



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE TECNOLOGIA E GEOCIÊNCIAS
DEPARTAMENTO DE ENGENHARIA MECÂNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

SAMUEL BEZERRA DA GAMA NETO

***QUANTUM MACHINE LEARNING APLICADO À DETECÇÃO DE FALHAS EM
TURBINAS EÓLICAS***

Recife

2024

SAMUEL BEZERRA DA GAMA NETO

***QUANTUM MACHINE LEARNING* APLICADO À DETECÇÃO DE FALHAS EM
TURBINAS EÓLICAS**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Mecânica do Departamento de Engenharia Mecânica, Centro de Tecnologia e Geociências da Universidade Federal de Pernambuco como parte dos requisitos parciais para obtenção do título de mestre em Engenharia Mecânica. Área de concentração: Energia.

Orientador: Prof. Dr. Alvaro Antonio Ochoa Villa

Coorientador: Prof. Dr. Gustavo de Novaes Pires Leite

Recife

2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Gama Neto, Samuel Bezerra da.

Quantum Machine Learning aplicado à detecção de falhas em turbinas eólicas / Samuel Bezerra da Gama Neto. - Recife, 2024. 116f.: il.

Dissertação (Mestrado) - Universidade Federal de Pernambuco, Centro de Tecnologia e Geociências, Programa de Pós-Graduação em Engenharia Mecânica, 2024.

Orientação: Alvaro Antonio Ochoa Villa.

Coorientação: Gustavo de Novaes Pires Leite.

Inclui referências e apêndices.

1. Turbinas eólicas; 2. Detecção de falha; 3. Quantum Machine Learning; 4. PCA; 5. Circuitos quânticos variacionais. I. Villa, Alvaro Antonio Ochoa. II. Leite, Gustavo de Novaes Pires. III. Título.

UFPE-Biblioteca Central

SAMUEL BEZERRA DA GAMA NETO

***QUANTUM MACHINE LEARNING APLICADO À DETECÇÃO DE FALHAS EM
TURBINAS EÓLICAS***

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Mecânica do Departamento de Engenharia Mecânica, Centro de Tecnologia e Geociências da Universidade Federal de Pernambuco como parte dos requisitos parciais para obtenção do título de mestre em Engenharia Mecânica. Área de concentração: Energia.

Aprovado em: 09/08/2024

BANCA EXAMINADORA

Prof. Dr. Alvaro Antonio Ochoa Villa (Orientador)
Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco - IFPE

Prof. Dr. Gustavo de Novaes Pires Leite (Coorientador)
Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco - IFPE

Prof. Dr. José Ângelo Peixoto da Costa (Examinador Interno)
Instituto Federal de Educação, Ciência e Tecnologia de Pernambuco - IFPE

Prof^ª. Dr^ª. Isis Didier Lins (Examinadora Externa)
Universidade Federal de Pernambuco – UFPE

AGRADECIMENTOS

Agradeço a Deus por me fortalecer e me guiar ao longo de mais esta etapa.

Agradeço aos meus pais, Madalena e José Carlos, por sempre me incentivarem a alcançar os meus objetivos e terem contribuído para minha formação humana. Aos meus irmãos, Lívia e Vinícius, pela motivação e encorajamento para seguir em frente nesta jornada. E aos meus familiares por todo o apoio durante a minha formação.

Agradeço a todos meus amigos, especialmente Ana, Eduardo e Juliana, pela amizade desde o período da graduação e por terem escutado todos os meus esforços para concluir esta etapa.

Agradeço aos meus orientadores, Prof. Alvaro e Prof. Gustavo, por terem sugerido este desafio e pelos ensinamentos compartilhados ao longo do mestrado.

Agradeço aos professores do Programa de Pós-Graduação em Engenharia Mecânica (PPGEM/UFPE) por terem contribuído para minha formação, e ao Centro de Energias Renováveis (CER/UFPE) pela disponibilização do cluster, que foi crucial para acelerar o treinamento dos modelos, e pela parceria no desenvolvimento de projetos no setor de energia eólica.

Por fim, agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo suporte financeiro, à Rio Amazonas S.A. pelo provimento dos dados durante o projeto de P&D ANEEL NEO.PROGFALHAS, e ao grupo CPFL Energia pelo apoio financeiro durante o projeto de P&D PD-00063-3090 ANEEL.

RESUMO

A energia eólica desempenha um papel fundamental na transição energética e na descarbonização global do setor energético. Para fortalecer sua eficiência, os parques eólicos demandam a implementação de estratégias de redução de custos, inovações na gestão da operação e manutenção, e monitoramento contínuo da condição das turbinas eólicas. Na literatura, são exploradas diversas metodologias no setor da energia eólica, abrangendo métodos de previsão da velocidade do vento, estimativa da potência gerada, diagnóstico e prognóstico de falhas. Uma estratégia essencial para auxiliar na tomada de decisão com base nos dados é a detecção de falhas, que identifica desvios do comportamento padrão dos componentes, evitando tempo de inatividade das turbinas eólicas e prevenindo danos colaterais em outros componentes. Nesse contexto, este estudo propõe a implementação de modelos para detecção de falhas de rolamentos, utilizando dados reais de vibração de um parque eólico localizado no Brasil. A abordagem adotada utiliza o *Quantum Machine Learning* (QML), um campo ainda pouco explorado na área de turbinas eólicas, que utiliza dos fenômenos da mecânica quântica como a superposição e emaranhamento para construção dos modelos. Para a detecção de falhas foram considerados três tipos de modelos híbridos quântico-clássicos: *Quantum support vector machine* (QSVM), agrupamento espectral quântico (AEQ) e classificador quântico variacional. Os modelos foram treinados com atributos extraídos do sinal de vibração nos domínios do tempo, frequência e mecânico. Os resultados mostraram que o modelo de QSVM PauliFeatureMap-X apresentou desempenho igual ou superior aos modelos clássicos, com o valor de acurácia balanceada de 96,9%, 94,6% e 92,5% para os casos do rolamento da extremidade sem acionamento do gerador, rolamento principal e do rolamento do eixo de alta velocidade da caixa de engrenagens, respectivamente. Além disso, para o problema de agrupamento, os modelos de AEQ demonstraram uma performance superior em comparação aos modelos clássicos pela métrica da informação mútua normalizada. Os resultados demonstram o potencial do QML para auxiliar na operação e manutenção de turbinas eólicas.

