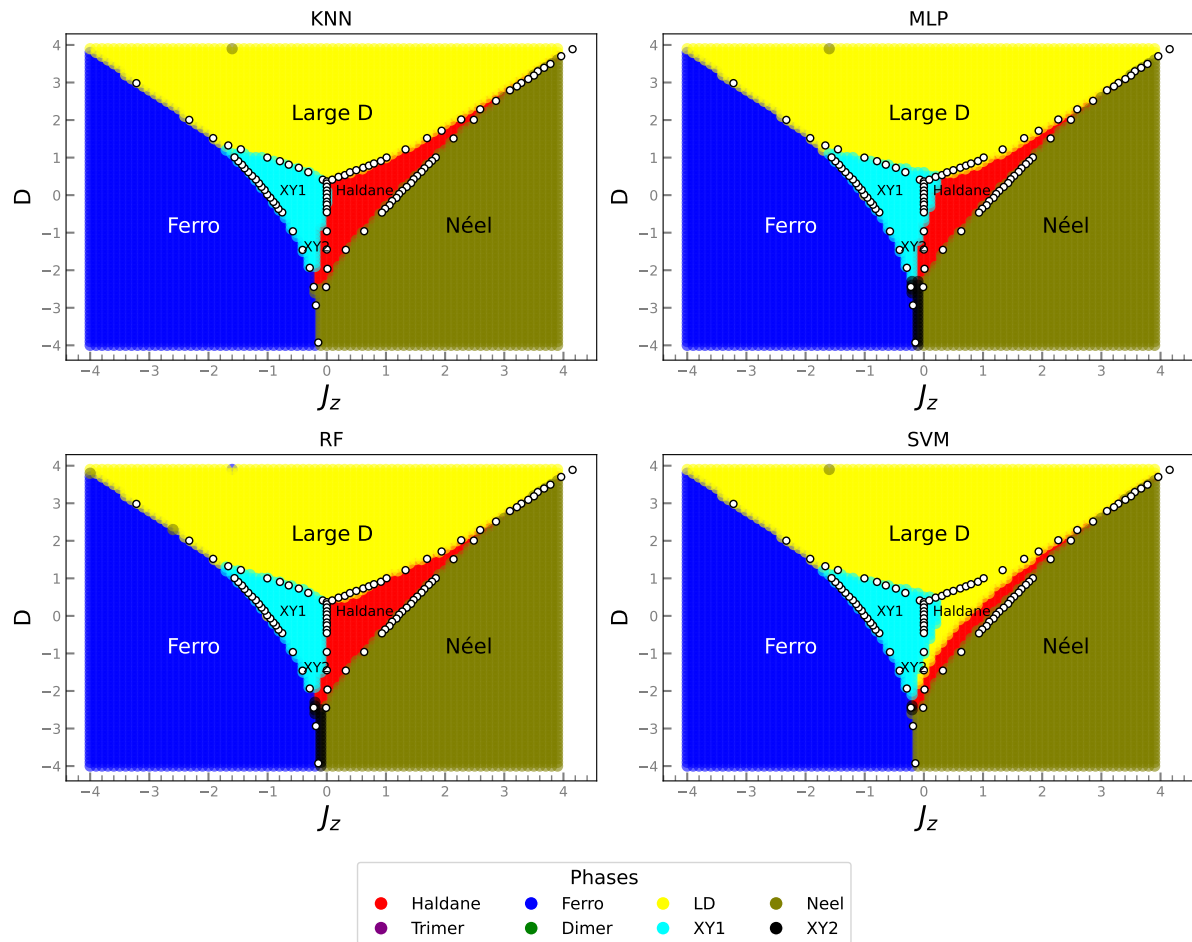


Figura 5 – Tentativas dos algoritmos de classificar o diagrama de fases de  $\mathcal{H}_1$  usando dados fragmentados do próprio modelo.

utilizando apenas uma pequena fração dos dados disponíveis.

No modelo Biquadrático Bilinear, representado na Fig. 7, foram utilizados 10% dos dados disponíveis, correspondendo a 452 amostras de um total de 4520. Todos os algoritmos apresentaram excelente desempenho, com destaque para o Random Forest (RF), que alcançou uma média de acurácia de 99.6%, seguido pelo K-Nearest Neighbors (KNN) com 99.4%, e pelo Multi-Layer Perceptron (MLP), que obteve 99.3%. O Support Vector Machine (SVM) teve o menor desempenho entre os algoritmos avaliados, mas ainda assim alcançou uma média de 97.6%. Os resultados superiores observados neste modelo podem ser atribuídos à natureza unidimensional do problema, em contraste com os casos bidimensionais analisados anteriormente, além da menor quantidade de dados a classificar, o que gerou contornos mais bem definidos e altas taxas de sucesso nos algoritmos utilizados.

A seguir, a Fig. 8 indica a relação entre a precisão obtida pelos quatro algoritmos de Aprendizado de Máquina utilizados em relação à quantidade de dados em suas amostras de treinamento.

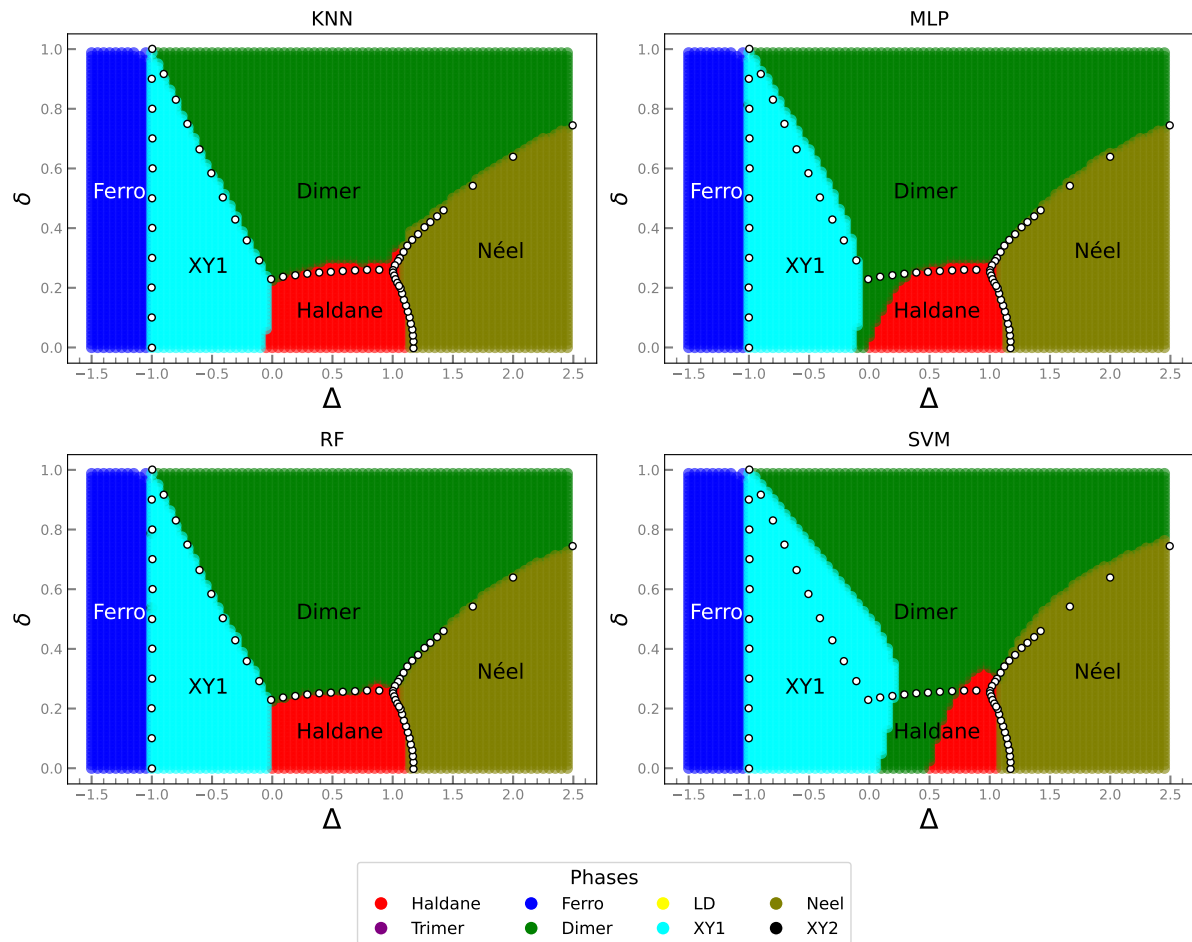


Figura 6 – Tentativas dos algoritmos de classificar o diagrama de fases de  $\mathcal{H}_2$  usando dados fragmentados do próprio modelo.

Observa-se que na figura 8 a), referente a  $\mathcal{H}_1$ , os algoritmos têm um desempenho semelhante, onde o SVM fica um pouco abaixo dos demais quando a quantidade de dados é baixa. Após a marca de aproximadamente 250 pontos ( $\approx 4\%$  do banco de dados), todos os demais têm precisões superiores a 95%. Já para  $\mathcal{H}_2$ , representado em 8 b), o SVM também apresenta um desempenho inferior aos demais, porém menos acentuado. Também há um desempenho inferior das redes neurais (comparado ao KNN e RF) ao longo do diagrama, sendo que, em aproximadamente 300 dados ( $\approx 5\%$ ), KNN, MLP e RF já apresentavam mais de 90% de precisão, enquanto o MLP estavam na zona de 90%. Por fim, para  $\mathcal{H}_3$  em 8 c), observamos um comportamento mais consistente de todos os algoritmos utilizados, permitindo uma taxa de sucesso na generalização dos diagramas de quase 100% a partir de aproximadamente 500 pontos. O estudo dessa relação entre a quantidade de dados utilizados e a precisão dos algoritmos permitiu observar que mesmo com uma quantidade muito pequena (da ordem de dezenas de dados), alguns modelos já eram capazes de generalizar todo o diagrama de fases do sistema com alta precisão, acima de 90% de exatidão, algo notável, que indica o poder

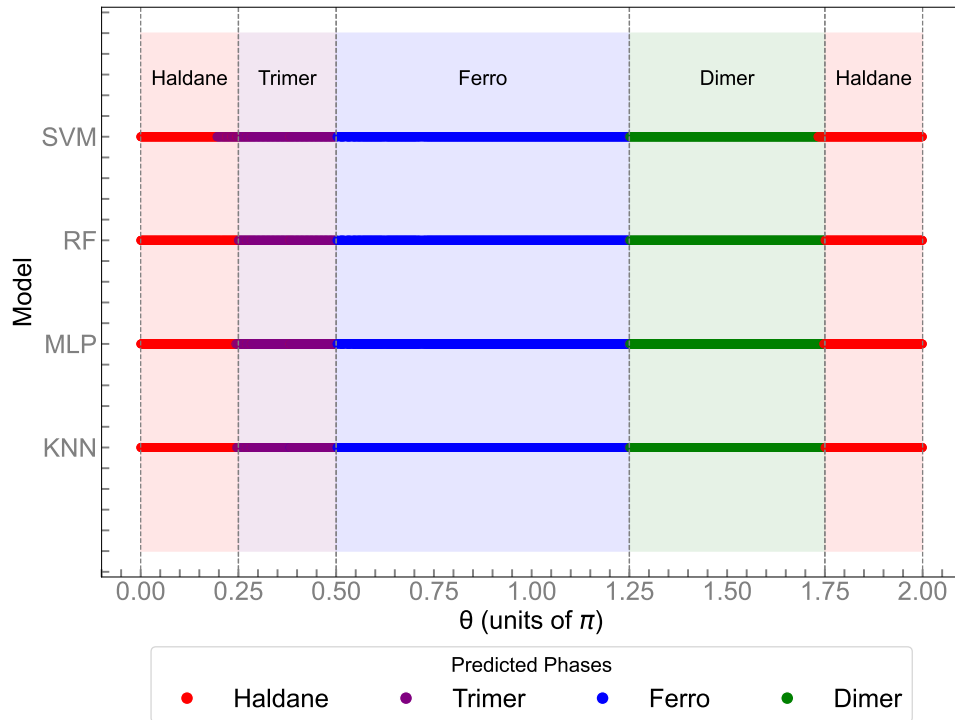


Figura 7 – Tentativas dos algoritmos de classificar o diagrama de fases de  $\mathcal{H}_3$  usando dados fragmentados do próprio modelo.

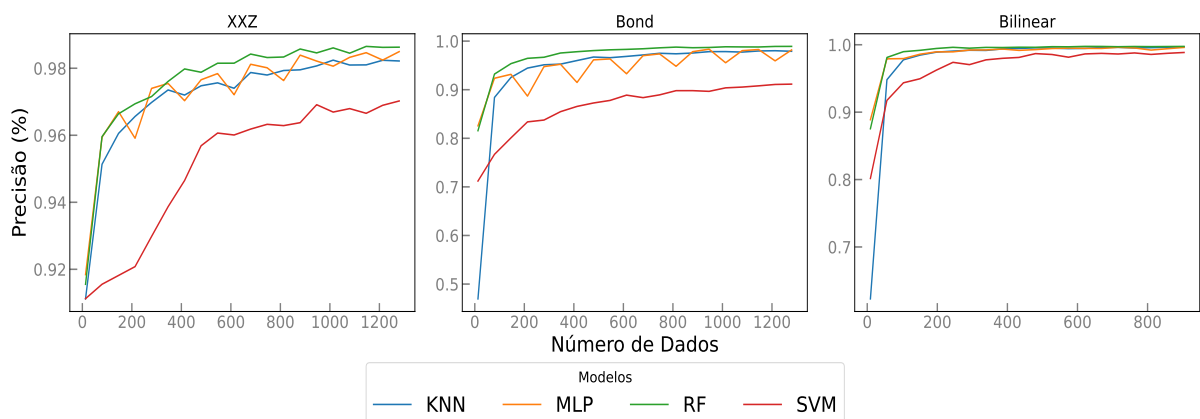


Figura 8 – Estudo da relação entre a quantidade de dados utilizados no conjunto de treinamento versus a precisão obtida pelos quatro classificadores analisados. **a)** Refere-se aos resultados para  $\mathcal{H}_2$ , **b)** é a precisão obtida na análise para  $\mathcal{H}_2$  e **c)** apresenta essa relação para  $\mathcal{H}_3$ .

preditivo dos métodos supervisionados de Aprendizado de Máquina. Esse ponto será central na seção seguinte, onde trataremos do próximo artigo.

### 5.1.3.3 Independência de Modelo

Após os resultados positivos obtidos na subseção anterior, procedemos com uma proposta mais ousada a seguir. Isso seria, conhecendo a possibilidade de classificar diagramas de fases a partir de fragmentos do próprio modelo, investigar a possibilidade de utilizar os bancos de dados de fases de modelos conhecidos para classificar novos diagramas totalmente desconhecidos. Para testar essa hipótese, vamos verificar a acurácia do classificador  $k$ -NN na tarefa desejada. O  $k$ -NN foi escolhido, por sua excelente precisão nas tarefas anteriores, aliado ao fato de ser um modelo de funcionamento simples, o que é desejável para facilitar a compreensão de físicos que possivelmente possam ler o trabalho e não estão familiarizados com algoritmos mais sofisticados de aprendizagem de máquina. Inicialmente, analisamos sua capacidade de redefinir as fases usando apenas um dos modelos como banco de dados de treino, para depois testá-lo no segundo. Em um segundo momento, o comportamento foi estudado por meio de 2 modelos para treinamento e um para teste.

Com o conjunto de dados dos três modelos, pudemos visualizar a relação das correlações com as fases quânticas. Como pretendemos usar um algoritmo classificador, a ideia é separar, no espaço multidimensional, fases distintas em posições distintas. Para 12 sítios, por exemplo, usando o conjunto de correlações descrito acima, temos 24 correlações distintas (um espaço com 24 dimensões) e visualizar essa quantidade de informações em um espaço 3D é complicado. A fim de obter insights sobre o comportamento dos dados no espaço de 24 dimensões, empregamos o algoritmo de Análise de Componentes Principais (PCA) para reduzir o conjunto de dados para duas dimensões, possibilitando a visualização em um plano. Na Figura (9), mostramos os dois componentes obtidos usando o método. Com o  $PCA_1$  no eixo  $X$  e o  $PCA_2$  no eixo  $Y$ , procurando a melhor representação gráfica do que acontece no espaço de correlação. Observe que  $PCA_1$  e  $PCA_2$  são calculados para todo o conjunto de dados, onde consideramos diferentes valores para os parâmetros dos Hamiltonianos  $\mathcal{H}_1$ ,  $\mathcal{H}_2$  e  $\mathcal{H}_3$ .

Um aspecto importante a ser considerado ao trabalhar com aprendizado de máquina é a transformação (ou normalização) do conjunto de dados. Em muitos casos, transformações apropriadas podem separar as classes no espaço de características, o que facilita o processo de classificação. Na Figura (9-a) a Figura (9-d) plotamos os dados brutos e na Figura (9-e) a Figura (9-h) plotamos os dados brutos após escalarmos o conjunto de dados para ter uma norma unitária, para cada modelo. Este é um procedimento de renormalização bem conhecido chamado pré-processamento de sinal espacial (SERNEELS; NOLF; ESPEN, 2006). Analisando a Figura (9), vemos que mesmo com as informações contidas em um único par de correlações, somos capazes de separar estados com diferentes fases com alta nitidez. De fato, há alguma sobreposição dessas informações para diferentes fases, especialmente ao considerar vários

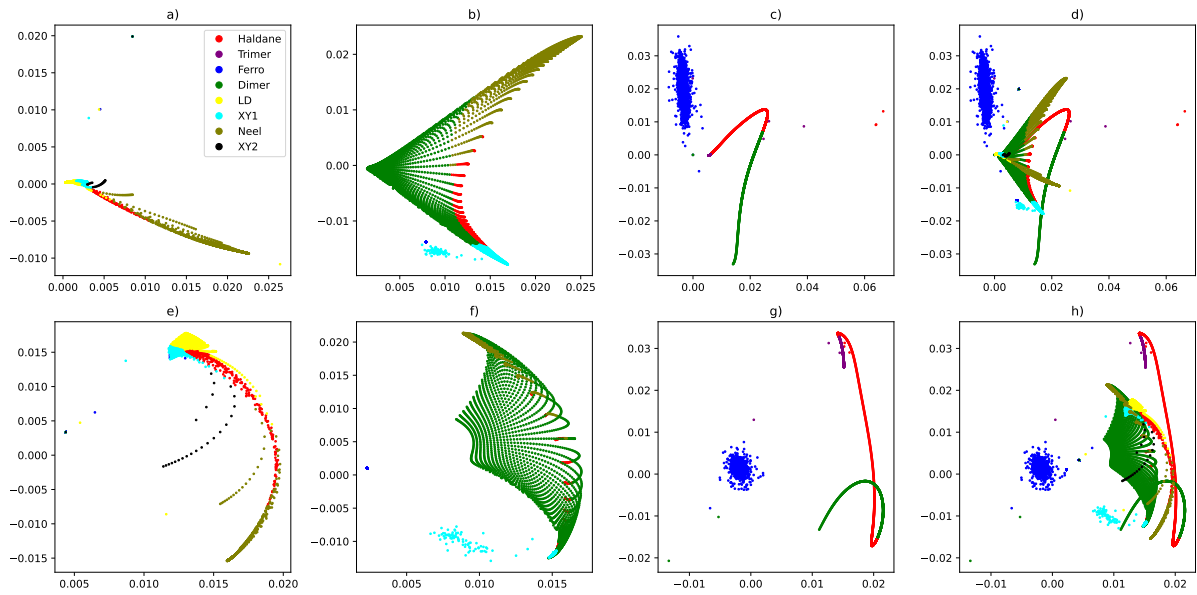


Figura 9 – Componente PCA 1 vs. Componente PCA 2, a cor representa a fase do ponto de dados correspondente. E as figuras (a) e (e) referem-se ao sistema descrito por  $\mathcal{H}_1$ , (b) e (f) referem-se a  $\mathcal{H}_2$ , (c) e (g) referem-se a  $\mathcal{H}_3$ , (d) e (h) são os dados dos três sistemas. As figuras (a), (b), (c) e (d) são os dados brutos, e (e), (f), (g), (h) são os dados após a realização da transformação.

modelos diferentes, Figura (9-d) e Figura (9-h). Além disso, para os dados normalizados, Figura (9-e) a Figura (9-h), a sobreposição de informações sobre as fases diminui para todos os modelos, quase desaparecendo para  $\mathcal{H}_3$ , Figura (9-g).

Apesar da melhora na redução da sobreposição, a Figura (9-h) mostra que ainda existem sobreposições de informações da fase Haldane com Large-D e das fases XY1, XY2, Néel e Haldane com a fase Dimer. No entanto, é importante observar que estamos ilustrando o procedimento considerando apenas dois recursos, o que é insuficiente para separar todas as fases, dada a complexidade do problema. Na verdade, o uso do conjunto completo de recursos, bem como o estudo e a aplicação de outras transformações para separar as fases no espaço de correlação, torna-se crucial e é o foco deste trabalho.

Para testar a precisão do algoritmo  $k$ -NN na transferência de aprendizado dos modelos conhecidos para o modelo desconhecido, começamos treinando o algoritmo para classificar os pontos de dados de  $\mathcal{H}_1$ . Para esse caso, treinamos o  $k$ -NN com dados dos modelos  $\mathcal{H}_2$  e  $\mathcal{H}_3$  (ou seja, o algoritmo  $k$ -NN conhece o conjunto de dados de  $\mathcal{H}_2$  e  $\mathcal{H}_3$  e seus rótulos de fases respectivos) e previmos as fases de  $\mathcal{H}_1$ . Em sequência, fazemos o mesmo para  $\mathcal{H}_2$  e  $\mathcal{H}_3$  usando os dados dos modelos  $\mathcal{H}_1$  e  $\mathcal{H}_3$ , e  $\mathcal{H}_1$  e  $\mathcal{H}_2$ , respectivamente.

No entanto, vale ressaltar que existem fases em  $\mathcal{H}_1$  (Large-D e XY2) que não existem em  $\mathcal{H}_2$  e  $\mathcal{H}_3$ , e uma fase em  $\mathcal{H}_3$  (Trimer) que não existe em  $\mathcal{H}_1$  e  $\mathcal{H}_2$ . Claramente, não há como aprender com os dados que não foram fornecidos (HUANG et al., 2021b; HUANG et al., 2021a), por esse motivo, foi introduzida uma transparência dentro da região da fase

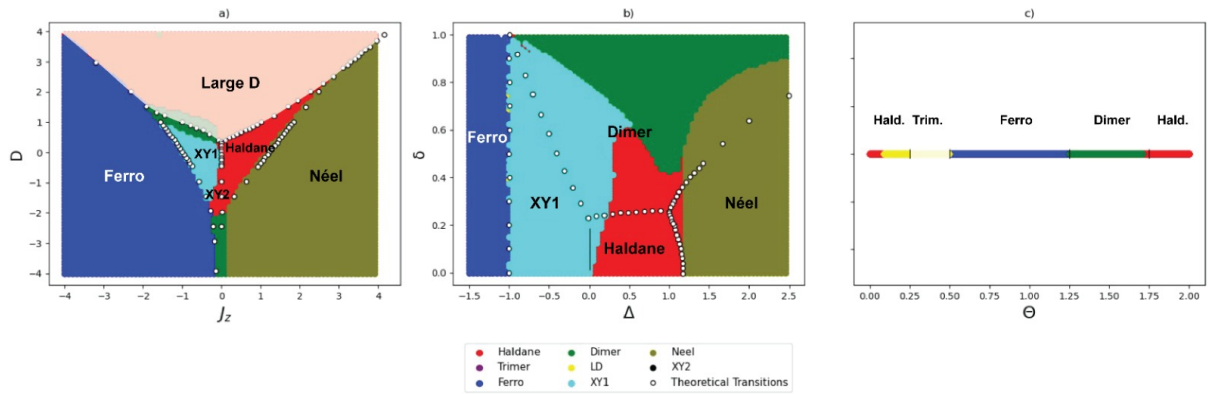


Figura 10 – A figura (a) mostra a previsão de fases do modelo  $\mathcal{H}_1$  usando o aprendizado dos modelos  $\mathcal{H}_2$  e  $\mathcal{H}_3$ . A figura (b) mostra a previsão de fases do modelo  $\mathcal{H}_2$  usando o aprendizado dos modelos  $\mathcal{H}_1$  e  $\mathcal{H}_3$ , e a figura (c) mostra a previsão de fases do modelo  $\mathcal{H}_3$  usando o aprendizado dos modelos  $\mathcal{H}_1$  e  $\mathcal{H}_2$ . Os pontos brancos nos painéis (a) e (b) e as linhas verticais no painel (c) são as fronteiras teóricas das fases fornecidas na literatura, para os modelos  $\mathcal{H}_1$  (CHEN; HIDA; SANCTUARY, 2003),  $\mathcal{H}_2$  (KITAZAWA; NOMURA; OKAMOTO, 1996) e  $\mathcal{H}_3$  (LÄUCHLI; SCHMID; TREBST, 2006), respectivamente. O nome em preto representa a fase real naquele local do modelo, e a cor representa a fase prevista. As fases com transparência (LD em (a) e trimer em (c)) são aquelas que não estavam presentes na base de treinamento do algoritmo.

LD para facilitar a distinção visual da fase atribuída pelo algoritmo, apesar de sua inerente impossibilidade de identificação correta. Além disso, quando previmos as fases de  $\mathcal{H}_3$ , fizemos o mesmo com a fase Trimer. Não contabilizamos essas fases ao calcular a precisão do algoritmo, pois possivelmente levariam a números enganosos.

Em nossa análise, transformamos nosso conjunto de dados seguindo um procedimento de renormalização chamado pré-processamento de sinal espacial, onde o conjunto de dados é escalonado para ter uma norma unitária (SERNEELS; NOLF; ESPEN, 2006). Como mencionado anteriormente, cada estado fundamental e sua respectiva fase rotulada são representados pelas correlações entre um par de spins,  $\langle S_1^k, S_i^k \rangle$ , o que fornece  $\{N/2 + 1\}$  características para cada variável  $k = \{x, y, z\}$ , e correlações globais  $\langle \prod_j S_j^k \rangle$ . Como observamos na Fig. (9), mesmo após a transformação do conjunto de dados, ainda há alguma sobreposição entre as fases. Por exemplo, na Fig.(9-f), é fácil notar a sobreposição da fase XY1 com as fases Haldane e Dimer, e a sobreposição entre as fases Néel e Dimer. A Fig. (10) mostra a previsão de fases do algoritmo  $k$ -NN considerando 12 sítios. Na Fig. (10-a), estimamos o diagrama de fases de  $\mathcal{H}_1$  com base nos dados de  $\mathcal{H}_2$  e  $\mathcal{H}_3$ , na Fig. (10-b), estimamos o diagrama de fases do modelo  $\mathcal{H}_2$  com base nos dados de  $\mathcal{H}_1$  e  $\mathcal{H}_3$ , e, de forma análoga, na Fig. (10-c), onde estimamos as fases de  $\mathcal{H}_3$  com base nos dados de  $\mathcal{H}_1$  e  $\mathcal{H}_2$ . A previsão do diagrama de fases de  $\mathcal{H}_1$ , apresentada na Fig. (10-a), foi incrivelmente bem-sucedida, com uma precisão de 96,77%. Todas as fases estão nas posições corretas com poucos erros nas transições de fases. Assim como em nosso exemplo ilustrativo bidimensional, a fase Haldane invade o espaço da fase XY1, e a confusão

entre as fases Néel e Dimer persiste. A região para a qual o algoritmo não tinha dados de treinamento (LD) foi quase inteiramente classificada como a fase Haldane. O motivo pode ser visto devido à proximidade das fases na Fig. (9-e). e (9-e). Quando consideramos  $\mathcal{H}_2$ , observamos que o algoritmo comete erros na separação das fases, o que fez sua precisão ser a mais baixa, cerca de 73,53%. Apesar disso, um aspecto precisa ser enfatizado. O conjunto de treinamento contém informações para  $\mathcal{H}_1$  e  $\mathcal{H}_3$ , que incluem todas as 8 fases distintas. No entanto, as fases Large-D, XY2 e Trimer não foram indicadas pelo algoritmo  $k$ -NN para todos os dados de  $\mathcal{H}_2$ , pois o algoritmo detecta todas as fases corretamente, cometendo erros apenas nas fronteiras. Finalmente, a previsão para o modelo  $\mathcal{H}_3$  é apresentada na Fig. (10-c), onde usamos os modelos  $\mathcal{H}_1$  e  $\mathcal{H}_2$  para aprendizado. Neste caso, ele classifica incorretamente a fase Haldane, misturando-a com a fase Large-D, que nem sequer está presente no modelo. No entanto, mesmo cometendo esse erro, o algoritmo alcança uma precisão de 88,27%. Apesar dos bons resultados, diferentes estratégias podem ser usadas para aumentar a precisão da previsão. Uma delas é adicionar novos modelos ao conjunto de dados do preditor, pois adicionar novas informações ao  $k$ -NN ajudaria a evitar previsões incorretas de fases. A segunda é encontrar uma transformação que possa separar as informações de fase no espaço de correlação. Por fim, diferentes algoritmos de aprendizado de máquina podem certamente ser implementados para aumentar a precisão.