Palavras-chave: turbinas eólicas; detecção de falha; *quantum machine learning*; pca; circuitos quânticos variacionais.

ABSTRACT

Wind energy plays a fundamental role in the energy transition and global decarbonization of the energy sector. To strengthen its efficiency, wind farms require the implementation of cost reduction strategies, innovations in operation and maintenance management, and continuous monitoring of the condition of wind turbines. In the literature, various methodologies in the wind energy sector are explored, covering methods for wind speed prediction, estimation of generated power, diagnosis and prognosis of failures. An essential strategy to assist decision-making based on data is fault detection, which identifies deviations from the standard behavior of components, avoiding turbine downtime and preventing collateral damage to other components. In this context, this study proposes the implementation of models for bearing fault detection, using real vibration data from a wind farm located in Brazil. The adopted approach utilizes Quantum Machine Learning (QML), a field still underexplored in the wind turbine area, which leverages quantum mechanical phenomena such as superposition and entanglement for model construction. For fault detection, three types of quantum-classical hybrid models were considered: Quantum support vector machine (QSVM), quantum spectral clustering (QSC), and variational quantum classifier. The models were trained with features extracted from the vibration signal in the time, frequency, and mechanical domains. The results showed that the QSVM PauliFeatureMap-X model performed equal to or better than classical models, with balanced accuracy values of 96.9%, 94.6%, and 92.5% for the non-drive end generator bearing, main bearing, and gearbox high-speed shaft bearing cases, respectively. Furthermore, for the clustering problem, the QSC models demonstrated superior performance compared to classical models according to the normalized mutual information metric. The results highlight the potential of QML to assist in the operation and maintenance of wind turbines.

Keywords: wind turbines; fault detection; quantum machine learning; pca; variational quantum circuits.

LISTA DE FIGURAS

Figura 1 - Consumo de energia elétrica por classe e ano no Brasil entre 2004 a 2023.	17
Figura 2 - Média mensal da velocidade do vento entre 1980 a 2022 no estado brasileiro onde o parque eólico está localizado.	17
Figura 3 - Número de publicações por área sobre QML até o ano de 2023 na base de dados Scopus.	25
Figura 4 - Principais componentes de turbinas eólicas.	28
Figura 5 - Principais componentes da Nacele.	29
Figura 6 - Porcentagem de falha dos principais componentes de turbinas eólicas.	30
Figura 7 - Geometria do rolamento.	31
Figura 8 - Estágios de degradação do rolamento no espectro de vibração. (a) Estágio inicial; (b) Estágio intermediário; (c) Estágio avançado; (d) Estágio crítico.	35
Figura 9 - Modelo de aprendizagem de máquina.	36
Figura 10 - Fluxograma da aprendizagem de máquina.	37
Figura 11 - Classificação linear pela técnica de SVM. (a) Conjunto de dados; (b) Hiperplanos de separação.	39
Figura 12 - Margem máxima de separação pelo SVM.	41
Figura 13 - SVM para classificação de dados não separáveis linearmente. (a) Baixa dimensão; (b) Alta dimensão.	43
Figura 14 - Comparação entre conjuntos de dados para agrupamento. (a) <i>clusters</i> bem agrupados; (b) <i>clusters</i> caracterizados por uma relação de similaridade.	44
Figura 15 - Agrupamento <i>k-means</i> .	46
Figura 16 - Pseudocódigo <i>k-means</i> .	46
Figura 17 - Rede neural artificial. (a) Neurônio artificial; (b) Perceptron multicamadas.	47
Figura 18 - Funções de ativação comuns. (a) Degrau; (b) ReLU; (c) Sigmoide; (d) Tangente hiperbólica.	48
Figura 19 - Fluxograma simplificado do treinamento de RNAs.	49
Figura 20 - Topologia das RNAs. (a) <i>Feedforward</i> ; (b) Recorrente.	50
Figura 21 - Representação da matriz de confusão.	51
Figura 22 - Representação geométrica de um <i>qubit</i> .	54
Figura 23 - Circuito quântico.	57
Figura 24 - Ação da porta quântica em um <i>qubit</i> . (a) Porta Pauli-X; (b) Porta Hadamard.	61
Figura 25 - Representação típica das portas para múltiplos <i>qubits</i> . (a) CNOT; (b) Toffoli.	62

Figura 26 - Tipos de abordagens da aprendizagem de máquina e computação quântica.	63
Figura 27 - Distribuição espacial das turbinas eólicas analisadas.	64
Figura 28 - Série temporal e espectro do envelope para o caso GnNDE. (a) TE normal; (b) TE com falha.	65
Figura 29 - Dispersão dos dados de cada domínio por caso de estudo. (a) MnBrg; (b) GbxHss; (c) GnNDE.	66
Figura 30 - Circuito quântico para determinação da função de kernel.	68
Figura 31 - Classificador quântico variacional.	70
Figura 32 - Pseudocódigo para combinação das estruturas do modelo CQV.	71
Figura 33 - Circuitos quânticos parametrizados utilizados no modelo CQV. (a) RealAmplitudes; (b) TwoLocal; (c) EfficientSU2.	72
Figura 34 - Fluxograma para os modelos de classificação de QML.	73
Figura 35 - Fluxograma para os modelos de agrupamento de QML.	74
Figura 36 - Variância explicada acumulada pelo número de componentes principais.	77
Figura 37 - Resultado da acurácia balanceada dos modelos clássicos para o conjunto de teste. (a) MnBrg; (b) GbxHss; (c) GnNDE.	78
Figura 38 - Métricas de performance do modelo SVM-polinomial para o conjunto de teste.	79
Figura 39 - Resultado da IMN dos modelos clássicos. (a) MnBrg; (b) GbxHss; (c) GnNDE.	80
Figura 40 - Resultado da acurácia balanceada dos modelos QSVM para o conjunto de teste. (a) MnBrg; (b) GbxHss; (c) GnNDE.	83
Figura 41 - Tempo de treinamento dos modelos de QSVM (a) MnBrg; (b) GbxHss; (c) GnNDE.	83
Figura 42 - Avaliação do número de repetições com 4 <i>qubits</i> para o modelo QSVM aplicado para o caso MnBrg.	84
Figura 43 - Resultado da IMN dos modelos de AEQ. (a) MnBrg; (b) GbxHss; (c) GnNDE.	86
Figura 44 - Avaliação do número de repetições com 4 <i>qubits</i> para o modelo AEQ aplicado para o caso MnBrg.	87
Figura 45 - Comparação da classe real pela posição dos <i>clusters</i> do modelo de AEQ de melhor performance por caso de estudo. (a) MnBrg; (b) GbxHss; (c) GnNDE.	87
Figura 46 - Boxplot da acurácia balanceada do conjunto de teste. (a) MnBrg; (b) GbxHss; (c) GnNDE.	90
Figura 47 - Resultado da métrica de acurácia balanceada do modelo CQV para o conjunto de teste. (a) Média; (b) Desvio padrão.	91

Figura 48 - Tempo médio de treinamento pelo aumento do número de *qubits* no modelo CQV.

LISTA DE QUADROS

Quadro 1 - Atributos do domínio do tempo para análise de vibração de rolamentos.	32
Quadro 2 - Comportamento característico de falha em rolamentos na análise de envelope.	34
Quadro 3 - Funções de kernel comuns.	44
Quadro 4 - Exemplos de funções de custo.	50
Quadro 5 - Resumo de algumas notações da mecânica quântica.	56
Quadro 6 - Principais bibliotecas do Python utilizadas na implementação dos modelos.	67
Quadro 7 - Mapas de atributos implementados pelas bibliotecas Pennylane e Qiskit no Python.	69
Quadro 8 - Identificação das combinações entre mapa de atributos, CQPs e otimizadores para os modelos CQV.	108

LISTA DE TABELAS

Tabela 1 - Métricas de performance de modelos para detecção de falhas de rolamentos em turbinas eólicas.	24
Tabela 2 - Descrição resumida do conjunto de dados de cada caso de estudo.	64
Tabela 3 - Distribuição das classes nos conjuntos de treinamento e teste.	76
Tabela 4 - Desvio padrão da acurácia balanceada dos modelos clássicos para o conjunto de teste.	79
Tabela 5 - Resultado da acurácia balanceada dos modelos de QSVM para o conjunto de teste.	82
Tabela 6 - Resultados da IMN para os modelos de AEQ.	85
Tabela 7 - Resultado da acurácia balanceada dos modelos de VQC para o conjunto de teste com 4 <i>qubits</i> .	89
Tabela 8 - Comparação dos modelos de melhor performance para a classificação.	93
Tabela 9 - Comparação dos modelos de melhor performance para o agrupamento.	93
Tabela 10 - Descrição do conjunto de dados para o caso MnBrg.	106
Tabela 11 - Descrição do conjunto de dados para o caso GbxHss.	106
Tabela 12 - Descrição do conjunto de dados para o caso GnNDE.	107
Tabela 13 - Métricas dos modelos clássicos para o conjunto de teste do caso MnBrg.	109
Tabela 14 - Métricas dos modelos clássicos para o conjunto de teste do caso GbxHss.	110
Tabela 15 - Métricas dos modelos clássicos para o conjunto de teste do caso GnNDE.	111
Tabela 16 - Resultado da IMN para os modelos clássicos de agrupamento.	112
Tabela 17 - Métricas dos modelos de QSVM para o conjunto de teste do caso MnBrg.	113
Tabela 18 - Métricas dos modelos de QSVM para o conjunto de teste do caso GbxHss.	114
Tabela 19 - Métricas dos modelos de QSVM para o conjunto de teste do caso GnNDE.	115
Tabela 20 - Métricas dos modelos de CQV para o conjunto de teste.	116