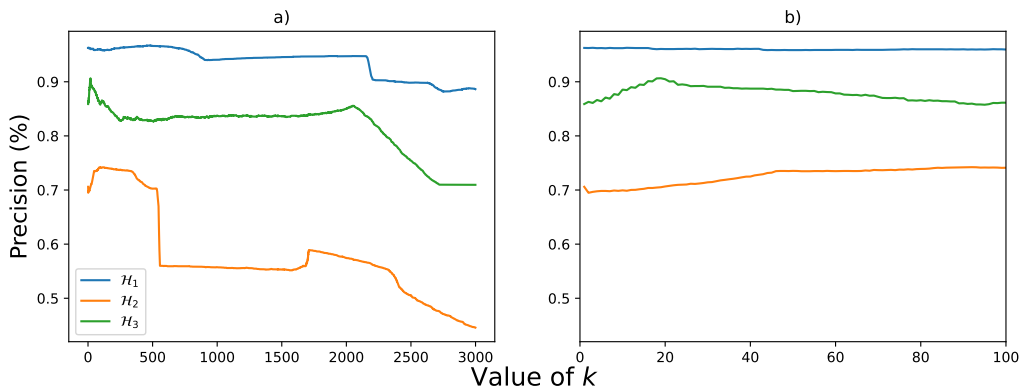


Figura 11 – Precisão obtida em função do valor de  $k$  ao classificar  $\mathcal{H}_1$  usando dados de  $\mathcal{H}_2$  e  $\mathcal{H}_3$  para treinamento, ao classificar  $\mathcal{H}_2$  usando dados de  $\mathcal{H}_1$  e  $\mathcal{H}_3$  para treinamento e ao classificar  $\mathcal{H}_3$  usando dados de  $\mathcal{H}_1$  e  $\mathcal{H}_2$  para treinamento. a) Variação da precisão para  $k$  até  $k = 3000$ , enquanto b) é um quadro na região entre  $0 < k < 100$ .

Outro componente levado em consideração em nossa análise é a influência do tamanho da cadeia na precisão do algoritmo. Os resultados obtidos podem ser observados na Tabela 2.

Como esperado, em geral, a precisão do algoritmo aumentou à medida que o número de spins aumentou, pois isso é uma representação melhor do sistema do ponto de vista físico. A única exceção a esse comportamento foi com o modelo XXZ para 12 sítios, no entanto, todos os erros do algoritmo estavam concentrados na região de transição de fases, o que naturalmente é uma tarefa mais complexa para ele.

	$\mathcal{H}_1$ (%)	$\mathcal{H}_2$ (%)	$\mathcal{H}_3$ (%)
4 Sítios	97,42	68,14	72,13
8 Sítios	97,77	69,14	79,44
12 Sítios	96,77	73,53	88,27

Tabela 2 – Precisão obtida pelo algoritmo  $k$ -NN na classificação das fases quânticas quando há variação no tamanho da cadeia de spins.

Além disso, apesar de desvendar transições de fases ser um tema de grande interesse na literatura, o método usado aqui não está voltado para revelar informações sobre esse aspecto. O mais próximo das transições que este trabalho aborda é a ideia de identificar seus limites, dada a alta precisão na classificação, como em  $\mathcal{H}_1$ . Por fim, outra discussão importante é a relação entre o valor  $k$  escolhido para o algoritmo (o único parâmetro especificado) e a precisão obtida. Esse problema é abordado na Fig. (11), que é uma representação gráfica da relação entre a precisão obtida pelo algoritmo e o número  $k$  de vizinhos usados. A partir disso, nossa análise mostra que, em nosso caso, o algoritmo funciona melhor para valores entre  $30 < k < 60$ , uma vez que, para valores muito pequenos, ele se torna mais suscetível a ruídos e, no caso de valores grandes de  $k$ , as fronteiras de fases distintas são incluídas e a classificação das fases se torna mais difícil. No quadro de detalhes da Fig. (11-b), podemos observar a região onde  $0 < k < 100$ , onde vemos que a precisão em cada modelo individual varia de acordo com o valor de  $k$ , de modo que a região em torno de  $k = 50$  foi escolhida porque é onde a soma da precisão ao prever as três combinações é a maior.

#### 5.1.4 Conclusão

Desenvolvemos um método para estudar as fases de sistemas magnéticos desconhecidos por meio de correlações de spins. Mostramos que algoritmos treinados com fragmentos de dados de um modelo são capazes de classificar fases do próprio modelo com precisão excelente, inclusive, precisando apenas de um número limitado de dados ( $\approx 2\%$ ) do diagrama de fases para realizar tais feitos. Concluimos também que os dados brutos de correlação carregam informações sobre as fases, que são independentes do modelo. Com as informações de correlação de spins, usamos um algoritmo  $k$ -NN para prever as fases de um modelo desconhecido com alta precisão. Apresentamos uma prova de conceito, mostrando que um algoritmo de aprendizado de máquina pode classificar fases desconhecidas de um Hamiltoniano por meio de fases conhecidas de outro Hamiltoniano, criando um classificador de fases quânticas independente do modelo. Nossa estratégia aborda o problema com ferramentas que são reproduzíveis e generalizáveis para novos Hamiltonianos e, embora não desenhe uma linha nítida para determinar os pontos de transição, fornece uma representação fiel do diagrama de fases quânticas para um modelo arbitrário. Uma possível vantagem desse método é que é possível treinar a partir de um conjunto de modelos numericamente menos exigentes, e a avaliação pode ser feita em modelos mais desafiadores. Ressaltamos que não foi feito uso explícito de parâmetros de ordem de fase, portanto, este

classificador independente do modelo abre a possibilidade de criar um classificador universal, à medida que mais e mais informações independentes do modelo são adicionadas ao banco de dados do classificador.

## 5.2 Classificando Fases Quânticas com Dados Escassos

Um dos métodos mais amplamente utilizados para investigar fases quânticas, especialmente em sistemas unidimensionais, é o *Grupo de Renormalização da Matriz de Densidade* (DMRG) (WHITE, 1992). Desde seu desenvolvimento, tornou-se o padrão-ouro para simular sistemas quânticos fortemente correlacionados com alta precisão (HALLBERG, 2006). No entanto, apesar de sua precisão, apresenta desafios computacionais significativos. O método requer uma quantidade substancial de memória e recursos de tempo, particularmente à medida que os tamanhos dos sistemas e o entrelaçamento dos estados quânticos aumentam (PAULETTI et al., 2024). Esse custo computacional limita a capacidade de explorar grandes espaços de parâmetros, especialmente em sistemas quânticos mais complexos ou de maior dimensionalidade. Consequentemente, identificar fases quânticas por meio de abordagens tradicionais de DMRG pode ser demorado.

Por outro lado, esforços recentes mostraram que ML oferece uma alternativa promissora para a classificação de fases (CARRASQUILLA; MELKO, 2017; DONG; POLLMANN; ZHANG, 2019; SHIINA et al., 2020; REM et al., 2019; BROECKER et al., 2017; CANABARRO et al., 2019; MAHLOW et al., 2023b), aprendendo padrões a partir de conjuntos de dados e generalizando por espaços de parâmetros sem a necessidade de cálculos exaustivos. A integração do DMRG com ML cria um pipeline eficiente, onde dados de alta qualidade das simulações podem ser usados para treinar modelos capazes de extrapolar fronteiras de fases além do que os métodos tradicionais podem facilmente alcançar. No entanto, um desafio nessa abordagem reside no fato de que gerar grandes quantidades de dados para treinar algoritmos de ML é computacionalmente caro, dado os significativos recursos exigidos pelas simulações. Essa limitação levanta preocupações sobre a viabilidade de usar o DMRG para gerar dados para algoritmos de ML. Nesse sentido, o foco principal deste trabalho é a exploração do potencial do ML para classificar fases quânticas com conjuntos de dados limitados, o que é particularmente relevante quando se utilizam métodos computacionalmente caros.

Aplicamos técnicas de ML a dois modelos canônicos de spin-1: a cadeia XXZ e a cadeia bilinear-biquadrática. Esses modelos exibem uma variedade de fases quânticas, incluindo as fases de Haldane, Néel, XY, Ferromagnética, dimerizada e trimerizada, que têm sido extensivamente estudadas (MALVEZZI et al., 2016). Ao aproveitar algoritmos de ML, visamos classificar essas fases de forma eficiente com poucos dados, usando resultados do DMRG como referência para validar nossa abordagem. Para mitigar ainda mais a questão de dados limitados, empregamos técnicas de geração de dados sintéticos utilizando *Redes Adversariais Generativas* (GANs) para *aumento de dados*. O conceito de *aumento de dados* (SHORTEN; KHOSHGOFTAAR, 2019) não é novo e tem sido amplamente utilizado em áreas como geração de imagens com Inteligência Artificial Generativa (GenAI), em modelos como o DALL-E da OpenAI (RAMESH et al., 2021), para melhorar conjuntos de dados e aprimorar o desempenho de modelos. No nosso caso, as GANs aplicam uma abordagem semelhante ao gerar dados adicionais de fases

quânticas, o que aumenta o conjunto de treinamento para os algoritmos de ML. Essa estratégia não apenas reduz o ônus computacional associado a simulações de DMRG em larga escala, mas também melhora a precisão dos modelos de classificação de fases quando enfrentam pequenos conjuntos de dados.

Para este estudo, utilizamos quatro algoritmos de AM: *k-Nearest Neighbors* ( $k$ -NN), Perceptron Multicamadas (MLP), Floresta Aleatória (RF) e *Gradient Boosting* (GB). Todos os modelos apresentaram bom desempenho na tarefa de classificar as fases com quantidades limitadas de dados, e os algoritmos  $k$ -NN e SVM mostraram melhoria consistente com o uso de aumento de dados.

A seção está estruturada da seguinte forma: a Seção 5.2.1 fornece uma visão geral dos modelos físicos e suas fases quânticas associadas. A Seção 5.2.2 detalha o processo de geração do conjunto de dados e sua estrutura. A Seção 5.2.3 apresenta os resultados do nosso estudo, destacando a precisão alcançada e o impacto dos dados sintéticos. Finalmente, a Seção 5.2.4 conclui com uma discussão sobre direções futuras na classificação de fases quânticas.

## 5.2.1 Modelos Físicos

Neste estudo, conduzimos uma análise computacional extensiva que produziu aproximadamente 500 pontos de dados para cada um dos dois modelos unidimensionais examinados: o modelo XXZ e o modelo bilinear-biquadrático. Nosso objetivo é utilizar um pequeno subconjunto desses pontos de dados para treinar um algoritmo de aprendizado de máquina, reservando a maioria para uma validação robusta dos resultados. Os estados fundamentais dos hamiltonianos foram calculados usando o método padrão de *Density Matrix Renormalization Group* (DMRG) de sistema infinito (WHITE, 1992). Essa técnica envolve a adição iterativa de dois sítios centrais a uma cadeia aberta, com o estado fundamental sendo recalculado e renormalizado após cada adição. Após inúmeras iterações, esses sítios centrais tornam-se parte de um sistema a granel, permitindo-nos extrair sua matriz densidade reduzida a partir do estado fundamental. É importante notar que tratamos as interações de spin entre esses sítios centrais de forma exata. Durante esse processo, os erros de truncamento variaram de  $10^{-10}$  a  $10^{-6}$ , com erros maiores observados perto de transições de fase de segunda ordem devido ao aumento das flutuações quânticas (MALVEZZI et al., 2016). Nossa discussão se concentra nas fases quânticas intrínsecas desses modelos e nas transições de fase que ocorrem entre elas.

### 5.2.1.1 Cadeia XXZ

Destacamos que esse modelo é um modelo unidimensional e não é o mesmo apresentado na equação 82. Para aprofundar na rica física dos sistemas quânticos unidimensionais, começamos examinando a cadeia de spin-1 XXZ, caracterizada pelo Hamiltoniano  $\mathcal{H}_1$ . Este modelo é fundamental para entender uma variedade de fenômenos quânticos, incluindo comportamento crítico em sistemas de baixa dimensionalidade e transições de fase. Essas transições, cada uma

pertencendo a diferentes classes de universalidade e associadas a distintas simetrias, fornecem uma plataforma rica tanto para simulações numéricas quanto para métodos analíticos. A fase de Haldane, em particular, caracterizada por sua ordem topológica não trivial e pela presença de estados de borda, tem sido objeto de extensa pesquisa. Estudos dessa fase contribuem para o campo mais amplo da matéria quântica topológica, oferecendo insights sobre emaranhamento quântico, parâmetros de ordem de corda e o papel das simetrias na proteção das fases topológicas.

O Hamiltoniano do modelo é definido como:

$$\mathcal{H}_1 = \sum_{i=1}^N \left( S_i^x S_{i+1}^x + S_i^y S_{i+1}^y + \Delta S_i^z S_{i+1}^z \right), \quad (85)$$

Aqui,  $S_i$  denota um operador de spin-1 no sítio  $i$  em uma rede unidimensional de  $N$  sítios, e o parâmetro  $\Delta$  representa a anisotropia nas interações de troca de spin. Dentro do intervalo  $\Delta \in [-1, 1.5]$ , o modelo exibe três fases distintas em seu diagrama de fases: a fase XY, a fase de Haldane e a fase de Néel (CHEN; HIDA; SANCTUARY, 2003). Quando  $\Delta < -1$ , o sistema está em uma fase ferromagnética. Em  $\Delta = -1$ , uma transição de fase de primeira ordem separa a fase ferromagnética da fase XY. À medida que  $\Delta$  aumenta, o sistema passa por uma transição de fase de ordem infinita (do tipo Kosterlitz-Thouless) em torno de  $\Delta \approx 0$ , marcando a transição da fase XY para a fase de Haldane. Esta fase de Haldane persiste até outro ponto crítico em  $\Delta \approx 1.185$ , onde ocorre uma transição de fase de segunda ordem para a fase de Néel, pertencente à classe de universalidade de Ising bidimensional.

### 5.2.1.2 Cadeia Bilinear Biquadrática

Em nossa segunda investigação, voltamos nossa atenção para a cadeia bilinear biquadrática de spin-1, descrita pelo Hamiltoniano  $\mathcal{H}_2$ :

$$\mathcal{H}_2 = \sum_{i=1}^N [\cos \theta (S_i \cdot S_{i+1}) + \sin \theta (S_i \cdot S_{i+1})^2], \quad (86)$$

Neste sistema,  $N$  representa o número total de sítios, e  $S_i$  denota o operador de spin-1 no sítio  $i$ . O parâmetro  $\theta \in [0, 2\pi)$  quantifica o acoplamento entre spins vizinhos mais próximos. No intervalo  $-0.25\pi < \theta < 0.25\pi$ , o sistema reside na fase de Haldane. Em  $\theta = 0.25\pi$ , ocorre uma transição do tipo Kosterlitz-Thouless (KT), separando a fase de Haldane de uma fase trimerizada sem gap. Uma transição de primeira ordem da fase trimerizada para a fase ferromagnética ocorre em  $\theta = 0.5\pi$ . Na região  $0.5\pi < \theta < 1.25\pi$ , o sistema está na fase ferromagnética. Outra transição de primeira ordem ocorre em  $\theta = 1.25\pi$ , marcando a fronteira entre a fase ferromagnética e a fase dimerizada com gap. Além disso, em  $\theta = 1.75\pi$ , ocorre uma transição de segunda ordem onde o sistema transita da fase dimerizada de volta para a fase de Haldane (MALVEZZI et al., 2016).

## 5.2.2 Estrutura dos Dados

Nesta seção, detalhamos o processo de aquisição de dados e a estrutura do conjunto de dados utilizado para o treinamento e teste dos modelos de aprendizado de máquina. Os dados foram gerados por meio de simulações numéricas focadas nos dois modelos de spin-1 mencionados.

Para o Hamiltoniano XXZ, conforme delineado na Eq. (85), o parâmetro de anisotropia  $\Delta$  foi variado dentro do intervalo especificado  $[-1, 1.5]$  em incrementos de 0.005, resultando em uma amostragem minuciosa do espaço de fases, culminando em um total de 500 pontos de dados. Para cada valor de  $\Delta$ , os observáveis físicos pertinentes foram determinados para os operadores de spin-1. Especificamente, foram avaliados os oito observáveis  $\langle \lambda_i \rangle$ , onde  $\lambda_i$  denota as oito matrizes de Gell-Mann (SCHERER; SCHINDLER, 2005). Esses observáveis, em conjunto com as fases associadas, formam o conjunto de características do nosso conjunto de dados aplicável a ambos os modelos. No que diz respeito à cadeia bilinear biquadrática, conforme descrito na Eq. (86), o ângulo  $\theta$  foi variado no intervalo  $[0, 2\pi)$  em incrementos de 0.004, resultando também em 500 pontos de dados. De forma análoga ao modelo XXZ, o Hamiltoniano foi calculado para cada valor de  $\theta$ , empregando o método DMRG para extrair as propriedades relevantes do estado fundamental. Mais uma vez, foram determinados os oito observáveis  $\langle \lambda_i \rangle$ .

Em nosso estudo, empregamos modelos de ML supervisionados. Essa abordagem implica que, durante o processo de treinamento, o modelo é fornecido com os observáveis coletados como entradas e as fases quânticas correspondentes como saídas, onde as fases são codificadas como rótulos inteiros. Para isso, foi necessário rotular o conjunto de dados descrito, concedendo ao modelo acesso à informação de fase nos dados de treinamento. As previsões do modelo são então comparadas com as fases conhecidas, previamente estabelecidas na literatura, cujas transições são discutidas em detalhes na Seção 5.2.1. Utilizamos uma CTGAN para gerar dados sintéticos, aumentando nosso conjunto de dados limitado. A implementação utilizou o pacote `ydata_synthetic.synthesizers.regular`<sup>1</sup> e a classe `RegularSynthesizer`. O procedimento experimental é descrito a seguir:

Começamos selecionando um subconjunto do conjunto de dados original para treinar a CTGAN. O número de linhas selecionadas para cada fase quântica por modelo foi variado sistematicamente de 1 a 30, e para cada quantidade específica de linhas, selecionamos aleatoriamente quais linhas incluir. Essa abordagem nos permitiu estudar como a variação no número inicial de pontos de dados usados para treinar a GAN impactou as previsões finais de aprendizado de máquina. A semente aleatória foi fixada para reprodutibilidade. O conjunto de dados foi então dividido em colunas numéricas e categóricas, com a coluna categórica especificada como `labels`, enquanto as colunas numéricas incluíam todas as outras características (observáveis e o conjunto de valores de  $\Delta$  e  $\theta$ , que são os parâmetros variados no Hamiltoniano).

<sup>1</sup> <https://github.com/ydataai/ydata-synthetic>

O modelo CTGAN foi configurado com hiperparâmetros específicos. O tamanho do lote foi definido como 500, determinando o número de amostras processadas antes de atualizar os parâmetros internos do modelo. O modelo foi treinado por 501 épocas, indicando o número de vezes que todo o conjunto de dados foi passado pelo modelo. Uma taxa de aprendizado de  $2 \times 10^{-4}$  foi escolhida para controlar o tamanho do passo em cada iteração ao minimizar a função de perda. O otimizador Adam foi empregado com  $\beta_1 = 0.5$  e  $\beta_2 = 0.9$ , que influenciaram os termos de momentum durante a otimização. A dimensionalidade do espaço latente (vetor de ruído) foi definida como 32, enquanto a representação interna tinha uma dimensionalidade de 128.

Após especificar os parâmetros da CTGAN, um objeto `RegularSynthesizer` foi instanciado, seguido pelo treinamento do modelo em um subconjunto selecionado do conjunto de dados por um número predeterminado de épocas. Subsequentemente, a CTGAN foi utilizada para gerar 1.000 pontos de dados sintéticos. Essas amostras sintéticas foram projetadas para replicar as características estatísticas do conjunto de dados original, fornecendo, assim, dados adicionais para aplicações subsequentes de aprendizado de máquina. Por fim, os dados gerados sinteticamente foram integrados ao conjunto de treinamento para potencialmente melhorar a eficácia do modelo na classificação de fases quânticas.

### 5.2.3 Resultados

Nesta pesquisa, avaliamos a eficácia de algoritmos de aprendizado de máquina supervisionado na classificação de fases quânticas com um conjunto de dados mínimo. O experimento preliminar, ilustrado nas Figuras 12 e 13, indica que uma alta precisão de classificação pode ser alcançada mesmo com dados limitados. Especificamente, executamos a classificação dos diagramas de fases para os modelos XXZ e Bilinear empregando os métodos SVM, Random Forest, Perceptron Multicamadas e *k-Nearest Neighbors*, conforme elucidado acima. Para o modelo XXZ, que abrange três fases distintas, os dados de treinamento compreenderam cinco amostras por fase, totalizando 15 pontos de dados. Para o modelo Bilinear, com quatro fases distintas, o treinamento foi realizado com 20 pontos de dados em condições idênticas (cinco pontos de dados por fase). Este experimento foi reiterado 100 vezes, com cada iteração envolvendo a seleção de uma amostra de dados aleatória, e uma votação por maioria foi utilizada para determinar a fase designada por cada modelo ao longo dessas iterações. Esse método foi imperativo devido à natureza estocástica da seleção dos dados de treinamento, facilitando uma avaliação mais autêntica do desempenho de cada modelo com dados mínimos.

Embora apenas uma pequena parte do conjunto de dados tenha sido utilizada—3% para o modelo XXZ e 4% para o modelo Bilinear—os algoritmos de aprendizado de máquina demonstraram alta precisão na classificação das fases quânticas. Especificamente, os modelos SVM, Random Forest, Perceptron Multicamadas e *k-Nearest Neighbors* alcançaram acurácias preditivas médias de aproximadamente 80.6%, 94.4%, 93.9% e 85.9% para o modelo XXZ, e

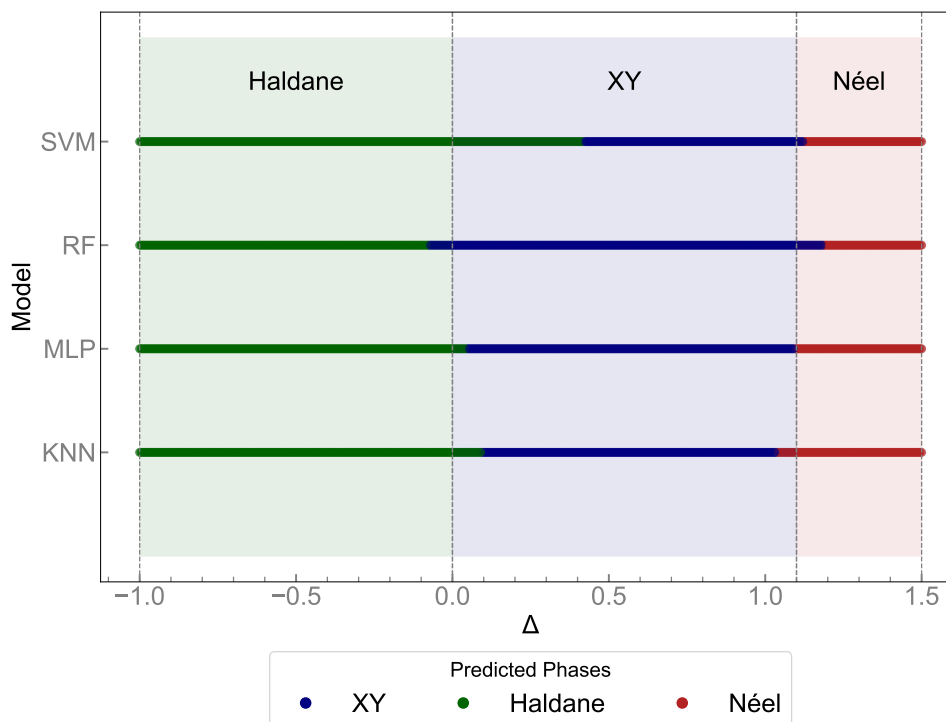


Figura 12 – Classificação do diagrama de fases XXZ usando modelos de aprendizado de máquina supervisionado com dados mínimos. Os resultados ilustram a precisão da previsão de fase alcançada pelos modelos SVM, Random Forest, Perceptron Multicamadas e *k-Nearest Neighbors*, utilizando apenas cinco pontos de dados por fase.

92.2%, 94.5%, 95.7% e 90.5% para o modelo Bilinear, respectivamente. Os resultados indicam claramente que, dado os dados limitados disponíveis, os modelos Perceptron Multicamadas e Random Forest consistentemente apresentaram a maior precisão para ambos os modelos XXZ e Bilinear.

Para determinar o conjunto de dados mínimo necessário para uma representação precisa dos diagramas de fases dos modelos, começamos treinando os modelos usando apenas um ponto de dado aleatório de cada fase. Em seguida, utilizamos esses pontos para classificar o restante do diagrama. Em iterações sucessivas, aumentamos o conjunto de dados de treinamento adicionando um ponto de dado adicional por fase para avaliar o efeito desse aumento incremental na precisão preditiva dos modelos. Para cada ponto, foram iteradas 100 vezes a seleção aleatória dos dados que iriam compor o conjunto de dados de treino, de forma a obter um panorama honesto da precisão média do algoritmo, independente dos dados iniciais.

A Figura 14 ilustra o desempenho dos modelos em termos de acurácia em função do número de amostras de treinamento para o modelo XXZ. Garantimos que pelo menos um ponto de dado de cada fase estivesse incluído em cada iteração, o que foi crucial, pois uma seleção aleatória de pontos de dados poderia, de outra forma, omitir fases inteiras do conjunto de treinamento. Tais omissões levariam à classificação incorreta de todos os pontos de dados dentro da fase negligenciada, comprometendo significativamente a precisão geral do modelo.

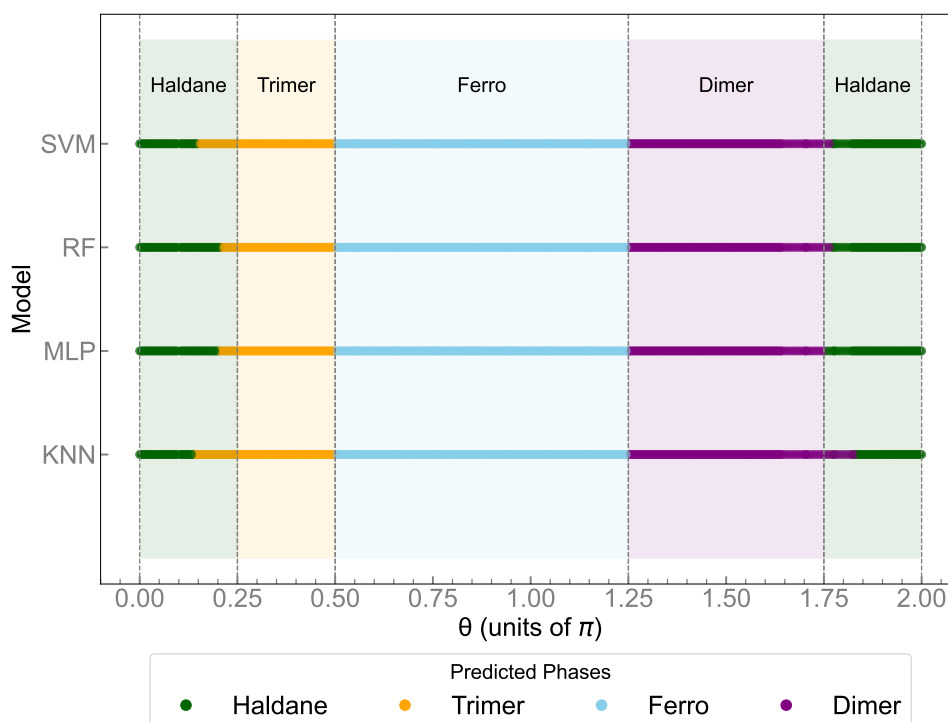


Figura 13 – Classificação do diagrama de fases Bilinear usando modelos de aprendizado de máquina supervisionado com dados mínimos. A figura mostra a precisão da previsão de fase para os modelos SVM, Random Forest, Perceptron Multicamadas e *k*-Nearest Neighbors, treinados com cinco pontos de dados por fase.

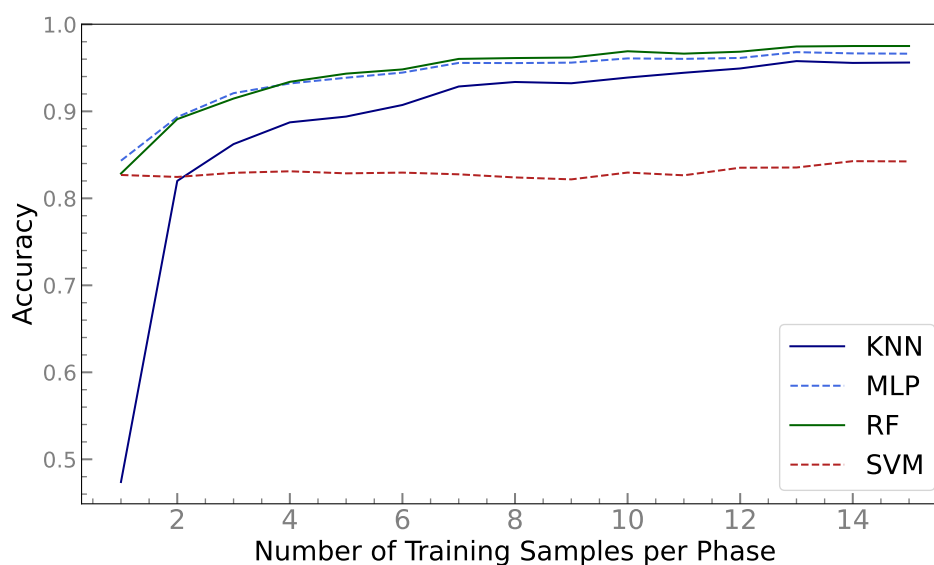


Figura 14 – Desempenho dos algoritmos SVM, Random Forest, Perceptron Multicamadas e *k*-Nearest Neighbors em termos de acurácia na classificação de fases do modelo XXZ, em função do número de amostras de treinamento por fase.