SUMÁRIO

1	INTRODUÇÃO	15
1.1	JUSTIFICATIVAS	17
1.2	OBJETIVOS	19
1.2.1	Objetivo geral	19
1.2.2	Objetivos específicos	19
1.3	ESTRUTURA DA DISSERTAÇÃO	20
2	ESTADO DA ARTE	21
2.1	CONTRIBUIÇÕES CIENTÍFICAS	26
3	REFERENCIAL TEÓRICO	28
3.1	ENERGIA EÓLICA	28
3.2	ANÁLISE DE VIBRAÇÃO DE ROLAMENTOS	31
3.3	APRENDIZAGEM DE MÁQUINA	35
3.3.1	Análise de componentes principais	37
3.3.2	<i>Support vector machine</i>	39
3.3.3	Agrupamento espectral	44
3.3.4	Redes neurais artificiais	46
3.3.5	Métricas de performance	51
3.4	COMPUTAÇÃO QUÂNTICA	53
3.4.1	Superposição quântica	57
3.4.2	Emaranhamento quântico	58
3.4.3	Interferência quântica	59
3.4.4	Portas quânticas	59
3.4.4.1	Portas quânticas para um <i>qubit</i>	60
3.4.4.2	Portas quânticas para múltiplos <i>qubits</i>	62
3.5	<i>QUANTUM MACHINE LEARNING</i>	63
4	METODOLOGIA	64
4.1	MÉTODOS DE KERNEL QUÂNTICO	67
4.2	CLASSIFICADOR QUÂNTICO VARIACIONAL	69
4.3	DESCRIÇÃO DA METODOLOGIA	72
4.3.1	Modelos quânticos para classificação	72
4.3.2	Modelos quânticos para agrupamento	74
5	RESULTADOS E DISCUSSÕES	76

5.1	PRÉ-PROCESSAMENTO	76
5.2	MODELOS CLÁSSICOS PARA DETECÇÃO DE FALHAS	78
5.3	MÉTODOS DE KERNEL QUÂNTICO	81
5.3.1	<i>Quantum support vector machine (QSVM)</i>	81
5.3.2	Agrupamento espectral quântico (AEQ)	85
5.4	CLASSIFICADOR QUÂNTICO VARIACIONAL (CQV)	88
5.5	COMPARAÇÃO DA PERFORMANCE ENTRE OS MODELOS CÁSSICOS E QUÂNTICOS	92
6	CONCLUSÕES	95
6.1	TRABALHOS FUTUROS	97
	REFERÊNCIAS	98
	APÊNDICE A – DESCRIÇÃO DETALHADA DO CONJUNTO DE DADOS POR CASO DE FALHA	106
	APÊNDICE B – IDENTIFICAÇÃO DAS COMBINAÇÕES UTILIZADAS PARA O MODELO CQV	108
	APÊNDICE C – MÉTRICAS DE PERFORMANCE DOS MODELOS CLÁSSICOS DE CLASSIFICAÇÃO PARA O CONJUNTO DE TESTE	109
	APÊNDICE D – MÉTRICA DE PERFORMANCE DOS MODELOS CLÁSSICOS DE AGRUPAMENTO	112
	APÊNDICE E – MÉTRICA DE PERFORMANCE DOS MODELOS QUÂNTICOS DE QSVM PARA O CONJUNTO DE TESTE	113
	APÊNDICE F – MÉTRICA DE PERFORMANCE DOS MODELOS QUÂNTICOS CQV PARA O CONJUNTO DE TESTE	116

1 INTRODUÇÃO

Atualmente, o aumento do consumo de energia elétrica tem ampliado as preocupações globais relacionadas ao aumento da emissão e concentração de gases de efeito estufa na atmosfera, além dos impactos ambientais adversos provocados por fontes não renováveis de energia. Diante desse cenário, as fontes de energia renovável têm se destacado como alternativas essenciais para reduzir os impactos prejudiciais da emissão de gases de efeito estufa. Dentre essas fontes, a energia eólica e a energia solar surgem como as mais populares e promissoras (SHER; CURNICK; AZIZAN, 2021). A energia eólica, em particular, tem sido cada vez mais explorada devido à sua disponibilidade abundante, baixo impacto ambiental e custos competitivos. No cenário mundial, entre os anos de 2023 e 2027, espera-se a adição de aproximadamente 680 GW na capacidade de geração de energia eólica em todo o mundo. Ademais, prevê-se que essa expansão seja parte de uma tendência mais ampla em direção a fontes de energia renovável. Até 2050, a energia renovável deve responder por 82% da geração de energia elétrica, com a energia eólica e solar representando 69% desse total (DNV, 2023; GWEC, 2023).

Um parâmetro comum utilizado para comparação entre fontes de energia é o custo nivelado de energia (LCOE). Este indicador relaciona todos os custos associados à geração de energia e a quantidade de energia gerada ao longo da vida útil de uma usina de geração de eletricidade. De acordo com IRENA (2023), entre 2010 e 2022, o LCOE da energia eólica *onshore* experimentou uma queda de 69%. Em 2022, a eólica se consolidou como a fonte renovável mais barata, com um LCOE de apenas 0,033 USD/kWh. Além disso, para o mesmo ano, o LCOE da energia eólica foi 52% inferior à opção mais barata de combustíveis fósseis.