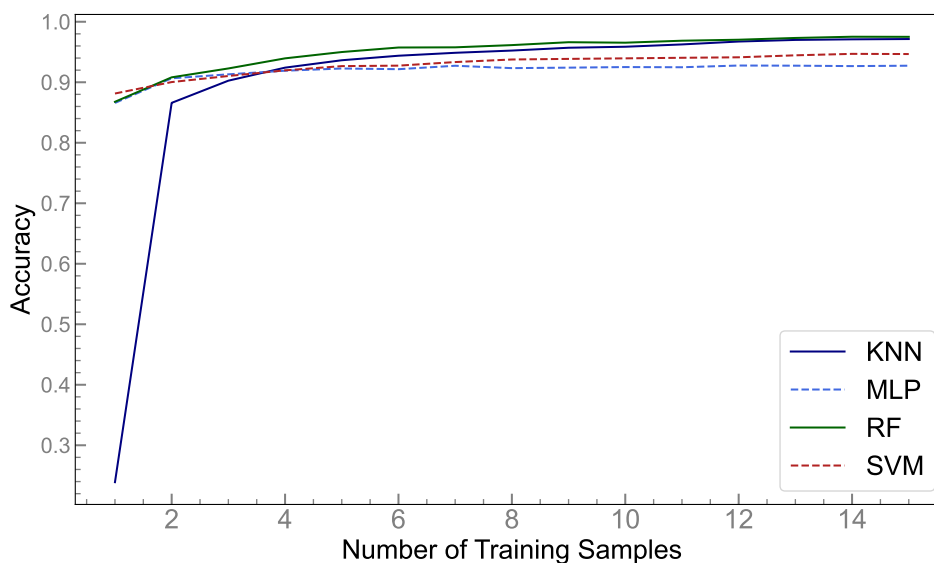


Figura 15 – Desempenho dos algoritmos SVM, Random Forest, Perceptron Multicamadas e *k*-Nearest Neighbors na classificação de fases do modelo Bilinear em função do número de amostras de treinamento por fase.

Como o modelo não pode prever com precisão a fase dos dados nos quais não foi treinado, essa abordagem foi essencial para manter a confiabilidade das previsões do modelo. Como pode-se observar, para o caso do XXZ, o algoritmo SVM manteve uma precisão abaixo dos demais, ficando durante todo o intervalo estudado com precisão próxima dos 80%. Os algoritmos com o melhor desempenho foram o XXZ e o RF, entretanto, para o caso estudado, quando tratamos de pouquíssimos dados o algoritmo KNN apresenta desempenho pobre, melhorando significativamente conforme a quantidade de dados disponíveis aumenta.

Em seguida, na figura 15, observamos que desta vez, o algoritmo com os piores resultados foram as redes neurais MLP, seguidos pelo SVM. Novamente os melhores desempenhos foram o KNN e o RF, sendo que novamente observamos que o KNN apresenta desempenho pobre quando alimentado com pouquíssimos dados. Apesar das observações feitas, destaca-se a capacidade de tais algoritmos de realizarem previsões altamente precisas com uma quantidade muito pequena de dados. Ressaltando a possibilidade de uso desses algoritmos em contextos com dados escassos.

Como pôde ser observado nas figuras 14 e 15, apesar de apresentar uma ótima precisão para uma quantidade pequena de dados, o algoritmo KNN acaba tendo uma precisão baixa nas circunstâncias em que se possui apenas 1 – 4 dados de cada fase (entre 3 e 16 dados totais). Por este motivo, o utilizamos como base para testar a hipótese de que o uso de GANs para a geração de dados sintéticos pode ser benéfica quando nos deparamos com *datasets* escassos.

Como pode ser observado na figura 16, observamos que com a adição dos dados sintéticos a precisão do KNN aumenta muito, principalmente no caso de pouquíssimos dados. A partir de 15 – 20 dados a precisão do modelo com e sem os dados começa a se equiparar.

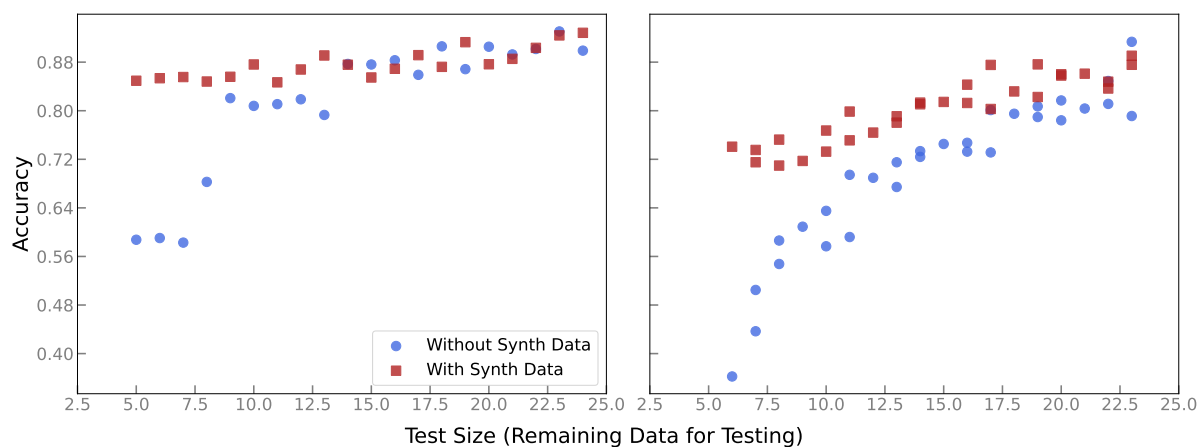


Figura 16 – Impacto da adição de dados sintéticos gerados por uma CTGAN no desempenho do algoritmo KNN para os modelos XXZ e Bilinear.

Desta forma, os resultados indicam que a adição de dados sintéticos ao dataset pode permitir que o algoritmo seja utilizado para realizar previsões mesmo em situações de extrema escassez de dados. Mesmo 1 dado por fase, aliado aos dados gerados sinteticamente pela CTGAN foram suficientes para que o KNN alcançasse cerca de 85% de precisão para o modelo XXZ e cerca de 75% para o Bilinear. Essas constatações contribuem no sentido de auxiliar com ferramentas para lidar com escassez de dados em física e mesmo para classificar fases de modelos calculados usando DMRG, que apresenta um alto custo computacional.

## 5.2.4 Conclusão

Este estudo explorou o uso de algoritmos de aprendizado de máquina supervisionado na classificação de fases quânticas, mesmo em cenários com conjuntos de dados limitados. Os resultados mostraram que os algoritmos Random Forest, Perceptron Multicamadas, SVM e *k-Nearest Neighbors* são capazes de classificar fases quânticas com alta precisão, especialmente em situações com dados moderadamente reduzidos. Observou-se que, enquanto algoritmos como Random Forest e Perceptron Multicamadas mantiveram uma precisão elevada independentemente da quantidade de dados, o desempenho do KNN era particularmente sensível ao número de amostras, apresentando dificuldades com pouquíssimos dados, mas melhorando significativamente à medida que o conjunto de treinamento aumentava.

A introdução de dados sintéticos gerados por GANs demonstrou ser uma solução eficiente para mitigar os desafios da escassez de dados. A adição desses dados foi especialmente benéfica para o KNN, permitindo alcançar níveis de precisão aceitáveis mesmo em condições de treinamento extremamente limitadas, como um único ponto por fase. Apesar disso, os ganhos trazidos pelos dados sintéticos se tornaram menos relevantes conforme o número de amostras reais aumentava, indicando que o uso de GANs é particularmente vantajoso em cenários de extrema limitação de dados.

Esses resultados abrem caminho para futuras investigações sobre o uso de aprendizado de máquina em classificações de fases quânticas. Trabalhos futuros podem explorar outros métodos de geração de dados sintéticos, como variações mais avançadas de GANs, para verificar se é possível melhorar ainda mais a qualidade dos dados gerados e o desempenho dos algoritmos. Além disso, o impacto de técnicas de pré-processamento, como aumento de dados ou seleção de características relevantes, merece uma análise mais detalhada, especialmente em problemas com alta dimensionalidade e conjuntos de dados reduzidos. Por fim, a aplicação dessas abordagens em modelos quânticos mais complexos e em experimentos reais é uma direção importante para validar e expandir os resultados obtidos, contribuindo para o avanço da interface entre aprendizado de máquina e física de sistemas quânticos.

## 5.3 Classificando Fases Quânticas com QML

Esta seção de aplicação é baseada no artigo:

- Santos, G., Mahlow, F., Prado, P., Pexe, G., Rattighieri, L., Fanchini, F. Quantum Phases Classification Using Quantum Machine Learning with SHAP-Driven Feature Selection. 2025. (**Submetido Quantum Machine Intelligence**).

Nesta seção, examinamos o modelo *Axial* (ou *anisotropic*) *Next-Nearest Neighbor Ising model* (ANNNI) (SELKE, 1988; CHAKRABARTI; DUTTA; SEN, 2008). Este modelo é particularmente significativo devido ao seu papel como um sistema simples que integra flutuações quânticas com interações competitivas de troca frustradas (CANABARRO et al., 2019). Propomos um método para classificar com precisão as fases quânticas dentro do modelo ANNNI utilizando SHAP com Máquinas de Vetores de Suporte Quânticas (QSVM) e Classificadores Quânticos Variacionais (VQC). Usando o algoritmo SHAP, identificamos as características mais influentes e, por meio de um pré-processamento de dados e mapeamento de características detalhado, preparamos nosso conjunto de dados para uma classificação eficaz. A aplicação de QSVM e VQC nos permite demonstrar o potencial desses modelos de QML (HAVLÍČEK et al., 2019b) para classificar as fases no modelo ANNNI de forma eficaz. Também estudamos o impacto do número de características na precisão de ambos os algoritmos. Essa abordagem não apenas destaca o potencial dos modelos de QML para analisar sistemas quânticos complexos, mas também examina o impacto da seleção de características na precisão desses modelos.

Esta seção está organizada da seguinte forma. A Seção 5.3.1 apresenta o modelo ANNNI, enfatizando sua importância e detalhando suas fases. A Seção 5.3.2 detalha a metodologia para seleção de características, pré-processamento de dados, mapeamento de características e tarefas de classificação usando os algoritmos de QML. A Seção 5.3.3 apresenta os resultados obtidos nas simulações computacionais, e a Seção 5.3.4 resume nossas descobertas, destacando possíveis pesquisas futuras e as implicações de nosso trabalho nas transições de fases quânticas.

### 5.3.1 O Modelo ANNNI

O modelo ANNNI de spin-1/2, foi utilizado neste estudo. As suas transições de fase são governadas pelo seguinte Hamiltoniano:

$$H = -J \sum_{j=1}^N (\sigma_j^z \sigma_{j+1}^z - \kappa \sigma_j^z \sigma_{j+2}^z + g \sigma_j^x). \quad (87)$$

Na expressão acima, as matrizes de Pauli  $\sigma_j^a$ , com  $a = x, z$ , atuam no grau de liberdade de spin-1/2 do sítio  $j$ -ésimo. A constante de acoplamento  $J > 0$ , que define a escala de energia (aqui usamos  $J = 1$ ), caracteriza a interação ferromagnética entre os vizinhos mais próximos.

Já  $\kappa$  é a constante de acoplamento adimensional para a interação entre os próximos-vizinhos mais próximos, e  $g$  representa o campo magnético transverso. O diagrama de fases do estado fundamental do modelo ANNNI inclui fases ferromagnética, antifase, paramagnética e *floating* (CANABARRO et al., 2019; YANG; LUO, 2023). Este diagrama resulta da frustração entre as interações ferromagnéticas de vizinhos mais próximos e as interações antiferromagnéticas de próximos-vizinhos mais próximos, combinadas com os efeitos desordenadores introduzidos pelo campo transverso.

A fase ferromagnética ocorre quando a interação entre vizinhos mais próximos predomina, resultando tipicamente em estados onde os spins se alinham em uma única direção ( $\uparrow\uparrow\uparrow\uparrow\uparrow\uparrow$ ). Quando a influência da interação entre próximos-vizinhos mais próximos prevalece, ela leva à antifase, caracterizada por uma periodicidade de quatro sítios ( $\uparrow\uparrow\downarrow\downarrow\uparrow\uparrow\downarrow\downarrow$ ). Na presença de um campo transverso forte, os spins tendem a se alinhar com o campo, resultando na fase paramagnética. A fase *floating* (RIEGER; UIMIN, 1996) é caracterizada por correlações de spin que decaem algebricamente, indicando uma ordem de longo alcance modulada que varia com o parâmetro de frustração  $\kappa$ , apesar de carecer de uma ordem estável de longo alcance.

Para gerar os dados de treinamento para nossos modelos de QML, calculamos as correlações entre pares do estado fundamental, o estado de menor energia do Hamiltoniano. Especificamente, considerando que estamos lidando com uma cadeia fechada (onde o primeiro e o último spins estão conectados, formando um loop), computamos os valores esperados dos observáveis ( $\langle\sigma_i^x\sigma_j^x\rangle$ ,  $\langle\sigma_i^y\sigma_j^y\rangle$ ,  $\langle\sigma_i^z\sigma_j^z\rangle$ ) para  $i = 1$  e  $j = [2, N/2+1]$ , onde  $N$  é o número de spins. Note que calcular apenas até  $j = N/2 + 1$  para  $i = 1$  evita redundância, pois os dados obtidos seriam duplicados. Assim, para sistemas com 8 e 12 sítios, obtemos 12 e 18 características, respectivamente. Incrementando os valores de  $g$  e  $\kappa$  em passos de 0.01 no intervalo de  $[0, 1]$ , calculamos 10.000 pontos de dados. Esses pontos de dados são subsequentemente analisados utilizando técnicas de aprendizado de máquina.

Dado que modelos de aprendizado supervisionado são empregados para a classificação de dados no âmbito do modelo ANNNI, um exame detalhado das transições de fase é essencial para nossa análise. Essas transições delimitam claramente as diferentes fases que o sistema atravessa sob diferentes regimes de parâmetros. Nesse sentido, uma compreensão abrangente dessas fronteiras é importante para a identificação e categorização precisas das fases, um ingrediente essencial para o treinamento eficaz dos modelos. No caso limite onde  $\kappa = 0$ , o modelo reduz-se ao bem conhecido modelo de Ising com campo transverso, que permite soluções analíticas exatas (SACHDEV, 2011). No ponto crítico  $g = 1$ , o sistema passa por uma transição de fase de segunda ordem, demarcando a mudança da fase ferromagnética ( $g < 1$ ) para a fase paramagnética ( $g > 1$ ). Uma expressão analítica aproximada para o valor crítico de  $g$  dentro do intervalo de parâmetros  $0 \leq \kappa \leq 1/2$  é fornecida por (SUZUKI; INOUE;

CHAKRABARTI, 2012):

$$g_I(\kappa) \approx \frac{1 - \kappa}{\kappa} \left( 1 - \sqrt{\frac{1 - 3\kappa + 4\kappa^2}{1 - \kappa}} \right) \quad (88)$$

Para  $\kappa > 1/2$ , o modelo passa por uma transição de Berezinsky-Kosterlitz-Thouless da fase *floating* para a fase paramagnética conforme  $g$  aumenta. O valor crítico aproximado de  $g$  para essa transição é dado por (BECCARIA; CAMPOSTRINI; FEO, 2007):

$$g_{BKT}(\kappa) \approx 1.05\sqrt{(\kappa - 0.5)(\kappa - 0.1)} \quad (89)$$

Como discutido, essas transições são fundamentais para nossa abordagem de aprendizado supervisionado, pois delimitam as fronteiras entre as diferentes fases do modelo. A classificação precisa dos dados por nossos algoritmos de QML depende de uma compreensão precisa dessas fronteiras de fase.

### 5.3.2 Métodos

Nesta seção, apresentamos a metodologia utilizada neste estudo. Começamos com a seleção de características utilizando valores SHAP para identificar as características mais relevantes, seguida por etapas de pré-processamento de dados, incluindo divisão em conjuntos de treino e teste e escalonamento de características. Em seguida, descrevemos o uso de mapas de características quânticos, como o *ZFeatureMap* (QUANTUM, 2021), para transformar dados clássicos em estados quânticos. Por fim, apresentamos o Classificador Quântico de Vetor de Suporte (QSVC) e o Classificador Quântico Variacional (VQC) para prever o diagrama de fases do modelo ANNNI.

#### 5.3.2.1 Seleção de Características com SHAP

Nesta seção, utilizamos o algoritmo SHAP para identificar as características mais relevantes no processo de aprendizado. Essa identificação é baseada em um modelo de aprendizado de máquina, neste caso o SVC, que avalia a importância de cada característica para as previsões. Essas informações são visualizadas em um gráfico resumo SHAP (ver Seção de Resultados), que destaca a contribuição de cada característica para as previsões e auxilia nas estratégias de seleção e otimização de características. Focamos nas características com as maiores pontuações de importância, visando melhorar a eficiência dos algoritmos de QML reduzindo o conjunto de características para as mais relevantes.

### 5.3.3 Resultados

Nesta seção, discutimos inicialmente a seleção das características mais importantes utilizando o algoritmo SHAP. Em seguida, apresentamos os resultados de classificação para

os modelos Quantum Support Vector Classification (QSVC) e Quantum Variational Classifier (VQC). Por fim, analisamos a precisão de ambos os modelos em função do número de características selecionadas pelo SHAP, demonstrando a eficácia dessa abordagem na melhoria do desempenho de classificadores quânticos.

### 5.3.3.1 Seleção de Características

Em nossa investigação, um passo crítico envolve a identificação e seleção de características (ou observáveis) que impactam significativamente o desempenho dos modelos estudados. Para isso, utilizamos o SHAP para definir a importância de diversos observáveis em nosso conjunto de dados. Como mencionado anteriormente, os dados utilizados para treinar nossos modelos de QML foram baseados em correlações entre pares de spins na estrutura. Especificamente, calculamos os valores esperados dos observáveis ( $\langle \sigma_i^x \sigma_j^x \rangle$ ,  $\langle \sigma_i^y \sigma_j^y \rangle$ ,  $\langle \sigma_i^z \sigma_j^z \rangle$ ) para  $j > i$  e  $i = [1, N - 1]$ , onde  $N$  representa o número de spins. Foi entre essas características que aplicamos o SHAP para identificar as mais relevantes, ou seja, aquelas que influenciam significativamente o desempenho do modelo.

A Figura 17 mostra o ranking de importância das características para a classificação das fases do modelo ANNNI utilizando o SVC em um sistema de 8 sítios. Este é conhecido como *SHAP summary plot*, onde características com valores SHAP próximos de zero indicam baixo impacto na saída do modelo. Como podemos observar, as três características mais importantes identificadas foram  $\langle \sigma_1^z, \sigma_4^z \rangle$ ,  $\langle \sigma_1^z, \sigma_3^z \rangle$  e  $\langle \sigma_1^z, \sigma_2^z \rangle$ , indicando que a projeção do spin ao longo do eixo  $z$  ( $\sigma^z$ ) desempenha um papel mais significativo na classificação de fases em comparação com outras projeções. Fisicamente, isso demonstra o papel relevante das correlações de spin longitudinais na determinação da estrutura de fases do modelo ANNNI. A forte influência dos pares  $\sigma^z$  envolvendo vizinhos mais próximos e próximos-vizinhos destaca a importância dessas interações específicas na indução de transições de fase, particularmente em relação à ordenação dos spins ao longo do eixo  $z$ .

A Figura 18 apresenta os resultados para um modelo com 12 sítios. As características mais importantes identificadas foram  $\langle \sigma_1^z, \sigma_6^z \rangle$ ,  $\langle \sigma_1^z, \sigma_7^z \rangle$  e  $\langle \sigma_1^y, \sigma_2^y \rangle$ . Notavelmente, com exceção da terceira característica, todas as outras envolvem correlações  $\sigma^z$ . Esse padrão sugere uma dominância consistente da projeção  $\sigma^z$  na determinação do comportamento de fase. A proeminência contínua das interações  $\sigma^z$  em diferentes tamanhos de sistema reforça seu papel crítico na definição das transições de fase no modelo ANNNI. Embora as contribuições das projeções  $\sigma^x$  e  $\sigma^y$  sejam menos pronunciadas, elas não são desprezíveis, com a projeção  $\sigma^y$  mostrando alguma relevância em um dos casos.

É importante enfatizar que, ao aplicar o SHAP com um SVC clássico, alinhamos o processo de seleção de características com o equivalente clássico do QSVM, potencialmente oferecendo ao QSVM uma vantagem em termos de relevância das características. Essa compatibilidade pode resultar em características mais significativas para o QSVM. Por outro lado, o

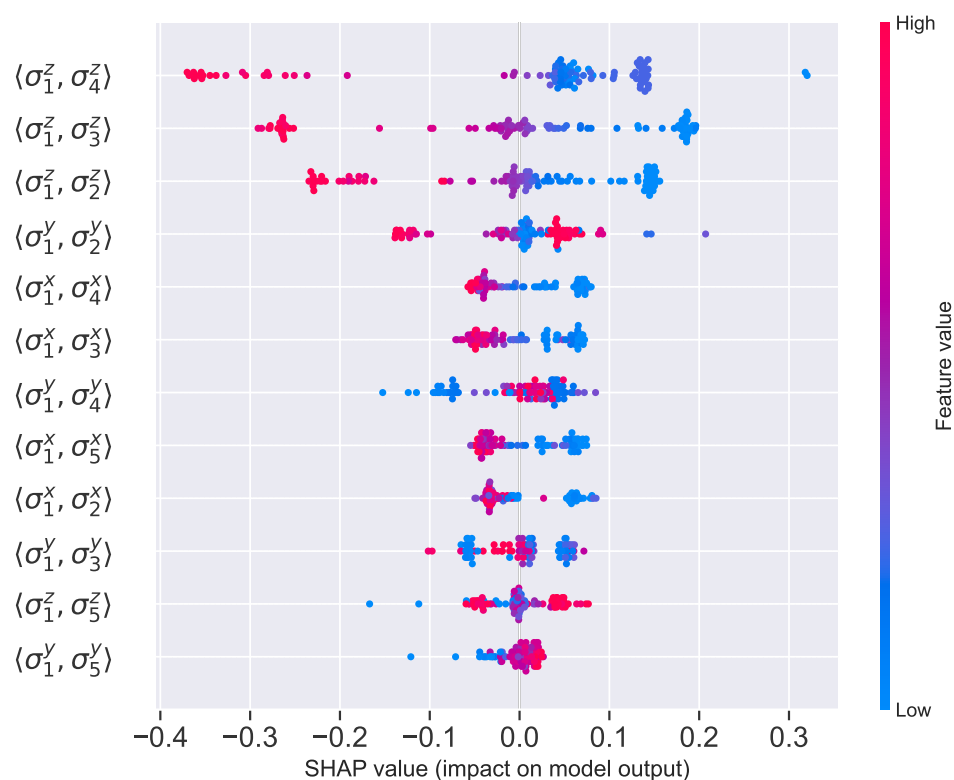


Figura 17 – Ranking de importância de características pelo SHAP para um modelo ANNNI de 8 sítios utilizando o Support Vector Classifier. O gráfico mostra a importância relativa das diferentes características na classificação das fases do modelo.

VQC carece de um equivalente clássico direto, tornando mais difícil determinar com precisão a importância das características dentro de seu framework. Essa discrepância aponta para a necessidade de técnicas de IA explicável específicas para modelos quânticos como o VQC. Com os principais observáveis agora identificados por meio do processo de seleção de características, podemos prosseguir para a próxima fase de nosso estudo: realizar tarefas de classificação utilizando essas características selecionadas.

### 5.3.3.2 Resultados de Classificação

Em nossa abordagem, implementamos o QSVM utilizando o kernel quântico baseado em fidelidade com o *ZFeatureMap* para codificação de dados. O *ZFeatureMap* foi configurado com três repetições, permitindo uma transformação eficaz de dados clássicos em estados quânticos. Para construir o kernel quântico, utilizamos o *FidelityStatevectorKernel* (JAVADI-ABHARI et al., 2024), que mede a similaridade entre esses estados quânticos. De forma similar, utilizamos o VQC com o *ZFeatureMap* para codificação de dados e o ansatz *EfficientSU2* (KANDALA et al., 2017), utilizando cinco repetições de entrelaçamento linear. Os parâmetros do circuito foram otimizados com o otimizador SPSA (SPALL, 1992) em 100 iterações, com valores esperados estimados a partir de 600 amostragens durante o treinamento. Ambos os modelos seguiram a mesma metodologia de treinamento, garantindo uma comparação justa, e foram avaliados

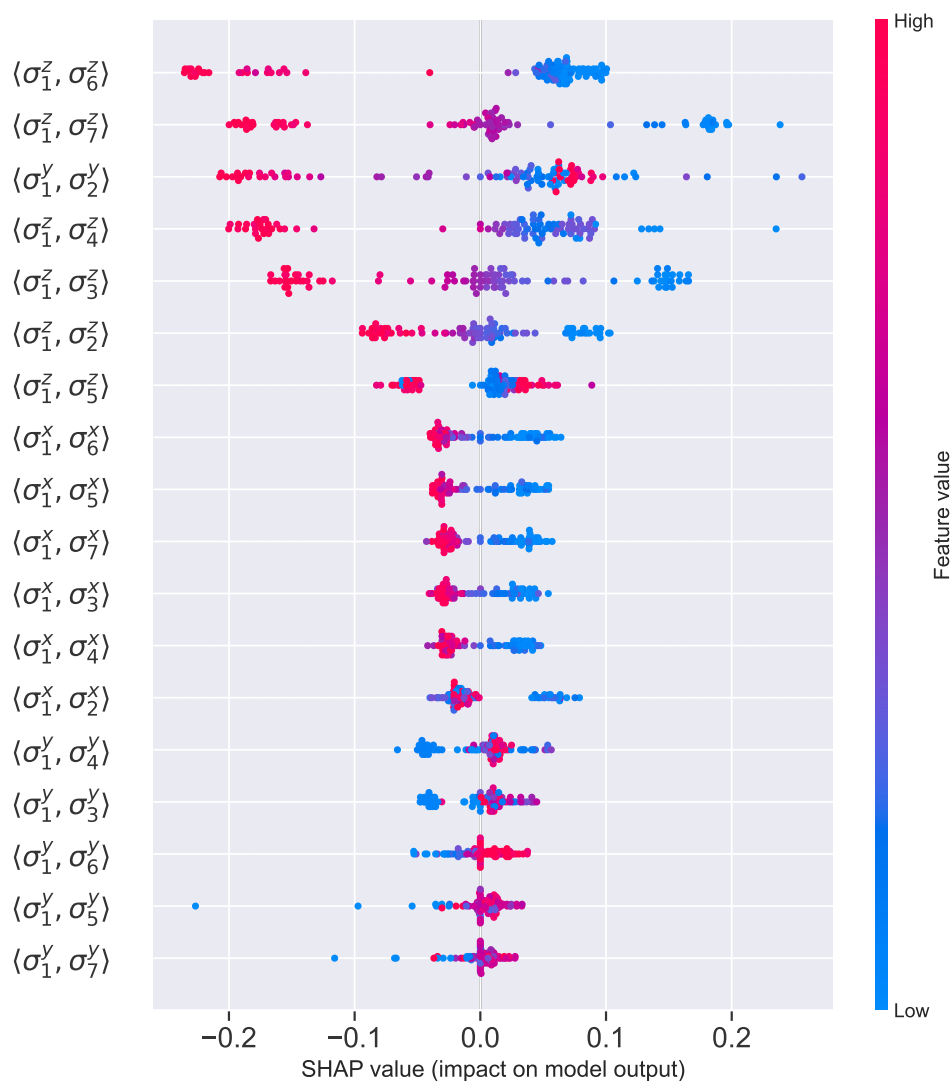


Figura 18 – Ranking de importância de características pelo SHAP para um modelo ANNNI de 12 sítios. O gráfico revela as características mais significativas para a classificação de fases neste sistema maior.

utilizando a mesma métrica de precisão.

Os resultados apresentados na Tabela 3 demonstram que ambos os algoritmos QSVM e VQC foram eficazes na classificação das fases quânticas, com a precisão em todos os casos superando 94%. Notavelmente, o QSVM superou o VQC nos casos estudados, particularmente no sistema de 8 sítios, onde alcançou uma precisão impressionante de 98,46

O desempenho superior do algoritmo QSVM pode ser atribuído à sua capacidade de explorar de forma mais eficaz o conjunto limitado de características disponíveis, especialmente em sistemas menores. A Figura 19 fornece uma representação visual da classificação de fases utilizando o QSVM e o VQC para sistemas de 8 e 12 sítios com 5 características. Como observado, os modelos classificaram as fases com precisão; no entanto, encontraram dificuldades em identificar as fronteiras de fase, com o QSVM de 8 sítios fornecendo a aproximação mais próxima das transições de fase observadas no modelo quântico subjacente. Isso sugere que o

Tamanho do Sistema	Algoritmo	Precisão
8 sítios	VQC	94,81%
12 sítios	VQC	96,49%
8 sítios	QSVM	98,46%
12 sítios	QSVM	97,73%

Tabela 3 – Desempenho na classificação dos algoritmos QSVM e VQC em sistemas de 8 e 12 sítios, utilizando as 5 características mais impactantes determinadas pelo algoritmo SHAP.

QSVM é mais adequado para capturar as características críticas necessárias para a identificação de fases nesses cenários específicos. O modelo VQC, apesar de seu desempenho geral competente, exibiu algumas classificações incorretas, notavelmente ao longo da linha  $g = 0$  para  $k > 0.5$ , onde classificou erroneamente a fase como ferromagnética em vez de antiferromagnética. Essa classificação incorreta foi consistente tanto no sistema de 8 sítios quanto no de 12 sítios. Tal limitação na detecção de fronteiras de fase contribuiu para a leve redução de precisão observada nos resultados do VQC.

A Figura 20 explora a relação entre o número de características selecionadas e a precisão de ambos os algoritmos QSVM e VQC. Os resultados indicam que o QSVM geralmente supera o VQC na maioria dos conjuntos de características. Para ambos os algoritmos, o desempenho ideal foi alcançado ao utilizar 5 características. Além desse intervalo, a precisão de ambos os modelos tende a diminuir, destacando a importância da seleção cuidadosa de características em tarefas de aprendizado de máquina quântico. Esses resultados enfatizam a eficácia do uso do SHAP para a seleção de características no aprendizado de máquina quântico. O SHAP foi instrumental na identificação das características mais relevantes para a classificação de fases, permitindo alta precisão mesmo com um conjunto reduzido de características.

No contexto da era NISQ (Noisy Intermediate-Scale Quantum), onde computadores quânticos são limitados por um número reduzido de qubits, a capacidade de alcançar alta precisão com um conjunto mínimo de características é particularmente valiosa. Os resultados demonstram que a classificação eficaz de fases quânticas pode ser alcançada sem a necessidade de um grande número de características, tornando o QSVM e o VQC opções viáveis para tarefas de classificação de fases nos ambientes de computação quântica atuais e futuros.

### 5.3.4 Conclusão

Este estudo demonstra que a seleção de características por meio do SHAP facilita a obtenção de alta precisão na classificação de fases quânticas utilizando algoritmos de QML, mesmo com um conjunto reduzido de características. A eficácia do SHAP em identificar as características mais relevantes permitiu que os algoritmos QSVM e VQC alcançassem uma classificação precisa das fases quânticas. Essa capacidade é particularmente valiosa na era NISQ,

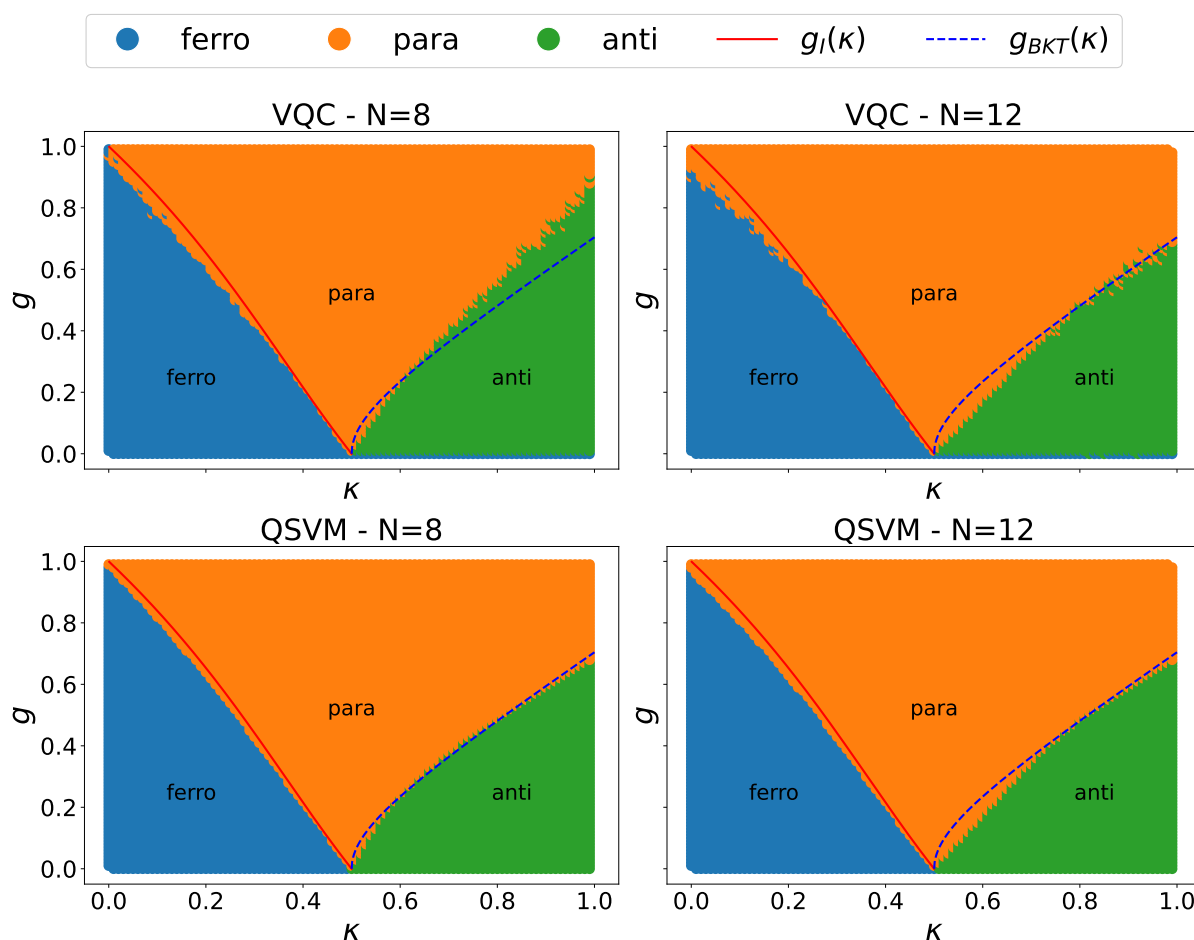


Figura 19 – Diagrama de fases do modelo ANNNI, delimitado por meio da classificação utilizando o VQC (Topo) e o QSVM (Base). Apresentado com cadeias de 8 sítios (esquerda) e 12 sítios (direita). Os marcadores denotam as previsões de fase de acordo com o modelo, enquanto as linhas representam as transições teóricas.

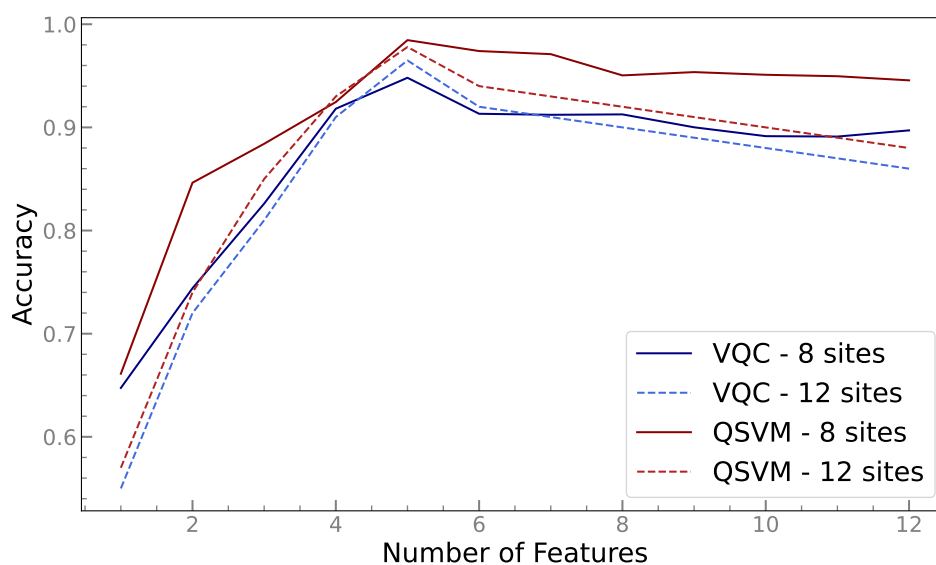


Figura 20 – Precisão em função do número de características para QSVM e VQC.

caracterizada pelas restrições impostas pela disponibilidade limitada de qubits nos recursos de computação quântica.

A capacidade de manter uma elevada precisão de classificação com um conjunto reduzido de características destaca a eficácia potencial dos algoritmos QSVM e VQC como ferramentas poderosas para a classificação de fases quânticas em ambientes com recursos limitados. Essa metodologia não apenas aumenta a eficiência computacional dos processos quânticos, mas também facilita o avanço dos algoritmos de QML para sistemas quânticos mais complexos.

Pesquisas futuras devem continuar a aprimorar as técnicas de seleção de características e explorar suas amplas aplicações dentro do aprendizado de máquina quântico, garantindo que esses algoritmos sejam utilizados de forma eficaz no domínio em constante progresso da computação quântica.

## 6 Estimando a Sincronização Quântica

Esta seção de aplicação é baseada no artigo:

- Mahlow, F., Çakmak, B., Karpát, G., Yalçinkaya, İ., Fanchini, F. F. (2024). Predicting the onset of quantum synchronization using machine learning. *Physical Review A*, 109(5), 052411. (**Editors' Suggestion**)

Sincronização é um fenômeno difundido que ocorre em muitos sistemas diferentes. Desde sistemas naturais, como as batidas do coração ou o piscar de vaga-lumes, até sistemas sociais, como o comportamento de uma multidão. Sistemas físicos podem exibir comportamento síncrono de duas formas, a saber, forçada e espontânea. Quando dois ou mais sistemas são forçados a oscilar em uníssono por uma influência externa, como a regulação da frequência cardíaca por um marca-passo externo através de pulsos elétricos, isso é conhecido como sincronização forçada. Por outro lado, a sincronização espontânea ocorre quando dois ou mais sistemas se sincronizam naturalmente na ausência de qualquer agente externo. Sincronização em sistemas clássicos tem sido estudada em uma variedade de contextos nas últimas décadas, com muitos resultados interessantes (PIKOVSKY; ROSENBLUM; KURTHS, 2001; ARENAS et al., 2008; OSIPOV; KURTHS; ZHOU, 2007). Consequentemente, o estudo desse fenômeno universal foi estendido para o domínio quântico.

No entanto, é importante enfatizar que sincronização é um termo de escopo amplo que pode, de fato, ser interpretado de várias maneiras (MARI et al., 2013; LI et al., 2017; GALVE; GIORGI; ZAMBRINI, 2017; ROULET; BRUDER, 2018; GIORGI; CABOT; ZAMBRINI, 2019). A sincronização forçada devido a um estímulo externo, também conhecida como “entrainment”, tem sido estudada em osciladores dissipativos impulsionados (ZHIROV; SHEPELYANSKY, 2008; ZHIROV; SHEPELYANSKY, 2009), osciladores de van der Pol (LEE; SADEGHPOUR, 2013; WALTER; NUNNENKAMP; BRUDER, 2014; SONAR et al., 2018) e modelos do tipo spin-bóson (GOYCHUK et al., 2006). Por outro lado, a sincronização espontânea induzida pelo ambiente também foi amplamente explorada, por exemplo, em osciladores harmônicos (GIORGI et al., 2012; MANZANO et al., 2013; MANZANO; GALVE; ZAMBRINI, 2013; BENEDETTI et al., 2016), osciladores de van der Pol (LEE; CHAN; WANG, 2014; WALTER; NUNNENKAMP; BRUDER, 2015), redes optomecânicas (HEINRICH et al., 2011; LUDWIG; MARQUARDT, 2013), íons frios em microarmadilhas (HUSH et al., 2015), redes atômicas (CABOT et al., 2019), spins acoplados a reservatórios comuns (ORTH et al., 2010; GIORGI et al., 2013; BELLOMO et al., 2017) e locais (GIORGI; GALVE; ZAMBRINI, 2016), sistemas atômicos de poucos corpos com dissipação coletiva (KARPAT; YALÇINKAYA; ÇAKMAK, 2020), modelos de colisão (KARPAT; YALÇINKAYA; ÇAKMAK, 2019; KARPAT et al., 2021; LI; LI; JIN, 2023b) e, mais recentemente, em muitos outros (SCHMOLKE; LUTZ, 2022; XIAO et al., 2023; KRITHIKA

et al., 2022; ES'HAQI-SANI et al., 2020; QIAO et al., 2020; LI; LI; JIN, 2023a; LI, 2022; WÄCHTLER; PLATERO, 2023; CATTANEO et al., 2021; CABOT; GIORGI; ZAMBRINI, 2021; BUČA; BOOKER; JAKSCH, 2022; IMPENS; GUÉRY-ODELIN, 2023; ŠTĚRBA; NOVOTNÝ; JEX, 2023; KALITA; CHAKRABORTY; SARMA, 2021).

Nos últimos anos, técnicas de aprendizagem de máquina começaram a ser empregadas em sistemas físicos, onde o comportamento síncrono entre seus constituintes emerge sob condições adequadas. Quase todos esses estudos focaram em redes complexas compostas por osciladores clássicos, com o objetivo de prever seu comportamento de sincronização (ITABASHI; TRAN; HASEGAWA, 2021; BASSI et al., 2022; FAN et al., 2022; GUTH; SAPSIS, 2019; CHOWDHURY et al., 2021; WENG et al., 2023; ZHANG et al., 2022; THIEM et al., 2020; BICCARI; ZUAZUA, 2020; FAN et al., 2021). Por outro lado, também foi demonstrado que um protocolo de aprendizagem de máquina quântica com feedback clássico pode melhorar a sincronização entre dois sistemas de dois níveis acoplados em duas cavidades (CÁRDENAS-LÓPEZ et al., 2019). Por fim, foi mostrado que uma rede neural artificial pode ser usada para inferir as propriedades de dissipação do ambiente através de um observável de sonda em um cenário de sistema aberto onde a sincronização se manifesta, e a emergência de sincronização melhora o desempenho de tarefas de classificação e regressão neste contexto (ESTARELLAS et al., 2019).

Neste capítulo, consideramos o fenômeno de sincronização induzido pelo ambiente, que emerge espontaneamente entre os valores esperados dos observáveis de spin de um par de qubits durante a evolução temporal do sistema aberto. Apresentamos uma abordagem baseada em aprendizagem de máquina que emprega um algoritmo de aprendizado supervisionado para identificar o comportamento de sincronização do par de qubits. Em particular, utilizamos o algoritmo de *k-Nearest Neighbors* para identificar se o par de qubits tende à sincronização, antissincronização, sincronização com atraso temporal ou à completa ausência desses fenômenos. Para quantificar o grau de sincronização entre os valores esperados dos observáveis locais dos qubits, calculamos o coeficiente de correlação de Pearson, que quantifica a associação entre duas variáveis medidas no mesmo intervalo. Enquanto estabelecemos o valor final do coeficiente de Pearson após uma longa evolução temporal como o alvo do nosso algoritmo, utilizamos os valores esperados dos observáveis dos qubits nos primeiros instantes da dinâmica como entrada para o algoritmo. Demonstramos a eficácia de nossa abordagem treinando e testando-a com dados quânticos calculados obtidos de três modelos distintos de sistemas abertos, e mostramos que ela é capaz de prever com precisão o comportamento de sincronização do par de qubits, mesmo na presença de erros aleatórios de medição. Estabelecemos uma relação de compromisso entre a porcentagem de erro na medição dos valores esperados e o número de medições necessárias para manter alta precisão.

Este capítulo está organizado da seguinte forma. Na Seção 6.1, apresentamos os modelos de sistemas abertos que consideramos. Na Seção 6.2, descrevemos nossa metodologia

para obter as bases de dados usadas para treinar e testar o algoritmo. Na Seção 6.3, apresentamos os principais resultados, demonstrando a eficácia de nossa abordagem na previsão do comportamento dos sistemas quânticos simulados, incluindo a adição de erros aleatórios nas medidas. Este capítulo é uma tradução do nosso segundo artigo (MAHLOW et al., 2023c), que foi publicado no *journal Physical Review A*, da *American Physical Society*, sendo escolhido como *editors' suggestion*.

## 6.1 Modelos Físicos

Nesta seção, apresentaremos três modelos físicos que exibem comportamentos sincronizados distintos (KARPAT; YALÇINKAYA; ÇAKMAK, 2019; KARPAT; YALÇINKAYA; ÇAKMAK, 2020; KARPAT et al., 2021). Nosso principal objetivo é estudar as propriedades de sincronização desses modelos sob a perspectiva de aprendizado de máquina.

### 6.1.1 Modelos de Colisão

Dois dos modelos que consideramos em nossa análise são baseados em modelos de colisão (CICCARELLO, 2017; CICCARELLO et al., 2022), que são ferramentas muito versáteis e eficazes na modelagem de sistemas quânticos abertos. No contexto dos modelos de colisão, cada interação, seja entre os subsistemas ou entre o sistema e o ambiente, é descrita como um acoplamento unitário curto entre as partes envolvidas, e o estado do sistema em questão é rastreado por meio desses acoplamentos sucessivos. Antes de detalharmos as propriedades distintas dos dois modelos de colisão que consideraremos, gostaríamos de delinear as partes comuns. A partir deste ponto, definiremos  $\hbar = 1$  em ambos os modelos, e temos dois qubits do sistema rotulados como  $s_1$  e  $s_2$  com seus hamiltonianos livres

$$H_{s_1} = -\frac{\omega_1}{2}\sigma_{s_1}^z, \quad H_{s_2} = -\frac{\omega_2}{2}\sigma_{s_2}^z, \quad (90)$$

onde  $\sigma^z$  é o operador Pauli usual na direção z e  $\omega_1$  e  $\omega_2$  representam as autoenergias dos qubits do sistema  $s_1$  e  $s_2$ , respectivamente, com seus operadores de auto-evolução,  $U_{s_1} = \exp(-iH_{s_1}\delta t_s)$  e  $U_{s_2} = \exp(-iH_{s_2}\delta t_s)$ . Além disso, supomos que o ambiente em ambos os modelos de colisão é composto por qubits idênticos, todos iniciados em seu estado fundamental, ou seja, os estados iniciais dos qubits do ambiente  $e_n, e_{n+1}, \dots$  são  $|e_n\rangle = |e_{n+1}\rangle = \dots = |0\rangle$ , e não existem correlações iniciais entre os qubits do sistema e os qubits do ambiente.

Agora, vamos continuar com os detalhes do primeiro modelo de colisão, que é esquematicamente descrito na Fig. 21(a). Neste modelo, o qubit do sistema  $s_2$  está acoplado ao ambiente, enquanto  $s_1$  está isolado dele e só pode interagir diretamente com  $s_2$ . Este modelo pode ser resumido em algumas etapas da seguinte forma. Primeiro, o qubit do sistema  $s_2$  interage com o qubit do ambiente  $e_n$  por meio do Hamiltoniano,

$$H_{s_2e_n} = \frac{J}{2}(\sigma_{s_2}^x \sigma_{e_n}^x + \sigma_{s_2}^y \sigma_{e_n}^y), \quad (91)$$

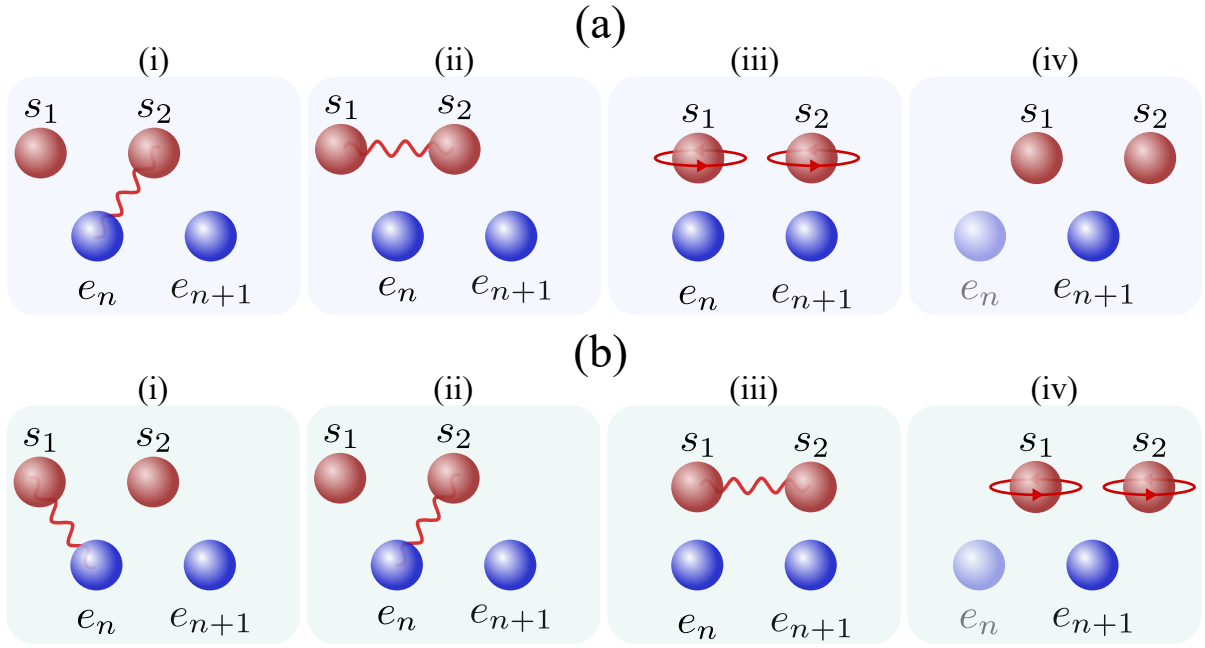


Figura 21 – (a) Primeiro modelo de colisão simulando dissipação local para o qubit do sistema  $s_2$ : (i) O qubit do sistema  $s_2$  interage com o qubit do ambiente  $e_n$ . (ii) Os qubits do sistema  $s_1$  e  $s_2$  interagem diretamente entre si. (iii) Os qubits do sistema  $s_1$  e  $s_2$  evoluem livremente. (iv) O qubit do ambiente que interagiu com o qubit do sistema  $s_2$  na etapa (i) é traçado, e o mesmo ciclo é repetido com o próximo qubit do ambiente  $e_{n+1}$ . (b) Segundo modelo de colisão simulando dissipação global para os qubits do sistema  $s_1$  e  $s_2$ : (i) O qubit do sistema  $s_1$  interage com o qubit do ambiente  $e_n$ . (ii) O qubit do sistema  $s_2$  interage com o mesmo qubit do ambiente  $e_n$ . (iii) Os qubits do sistema  $s_1$  e  $s_2$  interagem diretamente. (iv) Os qubits do sistema  $s_1$  e  $s_2$  evoluem livremente, e o qubit do ambiente que interagiu com os qubits do sistema na etapa (i-ii) é traçado para continuar o ciclo com o próximo qubit do ambiente  $e_{n+1}$ .

onde  $\sigma^x$  e  $\sigma^y$  são os operadores de Pauli padrão nas direções x e y, com o operador de evolução unitária correspondente dado por  $U_{se} = \exp(-iH_{se}\delta t_{se})$ . Aqui, o parâmetro  $J\delta t_{se}$  quantifica a intensidade do acoplamento entre o qubit do sistema aberto  $s_2$  e os qubits do ambiente. Em seguida, os qubits do sistema  $s_1$  e  $s_2$  interagem diretamente entre si, descritos pelo Hamiltoniano,

$$H_{s_1s_2} = \frac{\lambda}{2}(\sigma_{s_1}^x \sigma_{s_2}^x + \sigma_{s_1}^y \sigma_{s_2}^y), \quad (92)$$

onde o operador  $U_{ss} = \exp(-iH_{s_1s_2}\delta t_{ss})$  descreve sua dinâmica com  $\lambda\delta t_{ss}$  sendo a intensidade do acoplamento entre  $s_1$  e  $s_2$ . Na terceira etapa, os qubits do sistema  $s_1$  e  $s_2$  evoluem livremente com seus auto-Hamiltonianos. Na última etapa, uma única iteração do modelo de colisão é finalizada traçando o qubit do ambiente  $e_n$  e seguindo em frente para repetir o procedimento descrito acima com o qubit do ambiente  $e_{n+1}$ . Este ciclo é repetido  $N$  vezes para obter a evolução do sistema aberto.

No segundo modelo de colisão, diferentemente do primeiro modelo que discutimos, ambos os qubits do sistema,  $s_1$  e  $s_2$ , estão abertos para interação com o ambiente. Descrevemos

um único ciclo no modelo a seguir, que também é resumido visualmente na Fig. 21(b). Na etapa inicial, o qubit do sistema  $s_1$  interage com o spin do ambiente  $e_n$  por meio do Hamiltoniano de interação,

$$H_{s_1 e_n} = \frac{J}{2}(\sigma_{s_1}^x \sigma_{e_n}^x + \sigma_{s_1}^y \sigma_{e_n}^y). \quad (93)$$

Em seguida, o qubit do sistema  $s_2$  interage com o mesmo spin do ambiente  $e_n$  através do mesmo Hamiltoniano, ou seja,

$$H_{s_2 e_n} = \frac{J}{2}(\sigma_{s_2}^x \sigma_{e_n}^x + \sigma_{s_2}^y \sigma_{e_n}^y). \quad (94)$$

O operador de evolução unitária associado é dado por  $U_{se} = \exp(-iH_{se}\delta t_{se})$  para a interação de ambos os qubits do sistema com o ambiente, e  $J\delta t_{se}$  quantifica a intensidade do acoplamento. Após as interações sistema-ambiente, os qubits do sistema  $s_1$  e  $s_2$  interagem diretamente entre si com o Hamiltoniano,

$$H_{s_1 s_2} = \frac{\lambda}{2}(\sigma_{s_1}^x \sigma_{s_2}^x), \quad (95)$$

onde o operador de evolução unitária correspondente se torna  $U_{ss} = \exp(-iH_{s_1 s_2}\delta t_{ss})$ , e  $\lambda\delta t_{ss}$  é a intensidade da interação qubit-qubit. Por fim, os qubits do sistema  $s_1$  e  $s_2$  evoluem livremente com seus auto-Hamiltonianos dados na Eq. (90), e uma única iteração no modelo é concluída traçando o qubit do ambiente  $e_n$  para continuar a dinâmica com o próximo qubit do ambiente  $e_{n+1}$  que está por vir.

Por fim, destacamos brevemente a diferença fundamental entre os dois modelos de colisão que consideramos neste trabalho, conforme ilustrado na Fig. 21. Enquanto o primeiro modelo de colisão descreve um cenário de dissipação local, no qual apenas um qubit do sistema está aberto para interação com o ambiente, o segundo modelo permite que ambos os qubits do sistema se acoplem aos graus de liberdade do ambiente. Em outras palavras, enquanto o primeiro modelo na verdade simula uma dissipação local afetando um único qubit no sistema de dois qubits, o segundo serve como um exemplo de dissipação global, onde ambos os qubits estão sob o efeito do ruído. Por outro lado, também vale ressaltar que, apesar de ambos os modelos de colisão compartilharem vários parâmetros de interação em comum, o Hamiltoniano de interação qubit-qubit nesses modelos não é exatamente o mesmo, como pode ser visto ao comparar a Eq. (92) e a Eq. (95). De fato, consideramos esses Hamiltonianos diferentes para demonstrar a independência de nossos resultados de uma forma específica de interação qubit-qubit na previsão da sincronização com base em aprendizado de máquina.

### 6.1.2 Modelo de Equação Mestra

O modelo que consideramos nesta parte pode ser utilizado para descrever o acoplamento entre um par de átomos de dois níveis e um ambiente eletromagnético quantizado térmico. Aqui, definimos  $\hbar = 1$  e ajustamos as unidades dos outros parâmetros de acordo. Em seguida, os auto-Hamiltonianos dos átomos do sistema são dados por  $H_s = \sum_{i=1}^2 \omega_i \sigma_i^z$ , onde  $\omega_i$  representa

a frequência de transição entre os níveis de energia do  $i$ -ésimo átomo, e  $\sigma^z$  denota o operador de Pauli usual na direção  $z$ . Também supomos que os átomos do sistema possuem momentos de dipolo polarizados,  $d_{eg}$ , e que eles estão acoplados entre si por meio do Hamiltoniano de interação de troca  $H_d = \sum_{i \neq j}^2 f_{ij} \sigma_i^+ \sigma_j^-$ , onde  $f_{ij}$  quantifica a intensidade da interação átomo-átomo, e  $\sigma^\pm$  são os operadores de subida e descida de átomos de dois níveis. Considerando o acoplamento dos átomos a fótons térmicos e focando na dinâmica reduzida de sistemas atômicos de dois níveis, obtemos a conhecida equação mestra da óptica quântica dada por (BREUER; PETRUCCIONE, 2002; DAMANET; MARTIN, 2016)

$$\dot{\rho} = -i[(H_s + H_d), \rho] + \mathcal{D}_-(\rho) + \mathcal{D}_+(\rho) = \mathcal{L}(\rho), \quad (96)$$

onde  $\mathcal{D}_-(\rho)$  e  $\mathcal{D}_+(\rho)$  são definidos como

$$\mathcal{D}_-(\rho) = \sum_{i,j=1}^2 \gamma_{ij}(\bar{n} + 1)(\sigma_j^- \rho \sigma_i^+ - \frac{1}{2}\{\sigma_i^+ \sigma_j^-, \rho\}), \quad (97)$$

$$\mathcal{D}_+(\rho) = \sum_{i,j=1}^2 \gamma_{ij}\bar{n}(\sigma_j^+ \rho \sigma_i^- - \frac{1}{2}\{\sigma_i^- \sigma_j^+, \rho\}), \quad (98)$$

onde  $\rho$  representa o operador de densidade para o par de átomos. Na Eq. (96), como o primeiro termo representa a autoevolução unitária e a interação de troca entre os átomos, o segundo termo detalhado na Eq. (97) descreve os processos de emissão espontânea e induzida termicamente. Por fim, o terceiro termo fornecido na Eq. (98) é responsável pelo processo de absorção induzida termicamente. A taxa com que esses processos ocorrem é determinada pelo número médio de fótons  $\bar{n}$  na frequência de transição. As taxas de decaimento na equação mestra da óptica quântica são expressas como  $\gamma_{ij} = \sqrt{\gamma_i \gamma_j} a(k_0 r_{ij})$  onde  $\gamma_{i(j)} = \omega_{i(j)}^3 g$ ,  $k_0 = (\omega_1 + \omega_2)/2c$  e  $g = d_{eg}^2/3\pi\epsilon_0 c^3$ . Além disso, a condição de que  $a(k_0 r_{ij}) \leq 1$  precisa ser satisfeita para garantir a positividade da dinâmica. Aqui,  $d_{eg}$ ,  $c$ ,  $\epsilon_0$ , e  $r_{ij}$  representam respectivamente o momento de dipolo idêntico dos átomos, a velocidade da luz no vácuo, a constante de permissividade do espaço livre e a posição relativa dos átomos  $i$ -ésimo e  $j$ -ésimo. Enquanto  $a_{ij}$  controla o grau de comportamento coletivo na dinâmica dos átomos,  $f_{ij}$  quantifica a intensidade da interação de troca entre eles. As formas explícitas desses parâmetros do modelo podem ser encontradas em Refs. (DAMANET; MARTIN, 2016; BHATTACHARYA; BANERJEE; PATI, 2018; LEHMBERG, 1970). Em nossa análise, nos concentramos no caso de temperatura zero,  $\bar{n} = 0$ , e assumimos que o sistema de dois átomos passa por uma dinâmica completamente coletiva, ou seja,  $a_{ij} = 1$ .

## 6.2 Estrutura dos dados

Os dados são uma parte crucial de todas as técnicas de aprendizagem de máquina, pois os algoritmos de aprendizagem são treinados em conjuntos de dados. Para treinar esses modelos de aprendizagem, é necessária uma estrutura de dados que possa representar de forma

eficaz as informações obtidas a partir de simulações físicas. Tendo introduzido os três diferentes modelos de sistemas abertos que pretendemos considerar neste capítulo, iremos agora detalhar a formação dos bancos de dados a serem utilizados pelo algoritmo de aprendizagem de máquina para prever o comportamento de sincronização dos qubits.