Para assegurar a confiabilidade e a segurança no fornecimento de energia, os parques eólicos precisam implementar estratégias que visem à redução de custos e à integração de algoritmos inteligentes para operação e manutenção (O&M). Isso requer aprimoramentos nas áreas de projeto, industrialização e O&M das turbinas eólicas (VEERS *et al.*, 2023). A importância da otimização dos custos de O&M das turbinas eólicas se torna ainda mais evidente quando se considera que estes representam entre 20% e 25% do LCOE de sistemas eólicos atuais, conforme COSTA *et al.* (2021).

Turbinas eólicas podem ter seu funcionamento comprometido por diversos modos de falha como fadiga, rupturas e falhas mecânicas ou elétricas. As falhas em turbinas eólicas podem ser causadas por fatores intrínsecos, como corrosão, sobreaquecimento, vibrações excessivas, sobrecarga mecânica e desgaste, ou por condições externas, como a formação de

gelo nas pás, raios e rajadas de vento (SAIDI; BENBOUZID, 2021). Para lidar com esses problemas, o *Prognostic and Health Management* (PHM) utiliza tecnologias para prever, gerenciar e diagnosticar as falhas com base em dados históricos e do presente, minimizando o tempo de inatividade e melhorando a eficiência operacional (JAVED; GOURIVEAU; ZERHOUNI, 2017).

Uma técnica crucial nesse contexto é a manutenção baseada em condições e o diagnóstico de falhas, que monitora o sistema em busca de sinais de falha ou mau funcionamento, permitindo a realização de manutenções preventivas e evitando danos mais severos (GAO; LIU, 2021). Nessa perspectiva, as técnicas de aprendizado de máquina têm se tornado um campo popular no PHM, pois permitem uma análise mais precisa e rápida dos dados com a capacidade de aprender com as informações e identificar padrões de relacionamentos nos dados para prever falhas antes que elas ocorram (REZAMAND *et al.*, 2020).

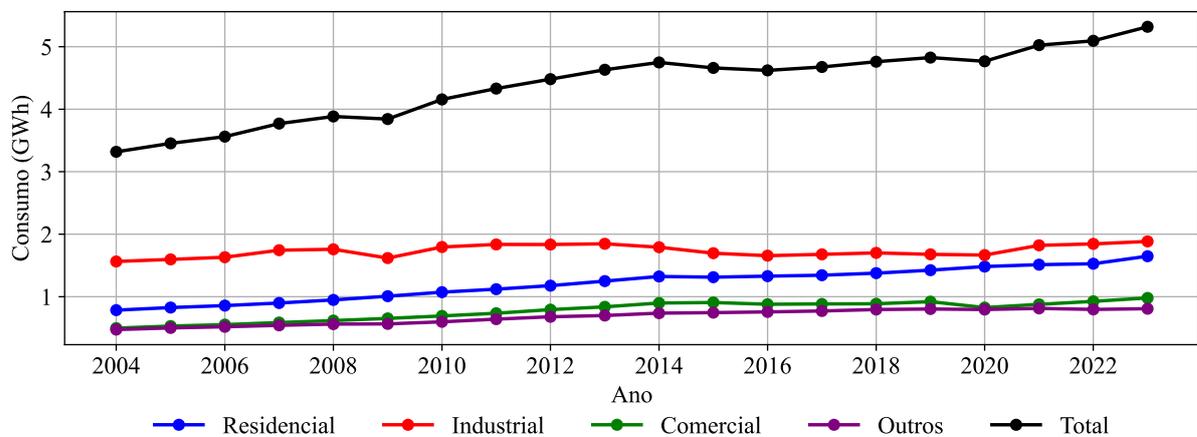
No que diz respeito à detecção de anomalias, ou seja, desvios inesperados do comportamento normal da turbina eólica, indicando potenciais problemas de funcionamento, como desgaste excessivo, falhas mecânicas ou condições ambientais adversas, os algoritmos de aprendizado de máquina utilizam os dados SCADA (*Supervisory Control and Data Acquisition*) e CMS (*Condition Monitoring System*) para o treinamento dos modelos. Os dados SCADA fornecem informações operacionais em tempo real, como velocidade do vento, temperatura, produção de energia e ajustes de controle da turbina eólica. Enquanto, os dados CMS são coletados por meio de sensores instalados na turbina eólica e monitoram parâmetros como vibração e temperatura dos rolamentos.

Um campo promissor no estudo para detecção de falhas é a exploração do potencial da computação quântica para a otimização de algoritmos clássicos de aprendizado de máquina (SCHULD; SINAYSKIY, I.; PETRUCCIONE, F., 2015). A computação quântica se baseia em fenômenos da mecânica quântica, como a superposição e o emaranhamento, e pode ser utilizada para a solução de problemas clássicos de maneira mais eficiente (ZHANG; NI, 2020). A concepção de integrar a computação quântica com aprendizado de máquina parte do pressuposto de que se os processadores quânticos de pequena escala conseguem produzir padrões estatísticos que são complexos para os processadores clássicos, então talvez sejam capazes de identificar padrões igualmente desafiadores para métodos clássicos (BIAMONTE *et al.*, 2017).