Em nossa análise, as variáveis em estudo são os valores esperados locais dos observáveis de qubit do sistema, pois pretendemos explorar a harmonia entre suas dinâmicas para observar o surgimento de sincronização espontânea. Devido ao fato de que a sincronização é uma propriedade da dinâmica do sistema aberto e não de uma escolha particular de observável, é, em princípio, possível escolher um observável arbitrário de qubit para estudar as propriedades de sincronização do sistema. Neste capítulo, focamos nos valores esperados dos observáveis dos qubits do sistema ao longo da direção  $x$ , ou seja,  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , para fins de concretude. Para determinar o eventual comportamento de sincronização do par de qubits, em geral, é necessário acompanhar a dinâmica de longo prazo dos valores esperados locais dos qubits e, assim, da própria dinâmica do sistema aberto. No entanto, consideraremos apenas de 5 a 100 valores esperados iniciais para cada qubit como características (entradas) do algoritmo para demonstrar a capacidade de nossa abordagem de aprendizagem de máquina de antecipar com sucesso o comportamento de sincronização de longo prazo do sistema aberto. Além disso, como é comum na literatura, quantificamos o grau de sincronização entre os qubits usando o coeficiente de Pearson. Particularmente, definimos o coeficiente de Pearson como o alvo (saída) do algoritmo de aprendizagem e o calculamos em simulações para os 100 últimos valores esperados dos observáveis dos qubits.

Vamos também introduzir a definição do coeficiente de Pearson e explicar brevemente sua importância na quantificação do comportamento síncrono entre os observáveis dos qubits. O coeficiente de Pearson é uma medida estatística que pode ser usada para quantificar o grau de associação linear entre duas variáveis discretas. Ele é definido como

$$C_{12} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (99)$$

onde  $\bar{x}$  e  $\bar{y}$  são os valores médios de  $x$  e  $y$ , respectivamente, e  $n$  é o número de valores que as variáveis podem assumir. O coeficiente de Pearson  $C_{12}$  pode assumir valores no intervalo  $[-1, 1]$ . Enquanto o caso  $C_{12} = 0$  indica que as variáveis não têm correlação,  $C_{12} = 1$  ( $C_{12} = -1$ ) significa que elas estão perfeitamente correlacionadas positivamente (negativamente). Em particular, enquanto a correlação positiva (negativa) perfeita indica que, à medida que uma variável muda, a outra variável muda da mesma forma (oposta). Em nosso trabalho, as variáveis estudadas são os valores esperados dos observáveis locais dos qubits  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ . No contexto do nosso estudo, estamos interessados no comportamento de sincronização assintótico do par de qubits no limite de longo prazo. Portanto, em simulações numéricas, geramos os valores esperados como amostras discretas que cobrem os últimos 100 valores esperados para cada observável de qubit local, garantindo também que escolhamos um valor suficientemente grande

de  $n$  para cada modelo. Em outras palavras, enquanto o valor assintótico de  $C_{12} \approx 0$  implica oscilações não correlacionadas para os valores esperados locais no limite de longo prazo,  $C_{12} \approx 1$  ( $C_{12} \approx -1$ ) significa que a sincronização completa (antissincronização) é estabelecida entre eles. Além disso, o coeficiente de Pearson  $C_{12}$  também pode atingir valores assintóticos diferentes de 1 ou -1. Isso sugere que as oscilações mútuas dos valores esperados ainda estão bloqueadas em fase no limite de longo prazo, entretanto, a diferença de fase entre elas não é nem 0, como na sincronização completa, nem  $\pi$ , como na antissincronização completa. Esses casos são conhecidos como sincronização com atraso temporal (GALVE; GIORGI; ZAMBRINI, 2017).

Antes de concluir esta seção, queremos discutir como levamos em consideração erros potenciais nos valores esperados dos qubits em nossas simulações, que podem ser causados por diversas fontes, como medições imperfeitas ou decoerência, a fim de enfatizar a robustez de nossa abordagem em um ambiente experimental. Para simular tais erros, adicionamos erros aleatórios de diferentes porcentagens aos valores esperados para verificar se o algoritmo de aprendizagem de máquina considerado ainda seria capaz de prever corretamente os valores para nosso alvo na presença de erro. O erro é adicionado somando-se os valores esperados a um valor aleatório  $\epsilon$  entre -1 e 1, multiplicado por uma taxa de erro entre 0,005 e 0,05. Isso nos permite determinar a precisão do algoritmo em termos de desvio em relação aos valores esperados calculados.

Vamos utilizar o algoritmo *k-Nearest Neighbors* (*k-NN*), que tem como objetivo calcular a proximidade entre pontos de dados usando uma métrica, como a distância euclidiana, e, em seguida, estima a saída de uma amostra de teste tirando a média das saídas de seus  $k$  vizinhos mais próximos nos dados de treinamento (ALPAYDIN, 2020). A implementação do algoritmo *k-NN* neste trabalho foi realizada utilizando a biblioteca `scikit-learn` em Python (PEDREGOSA et al., 2011). Todos os parâmetros do algoritmo são mantidos em suas configurações padrão.

## 6.3 Resultados

Nesta seção, apresentamos nossos principais resultados relacionados à implementação do algoritmo *k-NN* para prever o comportamento de sincronização de longo prazo para um par de qubits nos três modelos distintos de sistemas abertos introduzidos anteriormente.

### 6.3.1 Modelos de Colisão

Começaremos esta seção considerando o primeiro modelo de colisão, descrito na Fig. 21(a), onde apenas um dos qubits do sistema está acoplado diretamente ao ambiente. Definimos os estados iniciais dos qubits do sistema  $s_1$  e  $s_2$  como  $|\psi_{s_1}\rangle = |\psi_{s_2}\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ . Neste ponto, enfatizamos que o surgimento da sincronização não depende da escolha dos estados iniciais, pois é uma propriedade do processo dinâmico. Além disso, fixamos os parâmetros do

modelo como  $\delta t_s = \delta t_{ss} = 1.0$  e  $\delta t_{se} = 1.0$ . Para criar nosso conjunto de dados, consideramos os valores da *detuning*  $\omega_1/\omega_2$ , variando de 0.97 a 1.03 em passos de 0.0015, a intensidade de acoplamento direto  $\lambda$ , variando de 0.01 a 0.06 em passos de 0.002, e a intensidade de interação sistema-ambiente  $J$ , variando de 0.05 a 0.15 em passos de 0.01. Isso resulta em um total de 10.250 configurações diferentes para os parâmetros do modelo. Para cada uma dessas configurações, calculamos a dinâmica de sistema aberto do modelo e avaliamos os valores esperados dos observáveis para os qubits do sistema,  $\langle \sigma_{s_1}^x \rangle = \text{Tr}[\rho_{s_1} \sigma^x]$  e  $\langle \sigma_{s_2}^x \rangle = \text{Tr}[\rho_{s_2} \sigma^x]$ , após cada colisão, em um total de 4000 colisões, onde  $\rho_{s_1}$  e  $\rho_{s_2}$  denotam os operadores densidade reduzidos para os qubits do sistema  $s_1$  e  $s_2$ , respectivamente. Confirmando que o sistema atingiu seu comportamento dinâmico de longo prazo com esses dados, calculamos o coeficiente de correlação de Pearson  $C_{12}$  levando em conta os valores esperados dos qubits do sistema calculados para as últimas 100 colisões. Na Fig. 22(a), mostramos o coeficiente de Pearson  $C_{12}$  em função de  $\omega_1/\omega_2$  e da intensidade de acoplamento direto  $\lambda$  para uma intensidade de interação sistema-ambiente fixa, ou seja,  $J = 0.1$ , para fins de apresentação. Observa-se que as regiões de sincronização e antissincronização são separadas pela linha de ressonância onde a sincronização está ausente (KARPAT et al., 2021).

Para verificar a capacidade de nossa abordagem baseada em aprendizagem de máquina de prever as propriedades de sincronização deste modelo de sistema aberto, calculamos os valores esperados correspondentes  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$  para as primeiras 100 colisões (características) e o coeficiente de Pearson  $C_{12}$ , avaliado para as últimas 100 colisões (alvo) para treinar o algoritmo  $k$ - $NN$ . Para evitar que as previsões de  $C_{12}$  fossem baseadas diretamente nos dados de treino da Fig. 22, repetimos toda a simulação considerando pequenos ajustes  $+\varepsilon$  nos parâmetros  $\omega_1/\omega_2$  e  $\lambda$ , onde  $-0.0001 < \varepsilon < 0.0001$ . Isso nos permitiu treinar o algoritmo  $k$ - $NN$  usando o conjunto de dados original e avaliar seu desempenho de previsão em um novo conjunto de dados independente. Na Fig. 22(b), mostramos os valores previstos do coeficiente de Pearson  $C_{12}$ , novamente fixando  $J = 0.1$ . No entanto, observamos que essa escolha específica de  $J$  é feita apenas para fins de visualização, e que o algoritmo  $k$ - $NN$  foi de fato treinado com uma faixa de valores de  $J$ . Apresentamos os gráficos calculados e previstos de  $C_{12}$  lado a lado na Fig. 22. Como pode ser visto ao comparar os gráficos do coeficiente de Pearson calculado e previsto  $C_{12}$ , o algoritmo  $k$ - $NN$  parece desempenhar muito bem na previsão dos diferentes comportamentos de sincronização neste modelo, ou seja, regimes de sincronização/antissincronização e ausência de sincronização. Para quantificar a precisão do algoritmo, calculamos o erro médio absoluto (MAE) para os valores previstos em comparação com os valores calculados usando a expressão

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|, \quad (100)$$

onde  $n$  é o número de pontos de dados,  $y_i$  é o valor calculado de  $C_{12}$  e  $\hat{y}_i$  é o valor previsto de  $C_{12}$ . Ou seja, MAE é simplesmente o valor absoluto da diferença entre os valores calculados e previstos, médio sobre todos os pontos de dados. Para os resultados de previsão exibidos na Fig. 22, encontramos  $MAE = 0.009$ , o que demonstra o sucesso do algoritmo  $k$ - $NN$  em inferir

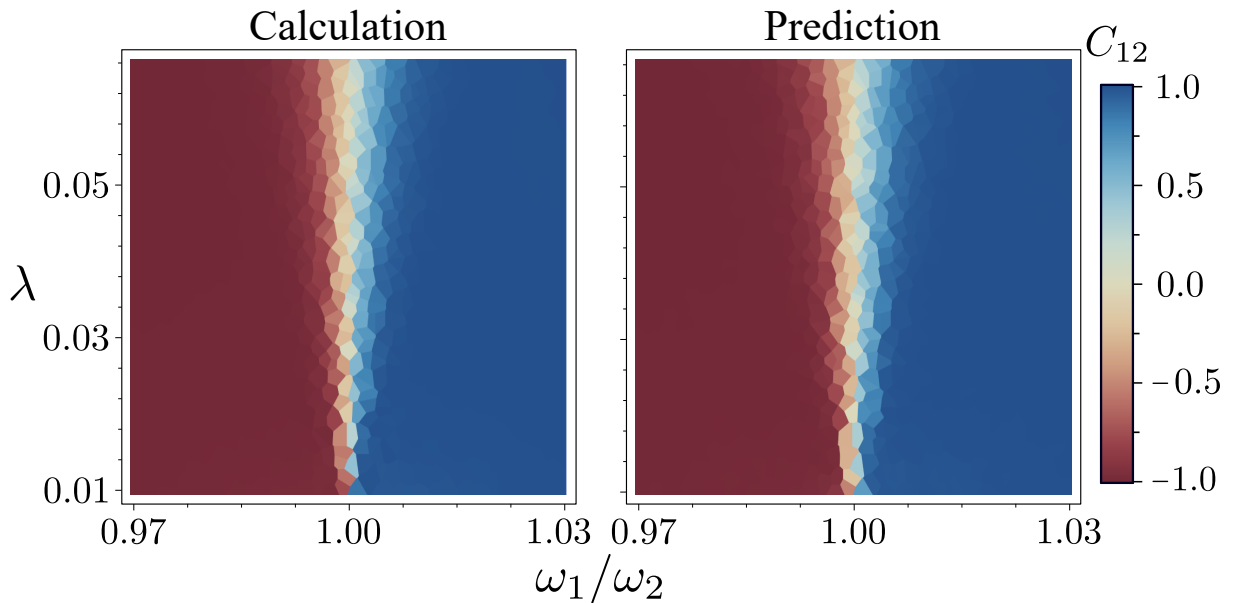


Figura 22 – O primeiro modelo de colisão: (a) O coeficiente de Pearson  $C_{12}$ , calculado para os últimos 100 valores esperados dos qubits do sistema  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , em função da intensidade de interação direta  $\lambda$  entre eles e do *detuning* entre suas autoenergias  $\omega_1/\omega_2$ , para a intensidade de acoplamento sistema-ambiente  $J = 0.1$ . (b) Valores previstos de  $C_{12}$  fornecidos pelo regressor  $k$ - $NN$  usando os primeiros 100 valores esperados do par de qubits em função de  $\lambda$  e  $\omega_1/\omega_2$  para  $J = 0.1$ .

o comportamento de sincronização de longo prazo do modelo, mesmo tendo acesso apenas aos valores esperados de curto prazo dos qubits do sistema calculados para as primeiras 100 colisões.

Além disso, para examinar um cenário mais realista, levamos em consideração erros potenciais na medição dos valores esperados locais dos qubits, que podem ocorrer naturalmente em um ambiente experimental. Conforme mencionado anteriormente, simulamos esses erros somando aos valores esperados um número aleatório  $\epsilon$  retirado do intervalo  $\epsilon \in [-1, 1]$ , que é então multiplicado por uma taxa de erro entre 0,005 e 0,05. Em outras palavras, consideramos uma taxa de erro percentual entre %0,5 e %5 nas medições dos valores esperados de ambos os qubits. Realizamos também a mesma análise exibida na Fig. 22, onde os primeiros 100 valores esperados são usados como entrada do regressor  $k$ - $NN$ , mas desta vez usando os primeiros 5, 10 e 50 valores esperados do par de qubits como entrada do algoritmo. Resumimos os resultados de nossas simulações na Fig. 23, onde as linhas vermelha pontilhada, verde ponto-tracejada, laranja tracejada e azul sólida mostram como o MAE para o coeficiente de Pearson  $C_{12}$  escala com o erro percentual adicionado aos valores esperados dos qubits,  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , quando os valores esperados após as primeiras 100, 50, 10 e 5 colisões são usados para previsão, respectivamente. Com base nesses resultados, é bastante claro que existe uma relação de *trade-off* entre o número de valores esperados dos qubits usados para a previsão de  $C_{12}$  e a robustez do algoritmo de regressão  $k$ - $NN$  contra erros. Em particular, as previsões do algoritmo em relação às propriedades de sincronização do modelo são bastante suscetíveis mesmo a

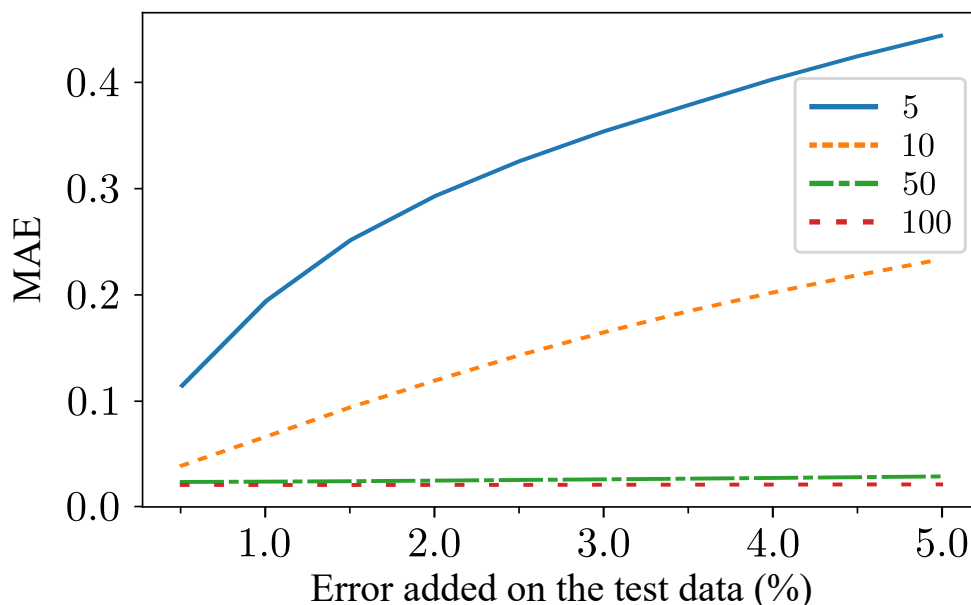


Figura 23 – Erro médio absoluto (MAE) para o coeficiente de Pearson  $C_{12}$  no primeiro modelo de colisão em função do erro percentual que é adicionado aos valores esperados calculados dos qubits  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ . Quatro linhas diferentes na legenda correspondem aos casos de fornecimento ao regressor  $k$ - $NN$  de 5, 10, 50 e 100 pares de valores esperados dos qubits para previsão.

um pequeno percentual de erros quando sua entrada envolve um número reduzido de valores esperados dos qubits, como pode ser visto no caso de 5 valores esperados para cada qubit mostrado pela linha azul sólida na Fig. 23. Por outro lado, observamos que, quando o algoritmo tem acesso a 50 ou 100 valores esperados para cada qubit, ele pode fazer previsões bastante precisas do coeficiente de Pearson  $C_{12}$ . De fato, nesses casos, o aumento do percentual de erro nos valores esperados de 0,5% para 5% muda apenas marginalmente o MAE, o que nem é visível a olho nu na Fig. 23. Em suma, esses resultados sugerem que, no caso do primeiro modelo de colisão que descreve a dissipação local de um dos qubits do par, nossa abordagem de aprendizagem de máquina baseada no algoritmo de regressão  $k$ - $NN$  é bastante eficaz em prever com precisão diferentes regimes de sincronização ou sua ausência. Por fim, enfatizamos mais uma vez que de 50 a 100 medições de valores esperados para cada qubit nas primeiras etapas da dinâmica são suficientes para prever o comportamento de sincronização de longo prazo deste modelo com alta precisão, apesar da presença de erros de medição em um determinado ambiente experimental.

Tendo demonstrado o desempenho de nossa abordagem no primeiro modelo de colisão, voltamos agora nossa atenção para o segundo modelo de colisão, representado na Fig. 21(b), que descreve a dissipação global do par de qubits no sentido de que ambos interagem com a mesma partícula ambiental em cada iteração. Da mesma forma que no primeiro modelo de colisão, definimos os estados iniciais dos qubits do sistema  $s_1$  e  $s_2$  como  $|\psi_{s_1}\rangle = |\psi_{s_2}\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$  e escolhemos os parâmetros do modelo como  $\delta t_s = \delta t_{ss} = 0.2$ ,  $\delta t_{se} = 0.05$  e  $J = 1.0$ . Após fixar os parâmetros do modelo e do sistema, criamos nosso conjunto de dados considerando os

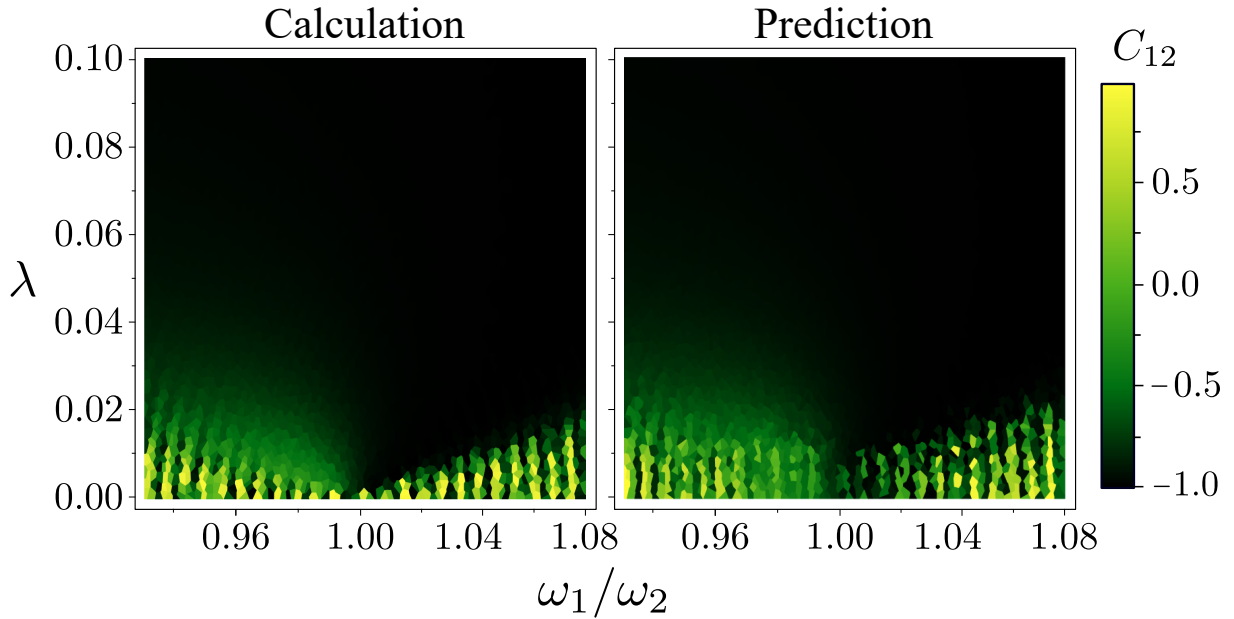


Figura 24 – O segundo modelo de colisão: (a) O coeficiente de Pearson  $C_{12}$ , calculado para os últimos 100 valores esperados dos qubits do sistema  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , em função da intensidade de interação direta  $\lambda$  entre eles e da *detuning* entre suas autoenergias  $\omega_1/\omega_2$ . (b) Valores previstos de  $C_{12}$  fornecidos pelo regressor  $k$ - $NN$  usando os primeiros 100 valores esperados dos qubits em função de  $\lambda$  e  $\omega_1/\omega_2$ .

valores de *detuning*  $\omega_1/\omega_2$  de 0.93 a 1.08 em passos de 0.002 e a intensidade de acoplamento direto  $\lambda$  de 0.0 a 0.01 em passos de 0.002. Isso nos dá um total de 3876 configurações diferentes para os parâmetros do modelo. Obtemos a evolução do sistema aberto do modelo e os valores esperados  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$  para cada uma dessas configurações durante 5000 colisões, após as quais os valores esperados tornam-se insignificantes. Em seguida, calculamos o coeficiente de Pearson  $C_{12}$  para os valores esperados dos qubits do sistema considerando as últimas 100 colisões para observar o comportamento de sincronização de longo prazo. Na Fig. 24(a), o coeficiente de Pearson  $C_{12}$  é representado em função da intensidade de acoplamento direto  $\lambda$  e da *detuning*  $\omega_1/\omega_2$  entre as autoenergias dos qubits. Diferentemente do primeiro modelo de colisão, onde tanto sincronização quanto antissincronização podem ocorrer dependendo da *detuning* entre as autoenergias dos qubits, aqui é possível observar apenas antissincronização, que surge se o acoplamento direto entre o par de qubits for suficientemente forte para compensar o *detuning*.

Usando a mesma abordagem seguida na análise do primeiro modelo de colisão, treinamos o regressor  $k$ - $NN$  com os valores esperados correspondentes  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$  calculados para as primeiras 100 colisões e o coeficiente de Pearson  $C_{12}$  avaliado para as últimas 100 colisões. Na Fig. 24(b), mostramos as previsões do coeficiente de Pearson  $C_{12}$  realizadas pelo algoritmo  $k$ - $NN$ . Também enfatizamos que evitamos que o regressor tenha acesso aos dados de treino ajustando-o novamente com pequenos desvios  $\varepsilon$  nos parâmetros do modelo  $\lambda$  e  $\omega_1/\omega_2$ , conforme foi feito para o caso do primeiro modelo de colisão. Comparando diretamente os gráficos dos valores calculados e previstos de  $C_{12}$  lado a lado na Fig. 24, observa-se que o algoritmo de

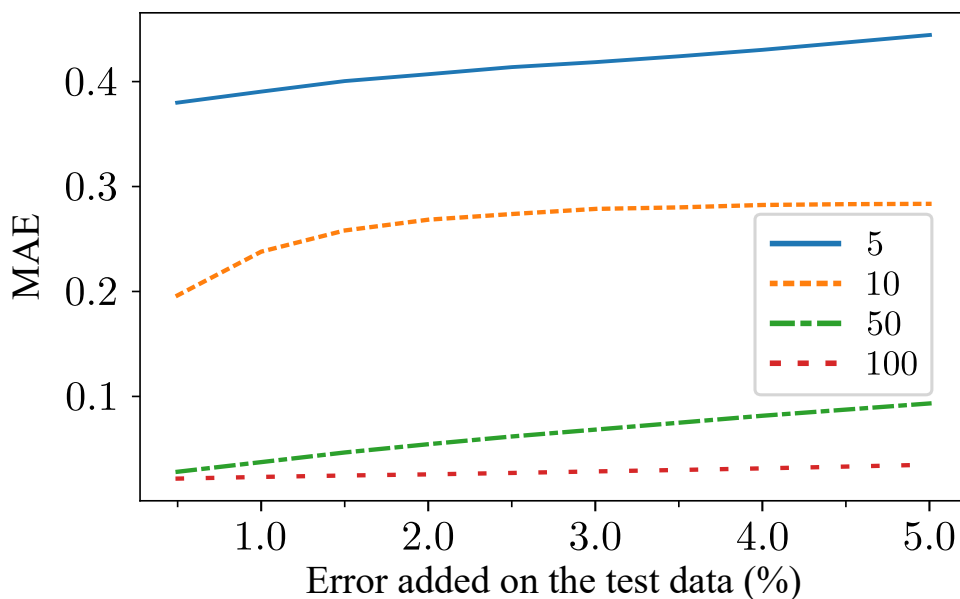


Figura 25 – Erro médio absoluto (MAE) para o coeficiente de Pearson  $C_{12}$  no segundo modelo de colisão em função do erro percentual adicionado aos valores esperados calculados  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ . Quatro linhas diferentes na legenda correspondem aos casos em que fornecemos ao regressor  $k$ - $NN$  de 5, 10, 50 e 100 pares de valores esperados dos qubits para previsão.

regressão  $k$ - $NN$  atinge alta precisão na previsão das propriedades de sincronização de longo prazo do modelo considerado. De fato, o MAE nesse caso é de 0.040. Por fim, consideramos possíveis imperfeições de medição nos valores esperados dos qubits. Levando em conta erros percentuais aleatórios nos valores esperados dos qubits  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$  entre %0,5 e %5, repetimos a mesma investigação para entender quão resiliente é nossa abordagem de aprendizagem de máquina, baseada no regressor  $k$ - $NN$ , contra potenciais erros na entrada do algoritmo. Apresentamos os resultados de nossos cálculos na Fig. 25, onde os primeiros 5, 10 e 50 valores esperados do par de qubits são considerados como entrada, além do caso dos primeiros 100 valores esperados. Na Fig. 25, as linhas vermelha pontilhada, verde ponto-tracejada, laranja tracejada e azul sólida demonstram como o MAE para  $C_{12}$  varia com o erro percentual nos valores esperados do par de qubits quando os primeiros 100, 50, 10 e 5 valores esperados são usados para previsão, respectivamente. Observamos que o algoritmo de regressão  $k$ - $NN$  pode fornecer previsões confiáveis sobre a sincronização também no segundo modelo de colisão, mesmo na presença de erros de medição em um cenário realista, desde que o número de valores esperados para cada qubit, usados como variáveis de entrada, seja suficiente.

Para resumir esta subseção, vemos que nossa abordagem de aprendizagem de máquina para a previsão de sincronização em dinâmicas de sistemas abertos mostra-se bem-sucedida em dois modelos distintos. Ou seja, tanto nos regimes de dissipação local quanto global descritos no contexto do modelo de colisão, o regressor  $k$ - $NN$  consegue prever com precisão o surgimento de (anti-)sincronização ou sua ausência nas dinâmicas de longo prazo, utilizando apenas os valores esperados iniciais dos observáveis dos qubits considerados.

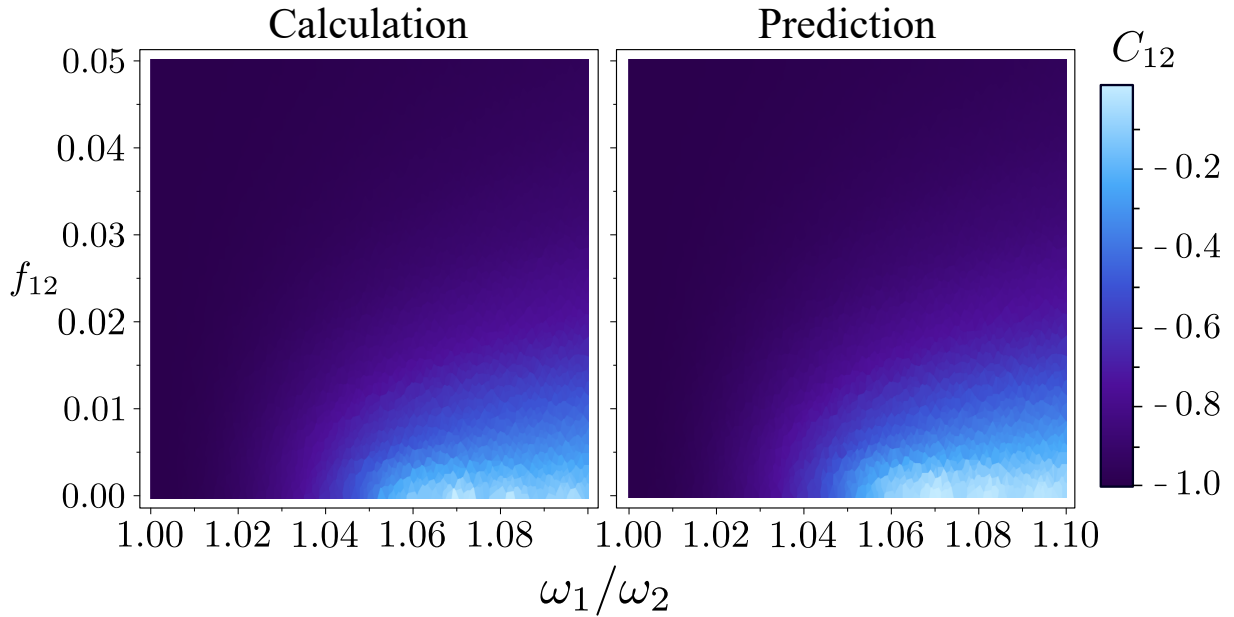


Figura 26 – O modelo de equação mestra: (a) O coeficiente de Pearson  $C_{12}$ , calculado para os últimos 100 valores esperados dos qubits  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , em função da intensidade de interação de troca  $f_{12}$  entre eles e da *detuning* entre suas autoenergias  $\omega_1/\omega_2$ . (b) Valores previstos de  $C_{12}$  produzidos pelo regressor  $k$ - $NN$  usando os primeiros 100 valores esperados do par de qubits em termos de  $f_{12}$  e  $\omega_1/\omega_2$ .

### 6.3.2 Modelo de Equação Mestra

Até este ponto, consideramos dois modelos de sistemas abertos que representam cenários de dissipação local e global, utilizando a estrutura do modelo de colisão. Vimos que nossa abordagem baseada em aprendizagem de máquina para a detecção do surgimento dinâmico de sincronização apresenta um desempenho bastante preciso em ambos os modelos, apesar de potenciais erros experimentais nas medições dos valores esperados dos qubits. Como último teste, consideramos um modelo de sistema aberto descrito por uma equação mestra, onde duas partículas atômicas de dois níveis (ou qubits) interagem de forma totalmente coletiva com um campo eletromagnético. Nesse contexto, dependendo do *detuning* entre as autoenergias dos qubits e do acoplamento direto entre eles, é possível observar oscilações de valores esperados com bloqueio de fase, onde a diferença de fase não é necessariamente  $\pi$  ou 0 graus, ou seja, sincronização com atraso temporal (KARPAT; YALÇINKAYA; ÇAKMAK, 2020).

Primeiro, vamos lembrar que consideraremos a dissipação coletiva do par de qubits no limite de temperatura zero, portanto fixamos  $\bar{n} = 0$  e  $a_{ij} = 1$ . Além disso, definimos os estados iniciais do par de qubits do sistema como  $|\psi_{s_1}\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$  e  $|\psi_{s_2}\rangle = (|0\rangle + e^{-i\pi/3}|1\rangle)/\sqrt{2}$ . Para criar nosso conjunto de dados, deixamos que o *detuning* entre as autoenergias dos qubits  $\omega_1/\omega_2$  adquira valores entre 1.0 e 1.1 em passos de 0.001, e a intensidade da interação de troca entre os qubits  $f_{12}$  entre 0 e 0.05 em passos de 0.001. Como resultado, obtemos 5151 diferentes configurações dos parâmetros do modelo, para cada uma das quais simulamos a dinâmica do sistema aberto e, em seguida, calculamos os valores esperados  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$  até

$t = 500$ . Observamos que, nesse ponto da dinâmica, o sistema já exibe seu comportamento de longo prazo e, aqui, simulamos a evolução temporal em passos de tempo discretos, ou seja, tomando  $t = 0, 1, \dots, 500$ . Em seguida, avaliamos o coeficiente de Pearson  $C_{12}$  para os valores esperados dos observáveis do par de qubits, levando em conta os últimos 100 passos de tempo na dinâmica, o que nos mostra o comportamento de sincronização de longo prazo do modelo de sistema aberto. Apresentamos os resultados de nossos cálculos na Fig. 26(a). É fácil notar que, quando a interação de troca entre os qubits é suficientemente forte para compensar o *detuning* entre suas autoenergias, surge uma antissincronização completa entre os valores esperados dos observáveis dos qubits. Além disso, para este modelo, sabe-se que, mesmo que os valores esperados não fiquem completamente antissincronizados, ainda ocorre sincronização com atraso temporal (KARPAT; YALÇINKAYA; ÇAKMAK, 2020). Em outras palavras, mesmo que o coeficiente de Pearson  $C_{12} \neq -1$  em longo prazo, ele ainda se estabiliza em um valor constante entre 0 e  $-1$ , conforme mostrado na Fig. 26(a), o que sinaliza o início de oscilações com uma diferença de fase fixa entre os valores esperados.

Finalmente, verificaremos a capacidade do regressor  $k$ - $NN$  de prever o comportamento de sincronização do par de qubits para o modelo de equação mestra considerado. Conforme feito para os modelos de colisão, alimentamos o regressor  $k$ - $NN$  com o conjunto de dados para essa configuração, que envolve os valores esperados dos qubits  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$  calculados nos primeiros 100 pontos de tempo, e o coeficiente de Pearson  $C_{12}$  avaliado considerando os últimos 100 pontos de tempo. Os valores previstos de  $C_{12}$  produzidos pelo algoritmo de regressão  $k$ - $NN$  são mostrados na Fig. 26(b) lado a lado com os valores calculados de  $C_{12}$ . É evidente a partir da comparação dos gráficos que nossa abordagem baseada no regressor  $k$ - $NN$  apresenta excelente desempenho na previsão das propriedades de sincronização do modelo. De fato, o MAE nesse caso é calculado como apenas 0.002.

Para completar esta seção, consideramos agora as possíveis imprecisões de medição nos valores esperados para avaliar a robustez de nossa abordagem baseada em aprendizagem de máquina contra erros. Semelhante aos modelos estudados anteriormente, adicionamos aleatoriamente erros percentuais aos valores esperados do par de qubits entre 0.5% e 5%, e realizamos a mesma análise considerando 5, 10, 50 e 100 valores esperados iniciais como entrada do regressor. A Fig. 27 exibe os resultados desta investigação, onde as linhas vermelha pontilhada, verde ponto-tracejada, laranja tracejada e azul sólida mostram como o MAE para o coeficiente de Pearson  $C_{12}$  varia em função do erro percentual adicionado aos valores esperados  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ , quando os valores esperados nos primeiros 100, 50, 10 e 5 pontos de tempo são dados ao regressor para previsão, respectivamente. É evidente a partir da Fig. 27 que, desde que 10 ou mais valores esperados iniciais do par de qubits sejam fornecidos ao algoritmo como entrada, o regressor  $k$ - $NN$  realiza uma excelente previsão do comportamento de sincronização de longo prazo do modelo, apesar da presença de erros nos valores esperados.

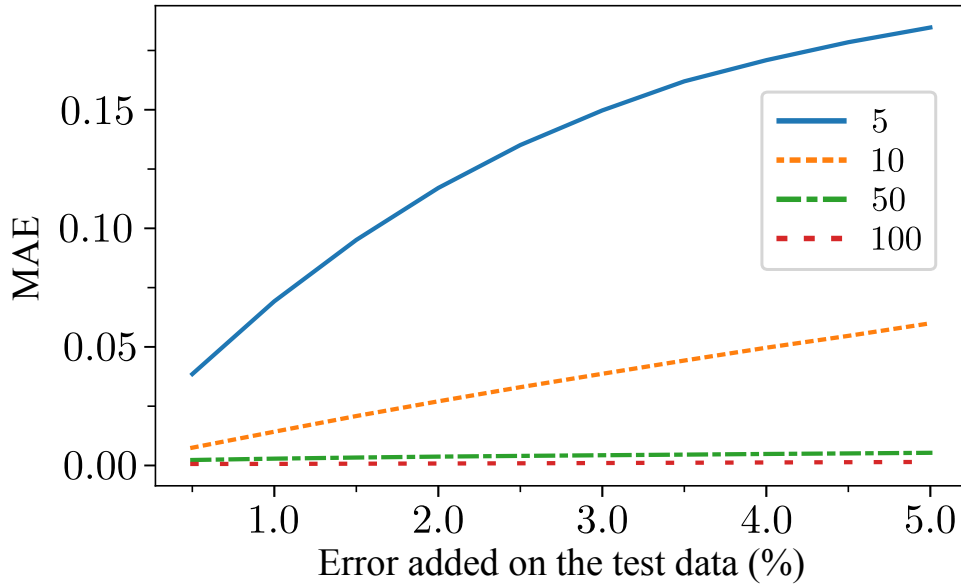


Figura 27 – Erro médio absoluto (MAE) para o coeficiente de Pearson  $C_{12}$  no modelo de equação mestra em função do erro percentual adicionado aos valores esperados calculados dos qubits  $\langle \sigma_{s_1}^x \rangle$  e  $\langle \sigma_{s_2}^x \rangle$ . Quatro linhas diferentes na legenda correspondem aos casos de fornecimento ao regressor  $k$ - $MN$  de 5, 10, 50 e 100 pares de valores esperados dos qubits para previsão.

### 6.3.3 Possíveis Mecanismos de Erro Físico

Nesta seção, investigaremos possíveis mecanismos de erro que podem afetar o estado do sistema quântico em questão, antes da etapa de previsão. Assumiremos que nosso algoritmo é treinado com os dados obtidos a partir do caso ideal e tentaremos avaliar o impacto de tais erros potenciais, quantificados via o MAE.

Primeiro, pretendemos examinar os efeitos de erros de inicialização, que podem ocorrer em experimentos, no estado inicial do sistema. Para levar em conta condições realistas, consideramos agora um estado inicial modificado para todos os três modelos, descrito pela matriz densidade,

$$\rho = p|++\rangle\langle ++| + (1-p)(|+-\rangle\langle +-| + |-+\rangle\langle -+|). \quad (101)$$

Essa modificação acomoda o potencial de os estados dos qubits serem incorretamente inicializados por meio de dois mecanismos específicos: um erro de fase em que um estado pretendido  $|+\rangle$  é inicializado como  $|-\rangle$ , ou um erro de bit-flip, uma vez que o estado  $|+\rangle$  pode ser gerado a partir do estado  $|0\rangle$  via uma porta de Hadamard. Neste último caso,  $1-p$  reflete a probabilidade de que a inicialização pretendida para o estado  $|0\rangle$  resulte, na verdade, no estado  $|1\rangle$ , levando à criação não intencional de um estado  $|-\rangle$  pela transformação de Hadamard. Tais erros são relativamente comuns, especialmente em plataformas de qubits supercondutores, onde sempre há uma probabilidade não nula de medir o estado  $|1\rangle$  ( $|0\rangle$ ) quando a intenção era criar o estado  $|0\rangle$  ( $|1\rangle$ ) (IBM Quantum, 2024). Na Fig. 28, apresentamos o MAE em função de  $1-p$ , ou seja, a probabilidade de ocorrência de um erro de inicialização, o que nos mostra

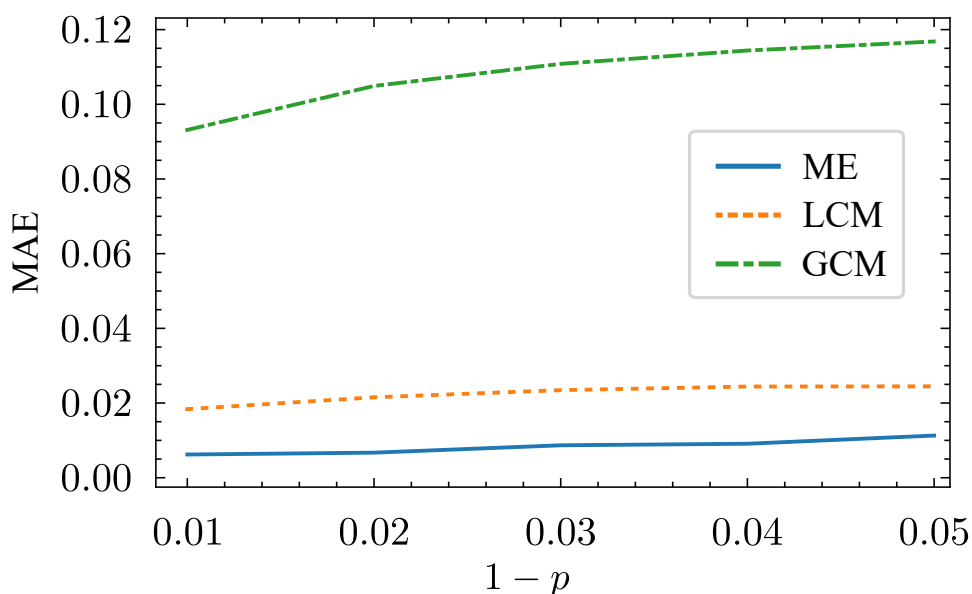


Figura 28 – Erro médio absoluto (MAE) para o coeficiente de Pearson  $C_{12}$  em função da probabilidade de erro no estado inicial fornecido ao regressor  $k$ - $NN$  para o modelo de equação mestra (ME), modelo de colisão local (LCM) e modelo de colisão global (GCM) usando 100 valores esperados iniciais.

a sensibilidade de nosso método a um possível erro no estado inicial na etapa de previsão para os três modelos diferentes.

Em todos os casos investigados nesta subseção, calculamos os valores esperados usando 100 pontos de tempo. Observamos também que aqui nos referimos ao primeiro e ao segundo modelos de colisão como modelo de colisão local (LCM) e modelo de colisão global (GCM) para fins de apresentação. Observamos na Fig. 28 que o modelo de equação mestra (ME) demonstra a menor sensibilidade a erros de inicialização, mantendo um MAE consistentemente baixo, com um aumento de 0.002 (na ausência de erros de inicialização) para valores ainda abaixo de 0.01. Em contraste, LCM e GCM exibem MAEs mais altos, com o MAE do LCM aumentando de 0.009 para aproximadamente 0.02 e o MAE do GCM subindo de 0.04 para a ordem de 0.1. Apesar desses aumentos, as variações gerais no MAE permanecem modestas, o que mostra a robustez de nosso método contra erros de inicialização. Notavelmente, o GCM experimenta um efeito relativamente mais significativo, com MAE agora na ordem de 10% ao se considerar erros de inicialização. Esses erros ainda permanecem relativamente baixos, destacando a eficiência de nossa abordagem.

Além disso, exploramos como a temperatura do ambiente influencia a capacidade do algoritmo de prever o surgimento de (anti-)sincronização em todos os três modelos considerados. Com esse objetivo, avaliamos o desempenho preditivo do regressor  $k$ - $NN$  na presença de temperaturas ambientais diferentes de zero, quando o algoritmo ainda é treinado usando os dados de temperatura zero. Para levar em conta os efeitos da temperatura, assumimos que as matrizes densidade do qubit ambiental nos modelos de colisão local e global são dadas por

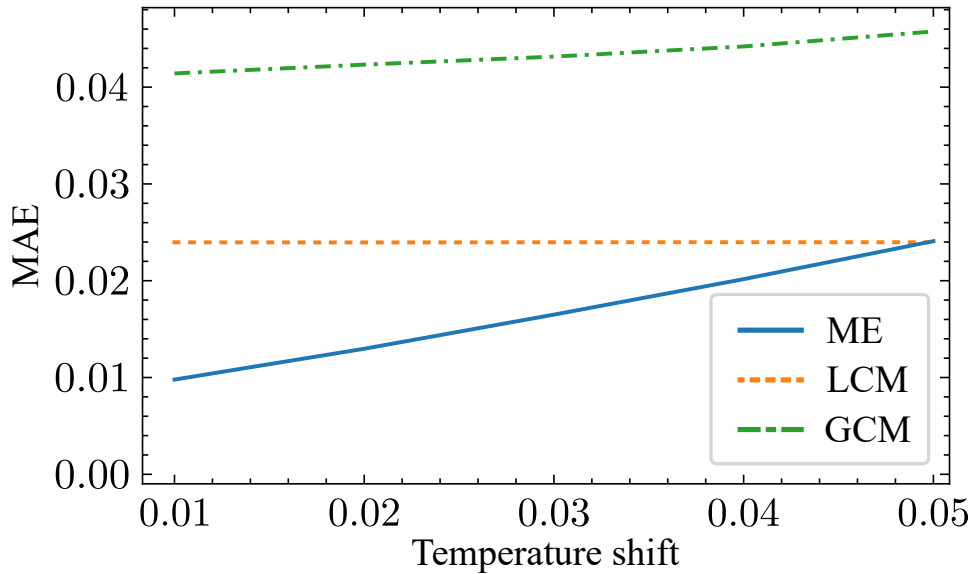


Figura 29 – Erro médio absoluto (MAE) para o coeficiente de Pearson  $C_{12}$  em função da temperatura do ambiente para o modelo de equação mestra (ME), modelo de colisão local (LCM) e modelo de colisão global (GCM), considerando 100 valores esperados iniciais dos qubits.

$\rho_E = (1 - p)|0\rangle\langle 0| + p|1\rangle\langle 1|$ , com  $p$  variando de 0.01 a 0.05. Fisicamente, isso implica que o ambiente está, na verdade, em um estado térmico misto, refletindo os efeitos da temperatura. Em particular, à medida que  $p$  varia de 0.01 a 0.05, uma elevação gradual na temperatura ambiental é simulada, onde a probabilidade do ambiente estar no estado fundamental diminui, enquanto ao mesmo tempo sua probabilidade de estar no estado excitado aumenta. Por outro lado, no caso do modelo de equação mestra, variamos ligeiramente o número médio de fótons, permitindo que  $\bar{n}$  varie de 0.01 a 0.05 nas Eq. (97) e Eq. (98), o que nos permite analisar o efeito da temperatura do ambiente sobre o desempenho do regressor  $k$ - $NN$ .

Na Fig. 29, exibimos o MAE para nossas previsões obtidas utilizando o regressor  $k$ - $NN$ , treinado com o conjunto de dados ideal, em função da temperatura do ambiente para todos os três modelos. Observamos que para o GCM e ME, o MAE aumenta ligeiramente à medida que a temperatura do ambiente sobe. No entanto, no caso do LCM, onde um dos qubits não está diretamente em contato com o ambiente, o desempenho de nossa abordagem é basicamente indiferente às mudanças na temperatura do ambiente.

## 6.4 Conclusão

Utilizamos uma técnica de aprendizado de máquina, baseada no algoritmo de regressão  $k$ - $NN$ , para prever o surgimento de sincronização espontânea induzida pelo ambiente entre os valores esperados dos observáveis para um par de qubits na dinâmica de sistemas quânticos abertos. Consideramos três modelos diferentes de sistemas abertos que abrangem regimes de dissipação local, global e coletiva. Enquanto os dois primeiros modelos são descritos dentro da

estrutura de modelos de colisão, o terceiro é representado por uma equação mestra na forma de Lindblad. Demonstramos que, em todos esses três modelos, nossa abordagem baseada em aprendizado de máquina é capaz de identificar o comportamento de sincronização de longo prazo do sistema aberto com bastante precisão, utilizando apenas os valores esperados iniciais dos observáveis dos qubits como entrada. De fato, mostramos que o algoritmo de aprendizado de máquina não apenas estima com muita precisão o início da sincronização ou antissincronização entre os valores esperados dos qubits, mas também as oscilações com bloqueio de fase entre eles com diferença de fase arbitrária, ou seja, sincronização com atraso temporal. A abordagem proposta tem o potencial de ser útil em experimentos, pois pode reduzir significativamente o número de medições necessárias para determinar o estabelecimento dinâmico da evolução síncrona. Para avaliar a robustez de nossa abordagem contra potenciais imperfeições de medição em um ambiente experimental, adicionamos erros aleatórios aos valores esperados simulados dos qubits e testamos a capacidade de nossa metodologia de reconhecer o comportamento de sincronização nos modelos considerados. Além disso, testamos a robustez de nosso modelo frente a pequenas variações no estado inicial do sistema, bem como à temperatura, mostrando que, apesar de um leve aumento no MAE, as previsões mantiveram-se bastante precisas. Estabelecemos uma relação de compromisso entre o grau de erro introduzido aleatoriamente nos valores esperados e a precisão do algoritmo em prever a sincronização. Em suma, nossos resultados demonstram que, apesar da presença de erros nos valores de entrada do regressor  $k$ - $NN$ , a abordagem proposta ainda funciona com alta precisão, desde que algumas dezenas de valores esperados iniciais dos qubits sejam introduzidos ao algoritmo para previsão. O fato de que o método empregado funciona com precisão em três diferentes cenários de sistemas abertos sugere que ele pode ser aplicado de maneira eficiente em uma ampla gama de modelos físicos. Por fim, motivados pelo desempenho da abordagem de aprendizado de máquina empregada em prever com sucesso o aparecimento de sincronização mútua entre dois qubits, pode-se tentar aplicar métodos similares a problemas mais complexos envolvendo um número maior de spins quânticos principais ou osciladores, onde estados dinâmicos exóticos, como estados de quimera, podem surgir sob condições adequadas (BASTIDAS et al., 2015; SENTHILKUMAR; CHANDRASEKAR, 2019).

## 6.5 Apêndice - Comportamento em tempos iniciais e surgimento da sincronização

Apresentamos a dinâmica explícita dos valores esperados de spin para todos os três modelos que consideramos no manuscrito na Fig. 30. Todas as subfiguras incluem a dinâmica pré-sincronização juntamente com o surgimento da dinâmica (anti-)sincronizada em todos os três modelos. Uma vez que o comportamento síncrono é estabelecido entre os observáveis, ele persiste, seja no estado estacionário ou até que os valores esperados decaiam a zero. A janela de tempo que usamos para treinar o algoritmo de aprendizado de máquina consiste nos dados

após o estabelecimento da sincronização. Em seguida, usando esse algoritmo, tentamos prever a sincronização observando os dados em tempos muito iniciais (até os primeiros 100 pontos), onde a sincronização ainda não está estabelecida, conforme descrito na Sec. 6.2. Portanto, acreditamos que nossa abordagem para o problema é uma maneira promissora de prever o comportamento em tempos tardios de uma propriedade de um sistema quântico observando a dinâmica em tempos iniciais.

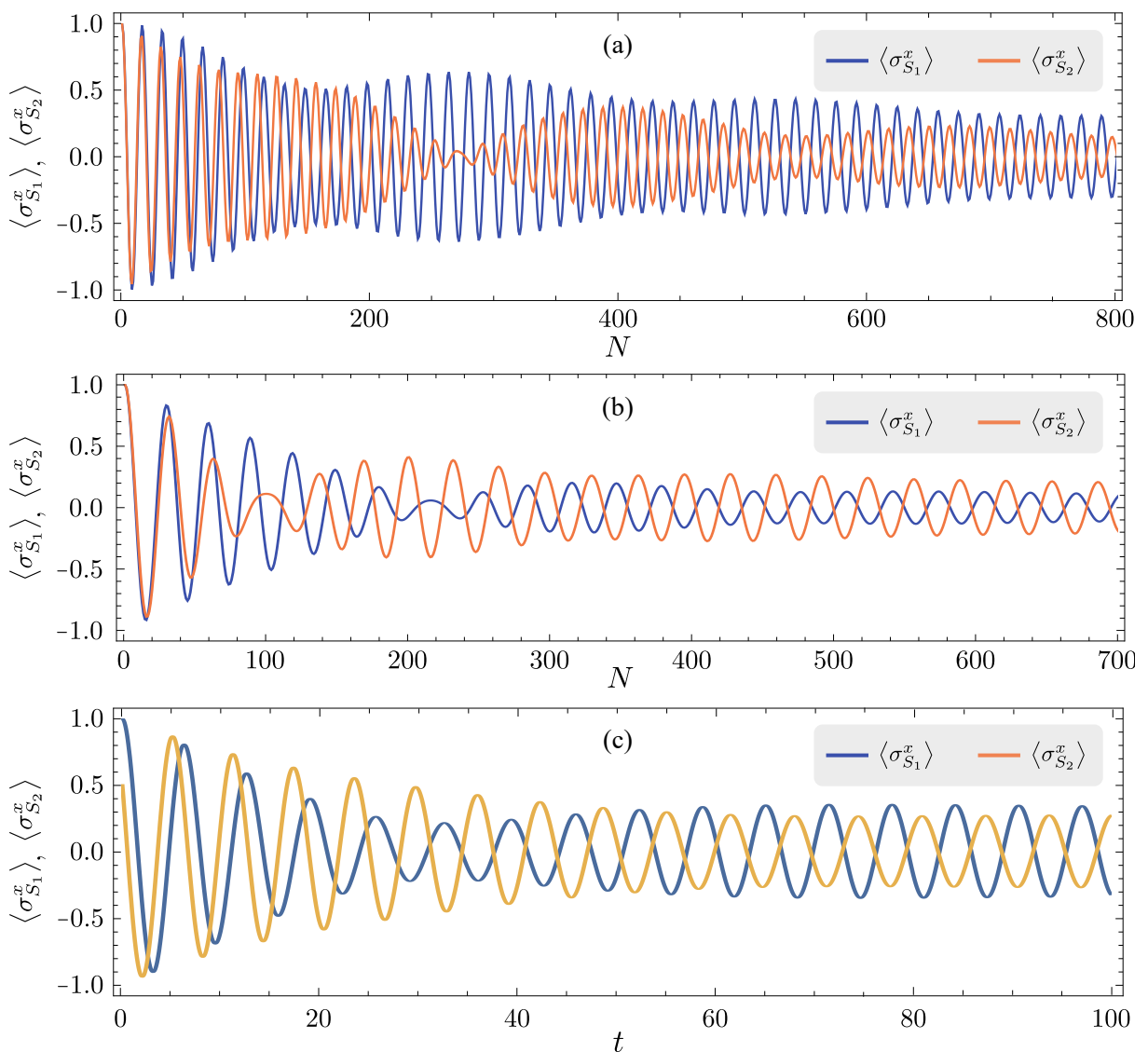


Figura 30 – Dinâmica dos valores esperados locais de spin dos dois qubits para o modelo de colisão local (LCM) (superior), modelo de colisão global (GCM) (meio) e modelo de equação mestra (ME) (inferior), que exibem tanto sua dinâmica pré-sincronização quanto o surgimento da transição dinâmica para o comportamento (anti-)síncrono.

# 7 Conclusão da Tese

Nesta tese, exploramos a interseção entre Informação Quântica e Aprendizado de Máquina para resolver problemas relevantes na área da Informação Quântica. Em um primeiro momento, desenvolvemos uma metodologia inovadora para classificar fases quânticas, independente do modelo subjacente, utilizando correlações de spins e técnicas de Machine Learning. Esse classificador independente de modelo demonstrou eficácia ao prever fases desconhecidas com base em informações de Hamiltonianos conhecidos. Embora tenha representado um avanço notável, reconhecemos a necessidade de aprimorar a resolução e a generalização dessa abordagem, bem como de investigar métodos para identificar com maior precisão os pontos de transição.

Ainda sobre a classificação de fases quânticas, mostramos a possibilidade de fazê-lo mesmo em situações com dados escassos, na qual a expansão do *dataset* utilizando dados sintéticos pode ser útil. Em seguida, utilizando algoritmos de aprendizado de máquina quânticos, em combinação com o SHAP para feature selection, fomos capazes de classificar as fases com alta precisão e apenas 5 features, reduzindo o poder computacional necessário para tal. Apesar de grandes avanços, reconhecemos a necessidade de explorar mais algoritmos e modelos, e no segundo caso de aplicar os algoritmos em hardware quântico real (onde a presença de erros pode impactar as previsões). Também reconhecemos que a implementação de uma versão quântica para o SHAP poderia contribuir com uma avaliação mais igual entre os diferentes modelos apresentados, mas se torna inviável por custo computacional.

Em seguida, aplicamos técnicas de Machine Learning para prever sincronização espontânea em sistemas quânticos abertos. Nosso algoritmo de regressão demonstrou alta precisão na detecção desse fenômeno, permitindo uma redução significativa na quantidade de medições necessárias para compreender a dinâmica da sincronização. Entretanto, para avanços futuros, consideramos a exploração de algoritmos mais avançados, além da incorporação de estratégias de interpretabilidade para uma compreensão mais profunda dos mecanismos subjacentes à sincronização quântica.

Nossos resultados demonstram o potencial do uso de Machine Learning em problemas complexos e desafiadores da Informação Quântica. À medida que avançamos para pesquisas futuras, exploramos a aplicação de algoritmos mais sofisticados e abordagens de inteligência artificial explicável para compreender profundamente os sistemas quânticos estudados. Esta tese oferece uma base sólida na compreensão e utilização de algoritmos de aprendizado de máquina e em informação quântica. Ela ressalta o potencial transformador do Aprendizado de Máquina na resolução de desafios na Ciência de Informação Quântica.

## 7.1 Direcionamentos para Trabalhos Futuros

Em um esforço contínuo para aprimorar nossa pesquisa, planejamos uma série de direcionamentos para trabalhos futuros. Estas atividades visam expandir nosso entendimento e contribuir de forma significativa com o campo.

Primeiramente, a investigação de modelos mais avançados de aprendizado de máquina, tanto clássicos quanto quânticos, pode oferecer melhorias na precisão das previsões e na identificação de padrões em sistemas quânticos complexos. Além disso, a implementação dos algoritmos desenvolvidos em hardware quântico real, como os disponíveis pela IBM Quantum ou Google Quantum AI, permitirá avaliar seu desempenho em cenários práticos.

Outra abordagem é a aplicação das metodologias aqui desenvolvidas em novos problemas, e mesmo a criação/adaptação de algoritmos quânticos para lidar com maior precisão ou eficiência com os problemas aqui descritos. Também seria relevante estudar a robustez e a escalabilidade dos algoritmos em cenários mais desafiadores, como dados ruidosos ou sistemas de maior escala. Outra linha de pesquisa seria o desenvolvimento de métodos para interpretar e explicar os resultados dos algoritmos quânticos, adaptando técnicas como o SHAP ao contexto quântico, para gerar insights mais claros sobre os padrões aprendidos.

À medida que encerramos esta jornada de pesquisa e investigação na interseção da Informação Quântica e Aprendizado de Máquina, humildemente reconhecemos que o caminho do conhecimento é infinito. Cada descoberta e desafio superado apenas revela a vastidão do que ainda não sabemos. A ciência, com sua infindável curiosidade e capacidade de questionar o desconhecido, nos ensinou que nosso trabalho é apenas uma pequena peça em um quebra-cabeça muito maior. A busca por compreender os mistérios da natureza é uma jornada contínua, e é nosso privilégio e responsabilidade como cientistas contribuir para esse esforço coletivo. Reverenciamos a ciência e celebramos o poder da mente humana de desvendar os segredos do cosmos.