1.1 JUSTIFICATIVAS

Conforme mencionado, o crescimento das atividades econômicas e um maior consumo residencial de energia elétrica têm intensificado as preocupações tanto em relação às condições ambientais quanto à capacidade de geração suficiente para atender às demandas ao longo do ano. Esse cenário é particularmente evidente no Brasil, conforme demonstram os dados históricos de consumo de energia elétrica de 2004 a 2023, apresentados na Figura 1.

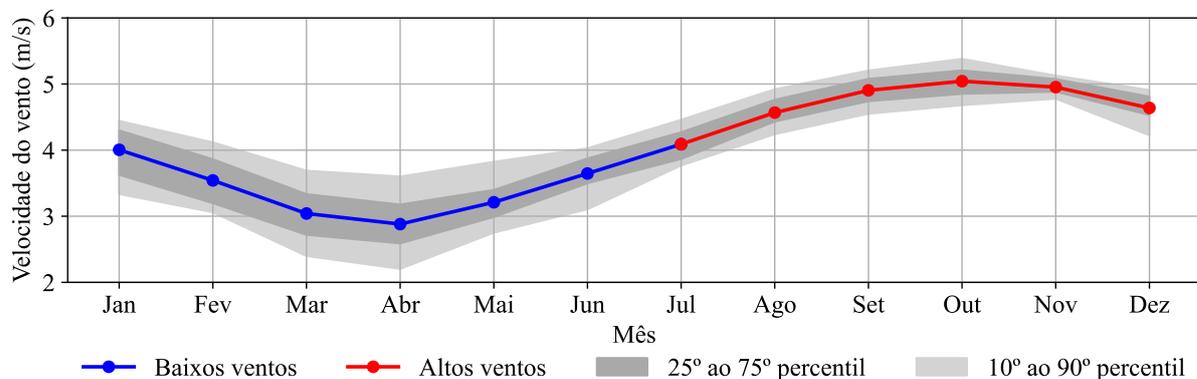
Figura 1 - Consumo de energia elétrica por classe e ano no Brasil entre 2004 a 2023.



Fonte: EPE (2024).

Adicionalmente, a energia eólica é uma fonte de energia intermitente, cuja capacidade de geração de eletricidade depende diretamente da disponibilidade e velocidade do vento. Um ano típico pode ser dividido em períodos de baixa e alta velocidade dos ventos, conforme apresentado na Figura 2 para o parque eólico analisado neste estudo.

Figura 2 - Média mensal da velocidade do vento entre 1980 a 2022 no estado brasileiro onde o parque eólico está localizado.



Fonte: RENEWABLES.NIJA (2024).

Dessa maneira, para otimizar o planejamento da manutenção das turbinas eólicas é essencial programar as atividades preventivas durante os períodos de baixos ventos. Isso evita o risco de falhas catastróficas durante épocas de ventos intensos, que poderiam resultar em períodos prolongados de inatividade e significativos prejuízos econômicos.

Além dos fatos anteriormente expostos, este estudo se motiva principalmente por:

- 1) A operação das turbinas eólicas ocorre em condições dinâmicas e adversas devido ao comportamento não linear do vento e o ambiente onde são geralmente instalados, o que os diferencia de outros equipamentos rotativos. Ademais, a detecção de falhas com base em dados de vibração enfrenta desafios adicionais devido à interação entre os componentes da turbina eólica e às variações de velocidade entre seus diferentes elementos;
- 2) A detecção de falhas em rolamentos por meio da análise do espectro de vibração é um procedimento complexo que exige uma compreensão detalhada do comportamento do equipamento, do tipo de falha presente e da expertise do analista. Como descrito por El-Thalji e Jantunen (2015), embora existam frequências características associadas a falhas em rolamentos, a análise de vibração no domínio da frequência pode ser dificultada pelos seguintes fatores: a) As amplitudes das frequências podem ser atenuadas devido a posição da falha em relação ao acelerômetro; b) As frequências podem ser moduladas com o surgimento de bandas laterais (*sidebands*). Com o agravamento do defeito, mais bandas laterais aparecem, até que, em determinado ponto, as frequências principais podem se tornar indetectáveis; c) No caso de múltiplos defeitos, as frequências geradas em rolamentos defeituosos podem se somar ou subtrair; e d) As frequências podem ser mascaradas por outras vibrações mais fortes. Desse modo, a detecção e diagnóstico de falha em rolamentos pode ser facilitado através de modelos de aprendizagem de máquina treinados com atributos extraídos do sinal de vibração;
- 3) As metodologias comumente utilizadas para detecção de falhas em rolamentos geralmente se baseiam em dados experimentais ou sintéticos. Esses dados são coletados em ambientes controlados, com o tipo de falha previamente estabelecido. Neste estudo, foram utilizados dados reais de operação de um parque eólico;
- 4) Por fim, as metodologias para detecção de falhas em rolamentos de turbinas eólicas são tradicionalmente baseadas na computação clássica, havendo poucos estudos desenvolvidos utilizando a computação quântica.

Nessa perspectiva, estudos no âmbito de detecção, diagnóstico e prognóstico de falhas em turbinas eólicas vem sendo desenvolvidos no Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Pernambuco (UFPE) na área de energia. Leite *et al.* (2021) aplicaram quantificadores da teoria da informação para monitorar e detectar mudanças nos sinais de vibração dos componentes de turbinas eólicas. Para o mesmo parque eólico abordado neste estudo, Leite *et al.* (2023) empregou onze modelos clássicos para detectar anomalias no rolamento sem acionamento do gerador de turbinas eólicas. Munguba *et al.* (2024) implementou técnicas de combinação de modelos de aprendizagem de máquina para detecção de falhas em componentes do *drivetrain* de turbinas eólicas. Por fim, Vieira *et al.* (2024) propôs uma estrutura para estimar o tempo de vida útil restante do rolamento principal de turbinas eólicas.