# Referências

ABRAMSON, N.; BRAVERMAN, D.; SEBESTYEN, G. Pattern recognition and machine learning. *IEEE Transactions on Information Theory*, v. 9, n. 4, p. 257–261, 1963.

AI'MEUR, E.; BRASSARD, G.; GAMBS, S. *Machine learning in a quantum world*. In: *Advances in Artificial Intelligence*. pp. 431–442. [S.l.]: Springer, 2006.

ALCALDE, A. *Creating Trees, Dependency Graphs and SVMs in TikZ*. 2017. <<https://elbauldelprogramador.com/en/creating-trees-dependency-graphs-svms-in-tikz/>>. Accessed: 2025-04-16. Disponível em: <<https://elbauldelprogramador.com/en/creating-trees-dependency-graphs-svms-in-tikz/>>.

ALI, S.; ABUHMED, T.; EL-SAPPAGH, S.; MUHAMMAD, K.; ALONSO-MORAL, J. M.; CONFALONIERI, R.; GUIDOTTI, R.; Del Ser, J.; DÍAZ-RODRÍGUEZ, N.; HERRERA, F. Explainable artificial intelligence (xai): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion*, v. 99, p. 101805, 2023. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253523001148>>.

ALPAYDIN, E. *Introduction to Machine Learning*. USA: MIT Press, 2004. ISBN 9780262012119. Disponível em: <<https://doi.org/10.1017/S0269888906220745>>.

ALPAYDIN, E. *Introduction to machine learning*. [S.l.]: MIT press, 2020.

ALTHNIAN, A.; ALSAEED, D.; AL-BAITY, H.; SAMHA, A.; DRIS, A. B.; ALZAKARI, N.; ELWAFI, A. A.; KURDI, H. Impact of dataset size on classification performance: an empirical evaluation in the medical domain. *Applied Sciences*, MDPI, v. 11, n. 2, p. 796, 2021.

ALZUBAIDI, L.; ZHANG, J.; HUMAIDI, A. J.; AL-DUJAILI, A.; DUAN, Y.; AL-SHAMMA, O.; SANTAMARÍA, J.; FADHEL, M. A.; AL-AMIDIE, M.; FARHAN, L. Review of deep learning: Concepts, cnn architectures, challenges, applications, future directions. *Journal of big Data*, Springer, v. 8, p. 1–74, 2021.

ARENAS, A.; DÍAZ-GUILERA, A.; KURTHS, J.; MORENO, Y.; ZHOU, C. Synchronization in complex networks. *Phys. Rep.*, v. 469, n. 3, p. 93 – 153, 2008. ISSN 0370-1573. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0370157308003384>>.

ARNDT, M.; NAIRZ, O.; VOS-ANDREAE, J.; KELLER, C.; ZOUW, G. Van der; ZEILINGER, A. Wave–particle duality of c60 molecules. *nature*, Nature Publishing Group UK London, v. 401, n. 6754, p. 680–682, 1999.

ARRUDA, L. G. E. *Computação quântica baseada em medidas projetivas em sistemas quânticos abertos*. Tese (Doutorado) — Universidade de São Paulo, 2011.

ASPURU-GUZIŁ, A.; WALTHER, P. Photonic quantum simulators. *Nature physics*, Nature Publishing Group, v. 8, n. 4, p. 285–291, 2012.

AWOYEMI, J. O.; ADETUNMBI, A. O.; OLUWADARE, S. A. Credit card fraud detection using machine learning techniques: A comparative analysis. In: IEEE. *2017 international conference on computing networking and informatics (ICCNi)*. [S.l.], 2017. p. 1–9.

BACH, S.; BINDER, A.; MONTAVON, G.; KLAUSCHEN, F.; MÜLLER, K.-R.; SAMEK, W. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PLoS ONE*, Public Library of Science, v. 10, n. 7, p. e0130140, 2015.

Barredo Arrieta, A.; DÍAZ-RODRÍGUEZ, N.; Del Ser, J.; BENNETOT, A.; TABIK, S.; BARBADO, A.; GARCIA, S.; GIL-LOPEZ, S.; MOLINA, D.; BENJAMINS, R.; CHATILA, R.; HERRERA, F. Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. *Information Fusion*, v. 58, p. 82–115, 2020. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253519308103>>.

BASSI, H.; YIM, R. P.; VENDROW, J.; KODULUKA, R.; ZHU, C.; LYU, H. Learning to predict synchronization of coupled oscillators on randomly generated graphs. *Scientific Reports*, v. 12, n. 1, p. 15056, Sep 2022. ISSN 2045-2322. Disponível em: <<https://doi.org/10.1038/s41598-022-18953-8>>.

BASTIDAS, V. M.; OMELCHENKO, I.; ZAKHAROVA, A.; SCHÖLL, E.; BRANDES, T. Quantum signatures of chimera states. *Phys. Rev. E*, American Physical Society, v. 92, p. 062924, Dec 2015. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.92.062924>>.

BECCARIA, M.; CAMPOSTRINI, M.; FEO, A. Evidence for a floating phase of the transverse annni model at high frustration. *Phys. Rev. B*, American Physical Society, v. 76, p. 094410, Sep 2007. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.76.094410>>.

BELL, A.; SOLANO-KAMAIKO, I.; NOV, O.; STOYANOVICH, J. It's just not that simple: An empirical study of the accuracy-explainability trade-off in machine learning for public policy. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. New York, NY, USA: Association for Computing Machinery, 2022. (FAccT '22), p. 248–266. ISBN 9781450393522. Disponível em: <<https://doi.org/10.1145/3531146.3533090>>.

BELLOMO, B.; GIORGI, G. L.; PALMA, G. M.; ZAMBRINI, R. Quantum synchronization as a local signature of super- and subradiance. *Phys. Rev. A*, American Physical Society, v. 95, p. 043807, Apr 2017. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.95.043807>>.

BENEDETTI, C.; GALVE, F.; MANDARINO, A.; PARIS, M. G. A.; ZAMBRINI, R. Minimal model for spontaneous quantum synchronization. *Phys. Rev. A*, American Physical Society, v. 94, p. 052118, Nov 2016. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.94.052118>>.

BENNETT, C. H.; BRASSARD, G. Quantum cryptography: Public key distribution and coin tossing. *Theoretical computer science*, Elsevier, v. 560, p. 7–11, 2014.

BENNETT, C. H.; BRASSARD, G.; CRÉPEAU, C.; JOZSA, R.; PERES, A.; WOOTTERS, W. K. Teleporting an unknown quantum state via dual classical and einstein-podolsky-rosen channels. *Physical review letters*, APS, v. 70, n. 13, p. 1895, 1993.

BENNETT, C. H.; BRASSARD, G.; EKERT, A. K. Quantum cryptography. *Scientific American*, JSTOR, v. 267, n. 4, p. 50–57, 1992.

BENNETT, C. H.; WIESNER, S. J. Communication via one-and two-particle operators on einstein-podolsky-rosen states. *Physical review letters*, APS, v. 69, n. 20, p. 2881, 1992.

BERRAR, D. et al. *Cross-Validation*. 2019.

- BHATTACHARYA, S.; BANERJEE, S.; PATI, A. K. Evolution of coherence and non-classicality under global environmental interaction. *Quantum Information Processing*, v. 17, n. 9, p. 236, Aug 2018. ISSN 1573-1332. Disponível em: <<https://doi.org/10.1007/s11128-018-1989-4>>.
- BIAMONTE, J.; WITTEK, P.; PANCOTTI, N.; REBENTROST, P.; WIEBE, N.; LLOYD, S. Quantum machine learning. *Nature*, Nature Publishing Group UK London, v. 549, n. 7671, p. 195–202, 2017.
- BICCARI, U.; ZUAZUA, E. A stochastic approach to the synchronization of coupled oscillators. *Frontiers in Energy Research*, v. 8, 2020. ISSN 2296-598X. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fenrg.2020.00115>>.
- BISHOP, C. M.; NASRABADI, N. M. *Pattern recognition and machine learning*. [S.l.]: Springer, 2006. v. 4.
- BOU-HAMAD, I. Bayesian credit ratings: A random forest alternative approach. *Communications in Statistics-Theory and Methods*, Taylor & Francis, v. 46, n. 15, p. 7289–7300, 2017.
- BREIMAN, L. Bagging predictors. *Machine learning*, Springer, v. 24, p. 123–140, 1996.
- BREIMAN, L. Random forests. *Springer*, 2001.
- BREIMAN, L. Random forests. *Machine learning*, Springer, v. 45, p. 5–32, 2001.
- BREUER, H.-P.; PETRUCCIONE, F. *The Theory of Open Quantum Systems*. [S.l.]: Oxford University Press, Oxford, 2002.
- BRO, R.; SMILDE, A. K. Principal component analysis. *Analytical methods*, Royal Society of Chemistry, v. 6, n. 9, p. 2812–2831, 2014.
- BROECKER, P.; CARRASQUILLA, J.; MELKO, R. G.; TREBST, S. Machine learning quantum phases of matter beyond the fermion sign problem. *Scientific Reports*, Springer Science and Business Media LLC, v. 7, n. 1, Aug 2017. ISSN 2045-2322. Disponível em: <<http://dx.doi.org/10.1038/s41598-017-09098-0>>.
- BROWNLEE, J. Classification accuracy is not enough: More performance measures you can use. *Machine Learning Mastery*, v. 21, 2014.
- BROWNLEE, J. Tactics to combat imbalanced classes in your machine learning dataset. *Machine Learning Mastery*, v. 19, 2015.
- BROWNLEE, J. How to backtest machine learning models for time series forecasting. *Machine Learning Mastery*, v. 19, 2016.
- BRUM, J. A. *Part VI: Transições de fase quânticas*. Notas de Aula. Universidade Estadual de Campinas - Unicamp, 2015. Disponível em: <<https://sites.ifi.unicamp.br/brum/files/2014/01/Ementa-FI105-2015.pdf>>. Acesso em: 20 de Janeiro de 2020.
- BUCHANAN, B. G. A (very) brief history of artificial intelligence. *Ai Magazine*, v. 26, n. 4, p. 53–53, 2005.
- BURGES, C. J. A tutorial on support vector machines for pattern recognition. *Data mining and knowledge discovery*, Springer, v. 2, n. 2, p. 121–167, 1998.

BUSCH, P.; LAHTI, P. Observable (compendium entry). 2008.

BUSUIOC, M. Accountable artificial intelligence: Holding algorithms to account. *Public Administration Review*, Wiley Online Library, v. 81, n. 5, p. 825–836, 2021.

BUČA, B.; BOOKER, C.; JAKSCH, D. Algebraic theory of quantum synchronization and limit cycles under dissipation. *SciPost Phys.*, SciPost, v. 12, p. 097, 2022. Disponível em: <<https://scipost.org/10.21468/SciPostPhys.12.3.097>>.

CABOT, A.; GIORGI, G. L.; GALVE, F.; ZAMBRINI, R. Quantum synchronization in dimer atomic lattices. *Phys. Rev. Lett.*, American Physical Society, v. 123, p. 023604, Jul 2019. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.123.023604>>.

CABOT, A.; GIORGI, G. L.; ZAMBRINI, R. Synchronization and coalescence in a dissipative two-qubit system. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 477, n. 2249, p. 20200850, 2021. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2020.0850>>.

CANABARRO, A.; FANCHINI, F. F.; MALVEZZI, A. L.; PEREIRA, R.; CHAVES, R. Unveiling phase transitions with machine learning. *Phys. Rev. B*, American Physical Society, v. 100, p. 045129, Jul 2019. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.100.045129>>.

CAO, Y.; ROMERO, J.; OLSON, J. P.; DEGROOTE, M.; JOHNSON, P. D.; KIEFEROVÁ, M.; KIVLICHAN, I. D.; MENKE, T.; PEROPADRE, B.; SAWAYA, N. P. et al. Quantum chemistry in the age of quantum computing. *Chemical reviews*, ACS Publications, v. 119, n. 19, p. 10856–10915, 2019.

CARLEO, G.; CIRAC, I.; CRANMER, K.; DAUDET, L.; SCHULD, M.; TISHBY, N.; VOGT-MARANTO, L.; ZDEBOROVÁ, L. Machine learning and the physical sciences. *Rev. Mod. Phys.*, American Physical Society, v. 91, p. 045002, Dec 2019. Disponível em: <<https://link.aps.org/doi/10.1103/RevModPhys.91.045002>>.

CARRASQUILLA, J.; MELKO, R. G. Machine learning phases of matter. *Nature Physics*, v. 13, n. 5, p. 431–434, 2017. Disponível em: <<https://doi.org/10.1038/nphys4035>>.

CASALICCHIO, G.; MOLNAR, C.; BISCHL, B. Visualizing the feature importance for black box models. In: SPRINGER. *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2018, Dublin, Ireland, September 10–14, 2018, Proceedings, Part I 18*. [S.l.], 2019. p. 655–670.

CATTANEO, M.; GIORGI, G. L.; MANISCALCO, S.; PARAOANU, G. S.; ZAMBRINI, R. Bath-induced collective phenomena on superconducting qubits: Synchronization, subradiance, and entanglement generation. *Annalen der Physik*, v. 533, n. 5, p. 2100038, 2021. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.202100038>>.

CEREZO, M.; ARRASMITH, A.; BABBUSH, R.; BENJAMIN, S. C.; ENDO, S.; FUJII, K.; MCCLEAN, J. R.; MITARAI, K.; YUAN, X.; CINCIO, L.; COLES, P. J. Variational quantum algorithms. *Nature Reviews Physics*, Springer Science and Business Media LLC, v. 3, n. 9, p. 625–644, aug 2021. Disponível em: <<https://doi.org/10.1038/s42254-021-00348-9>>.

CHAKRABARTI, B. K.; DUTTA, A.; SEN, P. *Quantum Ising phases and transitions in transverse Ising models*. [S.l.]: Springer Science & Business Media, 2008. v. 41.

CHEN, W.; HIDA, K.; SANCTUARY, B. C. Ground-state phase diagram of  $s = 1$  XXZ chains with uniaxial single-ion-type anisotropy. *Phys. Rev. B*, American Physical Society, v. 67, p. 104401, Mar 2003. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.67.104401>>.

CHEN, X.; GU, Z.-C.; WEN, X.-G. Complete classification of one-dimensional gapped quantum phases in interacting spin systems. *Physical Review B*, American Physical Society (APS), v. 84, n. 23, Dec 2011. ISSN 1550-235X. Disponível em: <<http://dx.doi.org/10.1103/PhysRevB.84.235128>>.

CHICCO, D.; WARRENS, M. J.; JURMAN, G. The coefficient of determination r-squared is more informative than smape, mae, mape, mse and rmse in regression analysis evaluation. *PeerJ Computer Science*, PeerJ Inc., v. 7, p. e623, 2021.

CHOWDHURY, S. N.; RAY, A.; MISHRA, A.; GHOSH, D. Extreme events in globally coupled chaotic maps. *Journal of Physics: Complexity*, IOP Publishing, v. 2, n. 3, p. 035021, sep 2021. Disponível em: <<https://dx.doi.org/10.1088/2632-072X/ac221f>>.

CICCARELLO, F. Collision models in quantum optics. *Quantum Measurements and Quantum Metrology*, v. 4, n. 1, p. 53–63, 2017. Disponível em: <<https://doi.org/10.1515/qmetro-2017-0007>>.

CICCARELLO, F.; LORENZO, S.; GIOVANNETTI, V.; PALMA, G. M. Quantum collision models: Open system dynamics from repeated interactions. *Physics Reports*, v. 954, p. 1–70, 2022. ISSN 0370-1573. Quantum collision models: Open system dynamics from repeated interactions. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0370157322000035>>.

CLAESSEN, M.; MOOR, B. D. *Hyperparameter Search in Machine Learning*. 2015.

CONTINENTINO, M. A. *Transições de fases quânticas*. Centro Brasileiro de Pesquisas Físicas. 25 slides., 2006. Disponível em: <<http://www2.im.ufrj.br/~coloquiomea/apresentacoes/continentino.pdf>>. Acesso em: 20 de Janeiro de 2020.

COOLEN, A. C. C.; NIKOLETOPOULOS, T.; ARAI, S.; TANAKA, K. Dynamical analysis of quantum annealing. In: \_\_\_\_\_. *Sublinear Computation Paradigm: Algorithmic Revolution in the Big Data Era*. Singapore: Springer Singapore, 2022. p. 295–317. ISBN 978-981-16-4095-7. Disponível em: <[https://doi.org/10.1007/978-981-16-4095-7\\_12](https://doi.org/10.1007/978-981-16-4095-7_12)>.

CORTES, C.; VAPNIK, V. Support-vector networks. *Machine learning*, Springer, v. 20, p. 273–297, 1995.

CUNNINGHAM, P.; CORD, M.; DELANY, S. J. Supervised learning. In: \_\_\_\_\_. *Machine Learning Techniques for Multimedia: Case Studies on Organization and Retrieval*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. p. 21–49. ISBN 978-3-540-75171-7. Disponível em: <[https://doi.org/10.1007/978-3-540-75171-7\\_2](https://doi.org/10.1007/978-3-540-75171-7_2)>.

CÁRDENAS-LÓPEZ, F. A.; SANZ, M.; RETAMAL, J. C.; SOLANO, E. Enhanced quantum synchronization via quantum machine learning. *Advanced Quantum Technologies*, v. 2, n. 7-8, p. 1800076, 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.201800076>>.

DAMANET, F.; MARTIN, J. Competition between finite-size effects and dipole–dipole interactions in few-atom systems. *Journal of Physics B: Atomic, Molecular and Optical Physics*, IOP Publishing, v. 49, n. 22, p. 225501, oct 2016. Disponível em: <<https://doi.org/10.1088%2F0953-4075%2F49%2F22%2F225501>>.

- DELUA, J. *Supervised vs. Unsupervised Learning: What's the Difference?* IBM Analytics, 2021. Disponível em: <<https://www.ibm.com/cloud/blog/supervised-vs-unsupervised-learning>>.
- DÍAZ-URIARTE, R.; ANDRÉS, S. Alvarez de. Gene selection and classification of microarray data using random forest. *BMC bioinformatics*, Springer, v. 7, p. 1–13, 2006.
- DIRAC, P. A. M. A new notation for quantum mechanics. In: CAMBRIDGE UNIVERSITY PRESS. *Mathematical Proceedings of the Cambridge Philosophical Society*. [S.l.], 1939. v. 35, n. 3, p. 416–418.
- DONG, X.-Y.; POLLMANN, F.; ZHANG, X.-F. Machine learning of quantum phase transitions. *Physical Review B*, American Physical Society (APS), v. 99, n. 12, Mar 2019. ISSN 2469-9969. Disponível em: <<http://dx.doi.org/10.1103/PhysRevB.99.121104>>.
- DUDANI, S. A. The distance-weighted k-nearest-neighbor rule. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6, n. 4, p. 325–327, 1976.
- EISBERG, R.; RESNICK, R. *Quantum physics of atoms, molecules, solids, nuclei, and particles*. [S.l.: s.n.], 1985.
- EKERT, A. K. Quantum cryptography based on bell's theorem. *Physical review letters*, APS, v. 67, n. 6, p. 661, 1991.
- ES'HAQI-SANI, N.; MANZANO, G.; ZAMBRINI, R.; FAZIO, R. Synchronization along quantum trajectories. *Phys. Rev. Res.*, American Physical Society, v. 2, p. 023101, Apr 2020. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevResearch.2.023101>>.
- ESTARELLAS, G. G.; GIORGI, G. L.; SORIANO, M. C.; ZAMBRINI, R. Machine learning applied to quantum synchronization-assisted probing. *Advanced Quantum Technologies*, v. 2, n. 7-8, p. 1800085, 2019. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/qute.201800085>>.
- EZHOV, A. A.; VENTURA, D. Quantum neural networks. *Future Directions for Intelligent Systems and Information Sciences: The Future of Speech and Image Technologies, Brain Computers, WWW, and Bioinformatics*, Springer, p. 213–235, 2000.
- FAN, H.; KONG, L.-W.; LAI, Y.-C.; WANG, X. Anticipating synchronization with machine learning. *Phys. Rev. Res.*, American Physical Society, v. 3, p. 023237, Jun 2021. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevResearch.3.023237>>.
- FAN, H.; WANG, L.; DU, Y.; WANG, Y.; XIAO, J.; WANG, X. Learning the dynamics of coupled oscillators from transients. *Phys. Rev. Res.*, American Physical Society, v. 4, p. 013137, Feb 2022. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevResearch.4.013137>>.
- FEYNMAN, R. P. Simulating physics with computers. In: *Feynman and computation*. [S.l.]: cRc Press, 2018. p. 133–153.
- FINN, C.; ABBEEL, P.; LEVINE, S. Model-agnostic meta-learning for fast adaptation of deep networks. *arXiv preprint arXiv:1703.03400*, 2017.
- FRADKOV, A. L. Early history of machine learning. *IFAC-PapersOnLine*, v. 53, n. 2, p. 1385–1390, 2020. ISSN 2405-8963. 21st IFAC World Congress. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S2405896320325027>>.

FUKUSHIMA, K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, Springer, v. 36, n. 4, p. 193–202, 1980.

GALVE, F.; GIORGI, G. L.; ZAMBRINI, R. Quantum correlations and synchronization measures. In: \_\_\_\_\_. *Lectures on General Quantum Correlations and their Applications*. Cham: Springer International Publishing, 2017. p. 393–420. ISBN 978-3-319-53412-1. Disponível em: <[https://doi.org/10.1007/978-3-319-53412-1\\_18](https://doi.org/10.1007/978-3-319-53412-1_18)>.

GANDHI, R. Support vector machine—introduction to machine learning algorithms. *Towards Data Science*, v. 7, n. 06, 2018.

GARCÍA, S.; RAMÍREZ-GALLEGO, S.; LUENGO, J.; BENÍTEZ, J. M.; HERRERA, F. Big data preprocessing: methods and prospects. *Big Data Analytics*, BioMed Central, v. 1, n. 1, p. 1–22, 2016.

GÉRON, A. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*. [S.l.]: "O'Reilly Media, Inc.", 2022.

GINZBURG, V. L.; LANDAU, L. D. On the theory of superconductivity. In: \_\_\_\_\_. *On Superconductivity and Superfluidity: A Scientific Autobiography*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. p. 113–137. ISBN 978-3-540-68008-6. Disponível em: <[https://doi.org/10.1007/978-3-540-68008-6\\_4](https://doi.org/10.1007/978-3-540-68008-6_4)>.

GIORDANO, C.; BRENNAN, M.; MOHAMED, B.; RASHIDI, P.; MODAVE, F.; TIGHE, P. Accessing artificial intelligence for clinical decision-making. *Frontiers in digital health*, Frontiers Media SA, v. 3, p. 645232, 2021.

GIORGI, G. L.; CABOT, A.; ZAMBRINI, R. Transient synchronization in open quantum systems. In: VACCHINI, B.; BREUER, H.-P.; BASSI, A. (Ed.). *Advances in Open Systems and Fundamental Tests of Quantum Mechanics*. Cham: Springer International Publishing, 2019. p. 73–89. ISBN 978-3-030-31146-9.

GIORGI, G. L.; GALVE, F.; MANZANO, G.; COLET, P.; ZAMBRINI, R. Quantum correlations and mutual synchronization. *Phys. Rev. A*, American Physical Society, v. 85, p. 052101, May 2012. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.85.052101>>.

GIORGI, G. L.; GALVE, F.; ZAMBRINI, R. Probing the spectral density of a dissipative qubit via quantum synchronization. *Phys. Rev. A*, American Physical Society, v. 94, p. 052121, Nov 2016. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.94.052121>>.

GIORGI, G. L.; PLASTINA, F.; FRANCICA, G.; ZAMBRINI, R. Spontaneous synchronization and quantum correlation dynamics of open spin systems. *Phys. Rev. A*, American Physical Society, v. 88, p. 042115, Oct 2013. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.88.042115>>.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, Massachusetts: The MIT Press, 2016.

GOODFELLOW, I. J.; POUGET-ABADIE, J.; MIRZA, M.; XU, B.; WARDE-FARLEY, D.; OZAIR, S.; COURVILLE, A.; BENGIO, Y. *Generative Adversarial Networks*. 2014. Disponível em: <<https://arxiv.org/abs/1406.2661>>.

GORI, M.; SCARSELLI, F. Are multilayer perceptrons adequate for pattern recognition and verification? *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE, v. 20, n. 11, p. 1121–1132, 1998.

GOYCHUK, I.; CASADO-PASCUAL, J.; MORILLO, M.; LEHMANN, J.; HÄNGGI, P. Quantum stochastic synchronization. *Phys. Rev. Lett.*, American Physical Society, v. 97, p. 210601, Nov 2006. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.97.210601>>.

GROVER, L. K. A fast quantum mechanical algorithm for database search. In: *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing*. [S.l.: s.n.], 1996. p. 212–219.

GUTH, S.; SAPSIS, T. P. Machine learning predictors of extreme events occurring in complex dynamical systems. *Entropy*, v. 21, n. 10, 2019. ISSN 1099-4300. Disponível em: <<https://www.mdpi.com/1099-4300/21/10/925>>.

HADFIELD, S.; WANG, Z.; O'GORMAN, B.; RIEFFEL, E. G.; VENTURELLI, D.; BISWAS, R. From the quantum approximate optimization algorithm to a quantum alternating operator ansatz. *Algorithms*, MDPI AG, v. 12, n. 2, p. 34, fev. 2019. ISSN 1999-4893. Disponível em: <<http://dx.doi.org/10.3390/a12020034>>.

HALLBERG, K. A. New trends in density matrix renormalization. *Advances in Physics*, Taylor & Francis, v. 55, n. 5-6, p. 477–526, 2006.

HART, P. E.; STORK, D. G.; DUDA, R. O. *Pattern classification*. [S.l.]: Wiley Hoboken, 2000.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. Overview of supervised learning. In: \_\_\_\_\_. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York, NY: Springer New York, 2009. p. 9–41. ISBN 978-0-387-84858-7. Disponível em: <[https://doi.org/10.1007/978-0-387-84858-7\\_2](https://doi.org/10.1007/978-0-387-84858-7_2)>.

HASTIE, T.; TIBSHIRANI, R.; FRIEDMAN, J. H.; FRIEDMAN, J. H. *The elements of statistical learning: data mining, inference, and prediction*. [S.l.]: Springer, 2009. v. 2.

HASTINGS, M. B. Classifying quantum phases with the kirby torus trick. *Physical Review B*, American Physical Society (APS), v. 88, n. 16, Oct 2013. ISSN 1550-235X. Disponível em: <<http://dx.doi.org/10.1103/PhysRevB.88.165114>>.

HAVLÍČEK, V.; CÓRCOLES, A. D.; TEMME, K.; HARROW, A. W.; KANDALA, A.; CHOW, J. M.; GAMBETTA, J. M. Supervised learning with quantum-enhanced feature spaces. *Nature*, 2019.

HAVLÍČEK, V.; CÓRCOLES, A. D.; TEMME, K.; HARROW, A. W.; KANDALA, A.; CHOW, J. M.; GAMBETTA, J. M. Supervised learning with quantum-enhanced feature spaces. *Nature*, Nature Publishing Group, v. 567, n. 7747, p. 209–212, 2019.

HEARST, M.; DUMAIS, S.; OSUNA, E.; PLATT, J.; SCHOLKOPF, B. Support vector machines. *IEEE Intelligent Systems and their Applications*, v. 13, n. 4, p. 18–28, 1998.

HEINRICH, G.; LUDWIG, M.; QIAN, J.; KUBALA, B.; MARQUARDT, F. Collective dynamics in optomechanical arrays. *Phys. Rev. Lett.*, American Physical Society, v. 107, p. 043603, Jul 2011. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.107.043603>>.

HICKEY, J. M.; STEFANO, P. G. D.; VASILEIOU, V. Fairness by explicability and adversarial shap learning. In: HUTTER, F.; KERSTING, K.; LIJFFIJT, J.; VALERA, I. (Ed.). *Machine Learning and Knowledge Discovery in Databases*. Cham: Springer International Publishing, 2021. p. 174–190. ISBN 978-3-030-67664-3.

HIDARY, J. D. A brief history of quantum computing. In: \_\_\_\_\_. *Quantum Computing: An Applied Approach*. Cham: Springer International Publishing, 2021. p. 15–21. ISBN 978-3-030-83274-2. Disponível em: <[https://doi.org/10.1007/978-3-030-83274-2\\_2](https://doi.org/10.1007/978-3-030-83274-2_2)>.

HOFFMAN, R. R.; MUELLER, S. T.; KLEIN, G.; LITMAN, J. *Metrics for Explainable AI: Challenges and Prospects*. 2019. Disponível em: <<https://arxiv.org/abs/1812.04608>>.

HOSSAIN, M. M.; ALI, M. S.; SWARNA, R. A.; HASAN, M. M.; HABIB, N.; RAHMAN, M. W.; AZAD, M. M.; RAHMAN, M. M. Analyzing the effect of feature mapping techniques along with the circuit depth in quantum supervised learning by utilizing quantum support vector machine. In: *2021 24th International Conference on Computer and Information Technology (ICCIT)*. [S.l.: s.n.], 2021. p. 1–5.

HUANG, H.-Y.; BROUGHTON, M.; MOHSENI, M.; BABBUSH, R.; BOIXO, S.; NEVEN, H.; MCCLEAN, J. R. Power of data in quantum machine learning. *Nat. Commun.*, Nature, v. 12, n. 12, p. 1–9, Feb 2021. Disponível em: <<http://dx.doi.org/10.1126/science.aag2302>>.

HUANG, H.-Y.; KUENG, R.; TORLAI, G.; ALBERT, V. V.; PRESKILL, J. *Provably efficient machine learning for quantum many-body problems*. 2021.

HUSH, M. R.; LI, W.; GENWAY, S.; LESANOVSKY, I.; ARMOUR, A. D. Spin correlations as a probe of quantum synchronization in trapped-ion phonon lasers. *Phys. Rev. A*, American Physical Society, v. 91, p. 061401, Jun 2015. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.91.061401>>.

IBM Quantum. *Computer resources*. 2024. <<https://quantum.ibm.com>>. Accessed: 2024-02-12.

IMPENS, F.; GUÉRY-ODELIN, D. Shortcut to synchronization in classical and quantum systems. *Scientific Reports*, v. 13, n. 1, p. 453, Jan 2023. Disponível em: <<https://doi.org/10.1038/s41598-022-27130-w>>.

ITABASHI, K.; TRAN, Q. H.; HASEGAWA, Y. Evaluating the phase dynamics of coupled oscillators via time-variant topological features. *Phys. Rev. E*, American Physical Society, v. 103, p. 032207, Mar 2021. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.103.032207>>.

JAVADI-ABHARI, A.; TREINISH, M.; KRSULICH, K.; WOOD, C. J.; LISHMAN, J.; GACON, J.; MARTIEL, S.; NATION, P. D.; BISHOP, L. S.; CROSS, A. W.; JOHNSON, B. R.; GAMBETTA, J. M. *Quantum computing with Qiskit*. 2024.

KALITA, S.; CHAKRABORTY, S.; SARMA, A. K. Switching of quantum synchronization in coupled optomechanical oscillators. *Journal of Physics Communications*, IOP Publishing, v. 5, n. 11, p. 115006, nov 2021. Disponível em: <<https://dx.doi.org/10.1088/2399-6528/ac3204>>.

KANDALA, A. et al. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, Springer, v. 549, p. 242–246, 2017.

KARPAT, G.; YALÇINKAYA, I.; ÇAKMAK, B. Quantum synchronization in a collision model. *Phys. Rev. A*, American Physical Society, v. 100, p. 012133, Jul 2019. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.100.012133>>.

KARPAT, G.; YALÇINKAYA, I.; ÇAKMAK, B. Quantum synchronization of few-body systems under collective dissipation. *Phys. Rev. A*, American Physical Society, v. 101, p. 042121, Apr 2020. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.101.042121>>.

KARPAT, G.; YALÇINKAYA, I.; ÇAKMAK, B.; GIORGI, G. L.; ZAMBRINI, R. Synchronization and non-markovianity in open quantum systems. *Phys. Rev. A*, American Physical Society, v. 103, p. 062217, Jun 2021. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.103.062217>>.

KHALIL, T.; RICHERT, J. Evolution of open quantum systems: Time scales, stochastic and continuous processes. *Quanta*, Quanta, v. 10, n. 1, p. 42–54, nov 2021. Disponível em: <<https://doi.org/10.12743%2Fquanta.v10i1.157>>.

KINGMA, D. P.; BA, J. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

KINGSFORD, C.; SALZBERG, S. L. What are decision trees? *Nature biotechnology*, Nature Publishing Group US New York, v. 26, n. 9, p. 1011–1013, 2008.

KITAZAWA, A.; NOMURA, K.; OKAMOTO, K. Phase diagram of  $S = 1$  bond-alternating xxz chains. *Phys. Rev. Lett.*, American Physical Society, v. 76, p. 4038–4041, May 1996. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.76.4038>>.

KOEHRSEN, W. Overfitting vs. underfitting: A complete example. *Towards Data Science*, v. 405, 2018.

KORZEKWA, K.; CZACHÓRSKI, S.; PUCHAŁA, Z.; ŻYCZKOWSKI, K. Coherifying quantum channels. *New Journal of Physics*, IOP Publishing, v. 20, n. 4, p. 043028, 2018.

KRITHIKA, V. R.; SOLANKI, P.; VINJANAMPATHY, S.; MAHESH, T. S. Observation of quantum phase synchronization in a nuclear-spin system. *Phys. Rev. A*, American Physical Society, v. 105, p. 062206, Jun 2022. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.105.062206>>.

LANDI, G. T. Quantum information and quantum noise. *Lecture notes*, July, 2018.

LÄUCHLI, A.; SCHMID, G.; TREBST, S. Spin nematics correlations in bilinear-biquadratic  $s=1$  spin chains. *Physical Review B*, American Physical Society (APS), v. 74, n. 14, Oct 2006. ISSN 1550-235X. Disponível em: <<http://dx.doi.org/10.1103/PhysRevB.74.144426>>.

LEE, N. T.; RESNICK, P.; BARTON, G. Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms. *Brookings Institute: Washington, DC, USA*, v. 2, 2019.

LEE, T. E.; CHAN, C.-K.; WANG, S. Entanglement tongue and quantum synchronization of disordered oscillators. *Phys. Rev. E*, American Physical Society, v. 89, p. 022913, Feb 2014. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.89.022913>>.

LEE, T. E.; SADEGHPOUR, H. R. Quantum synchronization of quantum van der pol oscillators with trapped ions. *Phys. Rev. Lett.*, American Physical Society, v. 111, p. 234101, Dec 2013. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.111.234101>>.

LEHMBERG, R. H. Radiation from an  $n$ -atom system. i. general formalism. *Phys. Rev. A*, American Physical Society, v. 2, p. 883–888, Sep 1970. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.2.883>>.

LI, W. Analyzing quantum synchronization through bohmian trajectories. *Phys. Rev. A*, American Physical Society, v. 106, p. 023512, Aug 2022. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.106.023512>>.

LI, W.; ZHANG, W.; LI, C.; SONG, H. Properties and relative measure for quantifying quantum synchronization. *Phys. Rev. E*, American Physical Society, v. 96, p. 012211, Jul 2017. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.96.012211>>.

LI, X.; LI, Y.; JIN, J. Synchronization of persistent oscillations in spin systems with nonlocal dissipation. *Phys. Rev. A*, American Physical Society, v. 107, p. 032219, Mar 2023. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.107.032219>>.

LI, Y.; LI, X.; JIN, J. Quantum nonstationary phenomena of spin systems in collision models. *Phys. Rev. A*, American Physical Society, v. 107, p. 042205, Apr 2023. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.107.042205>>.

LLOYD, S.; MOHSENI, M.; REBENTROST, P. Quantum algorithms for supervised and unsupervised machine learning. *arXiv preprint arXiv:1307.0411*, 2013.

LUDWIG, M.; MARQUARDT, F. Quantum many-body dynamics in optomechanical arrays. *Phys. Rev. Lett.*, American Physical Society, v. 111, p. 073603, Aug 2013. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.111.073603>>.

LUNDBERG, S. M.; ERION, G.; CHEN, H.; DEGRAVE, A.; PRUTKIN, J. M.; NAIR, B.; KATZ, R.; HIMMELFARB, J.; BANSAL, N.; LEE, S.-I. From local explanations to global understanding with explainable ai for trees. *Nature machine intelligence*, Nature Publishing Group, v. 2, n. 1, p. 56–67, 2020.

LUNDBERG, S. M.; ERION, G. G.; LEE, S.-I. Consistent individualized feature attribution for tree ensembles. *arXiv preprint arXiv:1802.03888*, 2018.

LUNDBERG, S. M.; LEE, S.-I. A unified approach to interpreting model predictions. In: GUYON, I.; LUXBURG, U. V.; BENGIO, S.; WALLACH, H.; FERGUS, R.; VISHWANATHAN, S.; GARNETT, R. (Ed.). *Advances in Neural Information Processing Systems*. Curran Associates, Inc., 2017. v. 30. Disponível em: <[https://proceedings.neurips.cc/paper\\_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf)>.

MADSEN, A.; REDDY, S.; CHANDAR, S. Post-hoc interpretability for neural nlp: A survey. *ACM Computing Surveys*, ACM New York, NY, v. 55, n. 8, p. 1–42, 2022.

MAHLOW, F.; LUIZ, F. S.; MALVEZZI, A. L.; FANCHINI, F. F. Model-independent quantum phases classifier. *Scientific Reports*, v. 13, p. 14411, 2023. Disponível em: <<https://doi.org/10.1038/s41598-023-33301-0>>.

MAHLOW, F.; LUIZ, F. S.; MALVEZZI, A. L.; FANCHINI, F. F. Model-independent quantum phases classifier. *Scientific Reports*, Springer Science and Business Media LLC, v. 13, n. 1, set. 2023. ISSN 2045-2322. Disponível em: <<http://dx.doi.org/10.1038/s41598-023-33301-0>>.

MAHLOW, F.; ÇAKMAK, B.; KARPAT, G.; YALÇINKAYA İskender; FANCHINI, F. *Predicting the Onset of Quantum Synchronization Using Machine Learning*. 2023.

MAIMON, O.; ROKACH, L. Introduction to supervised methods. In: \_\_\_\_\_. *Data Mining and Knowledge Discovery Handbook*. Boston, MA: Springer US, 2005. p. 149–164. ISBN 978-0-387-25465-4. Disponível em: <[https://doi.org/10.1007/0-387-25465-X\\_8](https://doi.org/10.1007/0-387-25465-X_8)>.

MALVEZZI, A. L.; KARPAT ÇAKMAK, B.; FANCHINI, F. F.; DEBARBA, T.; VIANNA, R. O. Quantum correlations and coherence in spin-1 heisenberg chains. *Phys. Rev. B*, American Physical Society, v. 93, p. 184428, May 2016. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.93.184428>>.

MANZANO, G.; GALVE, F.; GIORGI, G.; HERNÁNDEZ-GARCÍA, E.; ZAMBRINI, R. Synchronization, quantum correlations and entanglement in oscillator networks. *Sci. Rep.*, v. 3, p. 1439, 03 2013.

MANZANO, G.; GALVE, F.; ZAMBRINI, R. Avoiding dissipation in a system of three quantum harmonic oscillators. *Phys. Rev. A*, American Physical Society, v. 87, p. 032114, Mar 2013. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.87.032114>>.

MARI, A.; FARACE, A.; DIDIER, N.; GIOVANNETTI, V.; FAZIO, R. Measures of quantum synchronization in continuous variable systems. *Phys. Rev. Lett.*, American Physical Society, v. 111, p. 103605, Sep 2013. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.111.103605>>.

MERRICK, L.; TALY, A. The explanation game: Explaining machine learning models using shapley values. In: HOLZINGER, A.; KIESEBERG, P.; TJOA, A. M.; WEIPPL, E. (Ed.). *Machine Learning and Knowledge Extraction*. Cham: Springer International Publishing, 2020. p. 17–38. ISBN 978-3-030-57321-8.

MINSKY, M.; PAPERT, S. An introduction to computational geometry. *Cambridge tiass., HIT*, v. 479, n. 480, p. 104, 1969.

MISHEVA, B. H.; OSTERRIEDER, J.; HIRSA, A.; KULKARNI, O.; LIN, S. F. Explainable ai in credit risk management. *arXiv preprint arXiv:2103.00949*, 2021.

MOHAMMED, M.; KHAN, M.; BASHIER, E. *Machine Learning: Algorithms and Applications*. [S.l.: s.n.], 2016. ISBN 9781498705387.

MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. *Introduction to linear regression analysis*. [S.l.]: John Wiley & Sons, 2021.

MOORE, G. E. Cramming more components onto integrated circuits. *Electronics*, McGraw-Hill, v. 38, n. 8, p. 114–117, 1965.

MÜLLER, A. C.; GUIDO, S. *Introduction to machine learning with Python: a guide for data scientists*. [S.l.]: "O'Reilly Media, Inc.", 2016.

MÜLLER, A. C.; GUIDO, S. *Introduction to machine learning with Python*. 1st. ed. [S.l.]: O'Reilly Media, Inc, 2017.

NEVEN, H.; DENCHEV, V. S.; ROSE, G.; MACREADY, W. G. *Training a Large Scale Classifier with the Quantum Adiabatic Algorithm*. 2009.

NIELSEN, M. A.; CHUANG, I. L. *Quantum Computation and Quantum Information*. [S.l.]: Cambridge University Press, 2000.

NOVIKOFF, A. B. On convergence proofs on perceptrons. In: NEW YORK, NY. *Proceedings of the Symposium on the Mathematical Theory of Automata*. [S.l.], 1962. v. 12, n. 1, p. 615–622.

OLIVEIRA, L. G. S. de. Estudo da relação entre emaranhamento e transições de fases quânticas em sistemas de spins interagentes. Monografia apresentada ao Instituto de Física da Universidade Federal Fluminense - UFF, Niterói - RJ, 2014.

ORTH, P. P.; ROOSEN, D.; HOFSTETTER, W.; HUR, K. L. Dynamics, synchronization, and quantum phase transitions of two dissipative spins. *Phys. Rev. B*, American Physical Society, v. 82, p. 144423, Oct 2010. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.82.144423>>.

OSIPOV, G.; KURTHS, J.; ZHOU, C. *Synchronization in Oscillatory Networks*. Springer Berlin Heidelberg, 2007. (Springer Series in Synergetics). ISBN 9783540712695. Disponível em: <<https://books.google.com.tr/books?id=QHHbWbiB2hIC>>.

PAL, M. Random forest classifier for remote sensing classification. *International journal of remote sensing*, Taylor & Francis, v. 26, n. 1, p. 217–222, 2005.

PAULETTI, T.; SANINO, M.; GIMENES, L.; CARVALHO, I.; FRANÇA, V. Quantum phase transitions in one-dimensional nanostructures: a comparison between dft and dmrg methodologies. *Journal of Molecular Modeling*, Springer, v. 30, n. 8, p. 268, 2024.

PEDREGOSA, F.; VAROQUAUX, G.; GRAMFORT, A.; MICHEL, V.; THIRION, B.; GRISEL, O.; BLONDEL, M.; PRETTENHOFER, P.; WEISS, R.; DUBOURG, V.; VANDERPLAS, J.; PASSOS, A.; COURNAPEAU, D.; BRUCHER, M.; PERROT, M.; DUCHESNAY, E. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

PETERSON, L. E. K-nearest neighbor. *Scholarpedia*, v. 4, n. 2, p. 1883, 2009.

PIKOVSKY, A.; ROSENBLUM, M. G.; KURTHS, J. *Synchronization, A Universal Concept in Nonlinear Sciences*. Cambridge: Cambridge University Press, 2001. Disponível em: <<https://books.google.co.in/books?id=Fulv845q3QUC>>.

PRESKILL, J. Quantum computing in the nisq era and beyond. *Quantum*, Verein zur Förderung des Open Access Publizierens in den Quantenwissenschaften, v. 2, p. 79, 2018.

QIAO, G. J.; LIU, X. Y.; LIU, H. D.; SUN, C. F.; YI, X. X. Quantum  $\phi$  synchronization in a coupled optomechanical system with periodic modulation. *Phys. Rev. A*, American Physical Society, v. 101, p. 053813, May 2020. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.101.053813>>.

QUANTUM, I. Zfeaturemap: Quantum circuit documentation. *IBM Quantum Documentation*, 2021. Disponível em: <<https://docs.quantum.ibm.com/api/qiskit/qiskit.circuit.library.ZFeatureMap>>.

RAJKOMAR, A.; DEAN, J.; KOHANE, I. Machine learning in medicine. *New England Journal of Medicine*, Mass Medical Soc, v. 380, n. 14, p. 1347–1358, 2019.

RAMESH, A.; PAVLOV, M.; GOH, G.; GRAY, S.; VOSS, C.; RADFORD, A.; CHEN, M.; SUTSKEVER, I. *Zero-Shot Text-to-Image Generation*. 2021. Disponível em: <<https://arxiv.org/abs/2102.12092>>.

RANI, V.; NABI, S. T.; KUMAR, M.; MITTAL, A.; KUMAR, K. Self-supervised learning: A succinct review. *Archives of Computational Methods in Engineering*, Springer, v. 30, n. 4, p. 2761–2775, 2023.

RAY, S. A quick review of machine learning algorithms. In: *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. [S.l.: s.n.], 2019. p. 35–39.

REM, B. S.; KÄMING, N.; TARNOWSKI, M.; ASTERIA, L.; FLÄSCHNER, N.; BECKER, C.; SENGSTOCK, K.; WEITENBERG, C. Identifying quantum phase transitions using artificial neural networks on experimental data. *Nature Physics*, Springer Science and Business Media LLC, v. 15, n. 9, p. 917–920, Jul 2019. ISSN 1745-2481. Disponível em: <<http://dx.doi.org/10.1038/s41567-019-0554-0>>.

RIBEIRO, M. T.; SINGH, S.; GUESTRIN, C. Why should i trust you?: Explaining the predictions of any classifier. In: *ACM. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. [S.l.], 2016. p. 1135–1144.

RIEGER, H.; UIMIN, G. The one-dimensional annni model in a transverse field: analytic and numerical study of effective hamiltonians. *Zeitschrift für Physik B Condensed Matter*, Springer Science and Business Media LLC, v. 101, n. 4, p. 597–611, dez. 1996. ISSN 1431-584X. Disponível em: <<http://dx.doi.org/10.1007/s002570050252>>.

RINGNÉR, M. What is principal component analysis? *Nature biotechnology*, Nature Publishing Group US New York, v. 26, n. 3, p. 303–304, 2008.

ROMERO, J. et al. Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz. *Quantum Science and Technology*, IOP Publishing, v. 4, n. 1, p. 014008, 2018.

ROSENBLATT, F. *The perceptron, a perceiving and recognizing automaton Project Para*. [S.l.]: Cornell Aeronautical Laboratory, 1957.

ROSENBLATT, F. *Two theorems of statistical separability in the perceptron*. [S.l.]: United States Department of Commerce Washington, DC, USA, 1958.

ROSENBLATT, F. Perceptron simulation experiments. *Proceedings of the IRE*, IEEE, v. 48, n. 3, p. 301–309, 1960.

ROULET, A.; BRUDER, C. Synchronizing the smallest possible system. *Phys. Rev. Lett.*, American Physical Society, v. 121, p. 053601, Jul 2018. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.121.053601>>.

RUDER, S. An overview of gradient descent optimization algorithms. *arXiv preprint arXiv:1609.04747*, 2016.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. *nature*, Nature Publishing Group UK London, v. 323, n. 6088, p. 533–536, 1986.

RUPPERT, D.; CARROLL, R. J. Theory & methods: Spatially-adaptive penalties for spline fitting. *Australian & New Zealand Journal of Statistics*, Wiley Online Library, v. 42, n. 2, p. 205–223, 2000.

RUSSELL, S. J.; NORVIG, P. *Artificial Intelligence. A Modern Approach*. Englewood Cliffs, New Jersey: Prentice Hall, 1995.

SACHDEV, S. *Quantum phase transitions*. Cambridge, NY: Cambridge University Press, 2011.

SANTOS, A. C. O computador quântico da IBM e o IBM quantum experience. *Revista Brasileira de Ensino de Física*, FapUNIFESP (SciELO), v. 39, n. 1, sep 2016. Disponível em: <<https://doi.org/10.1590%2F1806-9126-rbef-2016-0155>>.

SARASWAT, D.; BHATTACHARYA, P.; VERMA, A.; PRASAD, V. K.; TANWAR, S.; SHARMA, G.; BOKORO, P. N.; SHARMA, R. Explainable ai for healthcare 5.0: Opportunities and challenges. *IEEE Access*, v. 10, p. 84486–84517, 2022.

SARKER, I. H. Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN Computer Science*, Springer, v. 2, n. 6, p. 420, 2021.

SCHERER, S.; SCHINDLER, M. R. *A Chiral Perturbation Theory Primer*. 2005. Disponível em: <<https://arxiv.org/abs/hep-ph/0505265>>.

SCHMOLKE, F.; LUTZ, E. Noise-induced quantum synchronization. *Phys. Rev. Lett.*, American Physical Society, v. 129, p. 250601, Dec 2022. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.129.250601>>.

SCHUCH, N.; PÉREZ-GARCÍA, D.; CIRAC, I. Classifying quantum phases using matrix product states and projected entangled pair states. *Physical Review B*, American Physical Society (APS), v. 84, n. 16, Oct 2011. ISSN 1550-235X. Disponível em: <<http://dx.doi.org/10.1103/PhysRevB.84.165139>>.

SCHULD, M. *Supervised quantum machine learning models are kernel methods*. 2021.

SCHULD, M.; KILLORAN, N. Is quantum advantage the right goal for quantum machine learning? *PRX Quantum*, American Physical Society (APS), v. 3, n. 3, jul 2022. Disponível em: <<https://doi.org/10.1103%2Fprxquantum.3.030101>>.

SCHULD, M.; PETRUCCIONE, F. *Supervised Learning with Quantum Computers*. 1st. ed. [S.l.]: Springer Publishing Company, Incorporated, 2018. ISBN 3319964232.

SCHULD, M.; SINAYSKIY, I.; PETRUCCIONE, F. The quest for a quantum neural network. *Quantum Information Processing*, Springer, v. 13, p. 2567–2586, 2014.

SELKE, W. The anni model — theoretical analysis and experimental application. *Physics Reports*, v. 170, n. 4, p. 213–264, 1988. ISSN 0370-1573. Disponível em: <<https://www.sciencedirect.com/science/article/pii/0370157388901408>>.

SENTHILKUMAR, D. V.; CHANDRASEKAR, V. K. Local and global chimera states in a four-oscillator system. *Phys. Rev. E*, American Physical Society, v. 100, p. 032211, Sep 2019. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevE.100.032211>>.

SERNEELS, S.; NOLF, E. D.; ESPEN, P. J. V. Spatial sign preprocessing: A simple way to impart moderate robustness to multivariate estimators. *Journal of Chemical Information and Modeling*, v. 46, n. 3, p. 1402–1409, 2006. PMID: 16711760. Disponível em: <<https://doi.org/10.1021/ci050498u>>.

SHIINA, K.; MORI, H.; OKABE, Y.; LEE, H. K. Machine-learning studies on spin models. *Scientific Reports*, v. 10, n. 1, p. 2177, 2020. Disponível em: <<https://doi.org/10.1038/s41598-020-58263-5>>.

SHOR, P. W. Algorithms for quantum computation: discrete logarithms and factoring. In: IEEE. *Proceedings 35th annual symposium on foundations of computer science*. [S.l.], 1994. p. 124–134.

SHOR, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, SIAM, v. 41, n. 2, p. 303–332, 1999.

SHORTEN, C.; KHOSHGOFTAAR, T. M. A survey on image data augmentation for deep learning. *Journal of big data*, Springer, v. 6, n. 1, p. 1–48, 2019.

SHRIKUMAR, A.; GREENSIDE, P.; KUNDAJE, A. Learning important features through propagating activation differences. In: *Proceedings of the 34th International Conference on Machine Learning*. [S.l.]: PMLR, 2017. v. 70, p. 3145–3153.

SMITH, M. J. Getting value from artificial intelligence in agriculture. *Animal Production Science*, CSIRO Publishing, v. 60, n. 1, p. 46–54, 2018.

SONAR, S.; HAJDUŠEK, M.; MUKHERJEE, M.; FAZIO, R.; VEDRAL, V.; VINJANAMPATHY, S.; KWEK, L.-C. Squeezing enhances quantum synchronization. *Phys. Rev. Lett.*, American Physical Society, v. 120, p. 163601, Apr 2018. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.120.163601>>.

SPALL, J. C. Multivariate stochastic approximation using a simultaneous perturbation gradient approximation. *IEEE Transactions on Automatic Control*, IEEE, v. 37, n. 3, p. 332–341, 1992.

STEIN, B. V.; RAPONI, E.; SADEGHI, Z.; BOUMAN, N.; HAM, R. C. H. J. V.; BÄCK, T. A comparison of global sensitivity analysis methods for explainable ai with an application in genomic prediction. *IEEE Access*, v. 10, p. 103364–103381, 2022.

STÖCKL, A. *Evaluating a Synthetic Image Dataset Generated with Stable Diffusion*. 2022. Disponível em: <<https://arxiv.org/abs/2211.01777>>.

SUZUKI, S.; INOUE, J.-i.; CHAKRABARTI, B. K. *Quantum Ising phases and transitions in transverse Ising models*. [S.l.]: Springer, 2012. v. 862.

THIEM, T. N.; KOOSHKBAGHI, M.; BERTALAN, T.; LAING, C. R.; KEVREKIDIS, I. G. Emergent spaces for coupled oscillators. *Frontiers in Computational Neuroscience*, v. 14, 2020. ISSN 1662-5188. Disponível em: <<https://www.frontiersin.org/articles/10.3389/fncom.2020.00036>>.

TURING, A. M. et al. On computable numbers, with an application to the entscheidungsproblem. *J. of Math*, v. 58, n. 345-363, p. 5, 1936.

VENTURA, D.; MARTINEZ, T. Quantum associative memory. *Information sciences*, Elsevier, v. 124, n. 1-4, p. 273–296, 2000.

WÄCHTLER, C. W.; PLATERO, G. Topological synchronization of quantum van der pol oscillators. *Phys. Rev. Res.*, American Physical Society, v. 5, p. 023021, Apr 2023. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevResearch.5.023021>>.

WALTER, S.; NUNNENKAMP, A.; BRUDER, C. Quantum synchronization of a driven self-sustained oscillator. *Phys. Rev. Lett.*, American Physical Society, v. 112, p. 094102, Mar 2014. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.112.094102>>.

WALTER, S.; NUNNENKAMP, A.; BRUDER, C. Quantum synchronization of two van der pol oscillators. *Ann. Phys.*, v. 527, n. 1-2, p. 131–138, 2015. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/andp.201400144>>.

WANG, A. X.; CHUKOVA, S. S.; SIMPSON, C. R.; NGUYEN, B. P. Challenges and opportunities of generative models on tabular data. *Applied Soft Computing*, v. 166, p. 112223, 2024. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494624009979>>.

WECKER, D. et al. Progress towards practical quantum variational algorithms. *Physical Review A*, American Physical Society, v. 92, n. 4, p. 042303, 2015.

WEN, X. G. Vacuum degeneracy of chiral spin states in compactified space. *Phys. Rev. B*, American Physical Society, v. 40, p. 7387–7390, Oct 1989. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.40.7387>>.

WEN, X.-G. *Quantum Field Theory of Many-Body Systems: From the Origin of Sound to an Origin of Light and Electrons*. [S.l.]: Oxford University Press, 2007.

WENG, T.; CHEN, X.; REN, Z.; YANG, H.; ZHANG, J.; SMALL, M. Synchronization of machine learning oscillators in complex networks. *Information Sciences*, v. 630, p. 74–81, 2023. ISSN 0020-0255. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0020025523002177>>.

WHITE, S. R. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, American Physical Society, v. 69, p. 2863–2866, Nov 1992. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.69.2863>>.

WIDROW, B. Pattern recognition and adaptive control. *IEEE Transactions on Applications and Industry*, IEEE, v. 83, n. 74, p. 269–277, 1964.

WIRIYATHAMMABHUM, P.; SUMMERS-STAY, D.; FERMÜLLER, C.; ALOIMONOS, Y. Computer vision and natural language processing: recent approaches in multimedia and robotics. *ACM Computing Surveys (CSUR)*, ACM New York, NY, USA, v. 49, n. 4, p. 1–44, 2016.

WITTEN, D.; JAMES, G. *An introduction to statistical learning with applications in R*. [S.l.]: springer publication, 2013.

WOSSNIG, L. Quantum machine learning for classical data. *arXiv preprint arXiv:2105.03684*, 2021.

XIAO, X.; LU, T.-X.; ZHONG, W.-J.; LI, Y.-L. Classical-driving-assisted quantum synchronization in non-markovian environments. *Phys. Rev. A*, American Physical Society, v. 107, p. 022221, Feb 2023. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevA.107.022221>>.

XU, L.; SKOULARIDOU, M.; CUESTA-INFANTE, A.; VEERAMACHANENI, K. *Modeling Tabular data using Conditional GAN*. 2019. Disponível em: <<https://arxiv.org/abs/1907.00503>>.

XU, Y.; ZOMER, S.; BRERETON, R. G. Support vector machines: a recent method for classification in chemometrics. *Critical Reviews in Analytical Chemistry*, Taylor & Francis, v. 36, n. 3-4, p. 177–188, 2006.

YANG, Y.; SUN, Z.-Z.; RAN, S.-J.; SU, G. Visualizing quantum phases and identifying quantum phase transitions by nonlinear dimensional reduction. *Physical Review B*, American Physical Society (APS), v. 103, n. 7, Feb 2021. ISSN 2469-9969. Disponível em: <<http://dx.doi.org/10.1103/PhysRevB.103.075106>>.

YANG, Y.-T.; LUO, H.-G. *Pattern description of the ground state properties of the one-dimensional axial next-nearest-neighbor Ising model in a transverse field*. 2023.

YOSHIDA, B. Classification of quantum phases and topology of logical operators in an exactly solved model of quantum codes. *Annals of Physics*, v. 326, n. 1, p. 15–95, 2011. ISSN 0003-4916. January 2011 Special Issue. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0003491610001867>>.

YOUNG, T.; HAZARIKA, D.; PORIA, S.; CAMBRIA, E. Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*, IEEE, v. 13, n. 3, p. 55–75, 2018.

ZHANG, H.; FAN, H.; DU, Y.; WANG, L.; WANG, X. Anticipating measure synchronization in coupled Hamiltonian systems with machine learning. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, v. 32, n. 8, 08 2022. ISSN 1054-1500. 083136. Disponível em: <<https://doi.org/10.1063/5.0093663>>.

ZHIROV, O. V.; SHEPELYANSKY, D. L. Synchronization and bistability of a qubit coupled to a driven dissipative oscillator. *Phys. Rev. Lett.*, American Physical Society, v. 100, p. 014101, Jan 2008. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevLett.100.014101>>.

ZHIROV, O. V.; SHEPELYANSKY, D. L. Quantum synchronization and entanglement of two qubits coupled to a driven dissipative resonator. *Phys. Rev. B*, American Physical Society, v. 80, p. 014519, Jul 2009. Disponível em: <<https://link.aps.org/doi/10.1103/PhysRevB.80.014519>>.

ZOUFAL, C. Generative quantum machine learning. *arXiv preprint arXiv:2111.12738*, 2021.

ŠTĚRBA, D.; NOVOTNÝ, J.; JEX, I. Asymptotic phase-locking and synchronization in two-qubit systems. *Journal of Physics Communications*, IOP Publishing, v. 7, n. 4, p. 045003, apr 2023. Disponível em: <<https://dx.doi.org/10.1088/2399-6528/acc0d4>>.