Neste contexto, este estudo visa aplicar técnicas de *Quantum Machine Learning* (QML) para detectar falhas em rolamentos de turbinas eólicas, utilizando dados reais de vibração de um parque eólico situado no Brasil. Além disso, pretende-se comparar o desempenho dos modelos quânticos com os modelos clássicos de *Machine Learning* (ML) em três cenários de falha distintos.

1.2 OBJETIVOS

1.2.1 Objetivo geral

Empregar técnicas de QML para a detecção de falhas em turbinas eólicas com base em dados reais de operação, e comparar seu desempenho com modelos clássicos de ML.

1.2.2 Objetivos específicos

- Investigar técnicas clássicas e quânticas de aprendizagem de máquina para detecção de falhas em turbinas eólicas, com base em estudos publicados na literatura;
- Aplicar técnicas de QML para o monitoramento da condição de turbinas eólicas, utilizando atributos extraídos dos dados de vibração em diferentes domínios;
- Analisar diferentes mapas de atributos para modelos quânticos baseado em kernels e avaliar a variação do número de *qubits* utilizados no treinamento dos modelos;

- Testar combinações entre mapa de atributos, circuito quânticos parametrizados e otimizadores para o classificador quântico variacional;
- Avaliar o desempenho dos modelos de QML em relação aos modelos clássicos, utilizando métricas de performance.

1.3 ESTRUTURA DA DISSERTAÇÃO

Esta dissertação está estruturada nos seguintes capítulos:

- No capítulo 1, são apresentadas as motivações e justificativas para o desenvolvimento deste estudo;
- No capítulo 2, são apresentados estudos relacionados à aplicação de técnicas de ML em turbinas eólicas;
- No capítulo 3, são apresentadas as bases teóricas, abrangendo as definições dos modelos de abordagem clássica e os conceitos de computação quântica aplicados ao QML;
- No capítulo 4, é detalhada a metodologia adotada para a detecção de falhas de rolamentos em três componentes de turbinas eólicas, utilizando tanto modelos tradicionais de aprendizado de máquina quanto modelos de QML;
- No capítulo 5, os resultados e as discussões da aplicação para os três cenários são apresentados;
- No capítulo 6, são expostas as conclusões finais do trabalho, ressaltando os principais resultados alcançados. Além disso, são apresentadas sugestões para pesquisas futuras, visando a expansão e aprimoramento das abordagens exploradas neste estudo.

2 ESTADO DA ARTE

Neste capítulo, serão apresentados estudos realizados nos últimos anos sobre a aplicação de técnicas de aprendizado de máquina em turbinas eólicas. A revisão abrange uma variedade de abordagens e métodos, destacando os principais resultados.

No contexto da energia eólica, modelos de aprendizagem de máquina tem sido empregados para auxiliar na tomada de decisão, previsão da potência gerada, e para detecção, diagnóstico e prognóstico de falhas (BENBOUZID *et al.*, 2021). No monitoramento da condição de equipamentos, para Gao e Liu (2021) e Encalada-Dávila *et al.* (2021), os modelos podem ser divididos em três áreas: detecção, diagnóstico e prognóstico de falhas. Modelos para detecção avaliam quando ocorrem mudanças significativas no comportamento normal do equipamento. O diagnóstico consiste na detecção, localização do componente e determinação do tipo de falha. E o prognóstico envolve a estimativa do tempo restante de operação depois da identificação da falha.

Demolli *et al.* (2019) utilizaram cinco modelos de ML de regressão, como *support vector regression* (SVR), *k-nearest neighbors* (KNN) *regression* e *extreme gradient boosting* (xGBoost), para previsão diária da produção da energia eólica, utilizando dados da velocidade do vento de quatro anos de cinco locais da Turquia e empregando o coeficiente de determinação (R^2), o erro médio absoluto (MAE) e a raiz do erro médio quadrático (RMSE), para avaliação dos modelos. Utilizando as mesmas métricas, Karaman (2023) combinou dados meteorológicos e do sistema SCADA para o treinamento de quatro arquiteturas de redes neurais, como a rede neural convolucional (CNN) e a rede neural recorrente (RNN), para previsão da produção de energia. Para ambas os estudos, os modelos de melhor desempenho alcançaram um valor de R^2 superior a 0,95.

Khosravi, Machado e Nunes (2018) aplicaram diferentes técnicas de ML para previsão da velocidade do vento para um parque eólico, localizado no sul do Brasil. Os modelos foram desenvolvidos com base nos dados históricos da velocidade do vento com diferentes intervalos de tempo. Para esta abordagem, os dados foram divididos em 70% para o conjunto de dados de treinamento e 30% para o conjunto de teste. O desempenho máximo foi obtido para os modelos de rede neural *feed-forward* multicamadas (RNFFM) e SVR com R^2 aproximadamente de 0,99.

Jiménez, Muñoz e Márquez (2017) propuseram modelos de aprendizagem de máquina para detecção e diagnóstico nas pás de turbinas eólicas devido à delaminação, que é um problema típico estrutural que ocorre em razão da separação das camadas do material compósito das pás. A aquisição dos dados foi realizada por meio de transdutores, em um experimento com

seis cenários de afastamento. Os modelos de classificação foram treinados com atributos extraídos por meio de um modelo autoregressivo, chamado de Yule–Walker. O modelo de rede neural artificial (RNA) alcançou uma acurácia de 91,50% na classificação correta dos casos positivos.

Xiang *et al.* (2021) desenvolveram uma abordagem com a combinação de CNN, da rede de memória de longo e curto prazo (LSTM) e de um mecanismo de atenção para fortalecer o impacto de informações importantes, para detecção de falha em três casos: a) falha na caixa de engrenagens; b) falha no gerado, devido à falha no isolamento do estator do alternador e baixa fase do rotor; c) turbina eólica saudável, avaliando a previsão da temperatura do rolamento da caixa de engrenagens. Os resultados mostraram que a abordagem obteve R^2 de 0,98 e 0,97 para os casos de falha.

Utilizando as séries temporais da temperatura do óleo e do rolamento, Dhiman *et al.* (2021) combinaram o *twin support vector machines* (TSVM) e uma técnica de limiar adaptativo para detecção de falha da caixa de engrenagens. A metodologia proposta foi comparada com quatro modelos de classificação binária, resultando em uma acurácia máxima de 95,84% para o modelo proposto. Para o mesmo componente, porém em uma tarefa de regressão, Heydari *et al.* (2021) sugeriram uma estrutura baseada na decomposição em modo variacional, RNA-GMDH e no algoritmo de otimização multiverso para a previsão das séries temporais das temperaturas do óleo e do rolamento, avaliada para dois casos de estudo.

No contexto do monitoramento de rolamentos em turbinas eólicas, Ali *et al.* (2018) apresentaram um método para o diagnóstico do rolamento do eixo de alta velocidade, que se baseia na extração de atributos nos domínios do tempo, frequência e tempo-frequência, visando o treinamento de uma rede neural adaptativa do tipo teoria da ressonância adaptativa 2 (ART2), para uma classificação não supervisionada. Os dados brutos de vibração foram coletados ao longo de um período de 50 dias, com 45 atributos sendo utilizados para o treinamento do modelo. O modelo alcançou uma acurácia de 90%, com uma taxa de falsos alarmes de 10%.

Elasha *et al.* (2019) propuseram o uso do ML para determinar o tempo de vida útil restante da caixa de engrenagens, utilizando os dados de vibração do rolamento do eixo de alta velocidade de uma turbina eólica de 2 MW. O desempenho da RNA foi comparado ao de modelos de regressão polinomial e exponencial, revelando que o modelo de RNA obteve resultados superiores, com um erro médio quadrático (MSE) de 5,62.

Para detecção da falha do rolamento principal, Beretta *et al.* (2021) implementaram um modelo de combinação, construído por uma RNA e o modelo de floresta de isolamento, treinado somente com dados do sistema SCADA. Neste procedimento, a temperatura do rolamento

principal foi selecionada como variável de interesse. A combinação dos modelos foi realizada através da avaliação de seus indicadores e comparada usando um algoritmo de decisão. Para um limite de decisão de 0,7 ou superior, observou-se que as melhores métricas de desempenho são obtidas para o modelo combinado, alcançando uma acurácia máxima de 0,986.

Considerando os dados brutos de vibração, Xu, Li; Yang (2021) desenvolveram um modelo de CNN multiescala aprimorado com um mecanismo de atenção de atributos. Para validar a metodologia, diferentes condições de ruído foram aplicadas ao conjunto de dados da Case Western Reserve University (CWRU). Posteriormente, foi considerado um caso de teste real de falha de rolamentos de duas turbinas eólicas. No caso de validação, o modelo proposto demonstrou um desempenho superior, com uma acurácia 6% maior em comparação aos outros modelos considerados.

Para determinar a melhor posição e o número de sensores de vibração para detecção de falha em rolamentos, Dibaj, Gao e Nejad (2023) usaram dados de vibração de uma turbina eólica *offshore* com três casos de falha: a) rolamento principal; b) rolamento do eixo de alta velocidade; c) rolamento do eixo de baixa velocidade. O método considerou a análise de componentes principais (PCA) e a arquitetura CNN para a extração dos atributos e a RNA do tipo perceptron multicamadas (MLP) para classificação.

Tian *et al.* (2023) sugeriram a aplicação de uma rede neural adaptativa multi-adversária dinâmica para detecção de falha de rolamentos, utilizando dados de laboratório. Para extração dos atributos, foi aplicado a estrutura CNN profunda. O método foi então comparado com cinco modelos de aprendizagem profunda. O modelo proposto alcançou a maior acurácia média entre os modelos avaliados, atingindo 83,33%.

Para o mesmo parque eólico investigado neste estudo, Leite *et al.* (2023) implementaram uma abordagem robusta, não supervisionada, para detecção de falha no rolamento da extremidade sem acionamento do gerador de turbina eólica, utilizando dados de vibração de doze turbinas eólicas. Atributos do domínio do tempo, frequência e mecânico foram extraídas e caracterizadas por meio da análise do espectro da frequência, conduzida por um especialista. A metodologia analisou onze modelos de aprendizagem de máquina, utilizando uma estratégia 5x2 de validação cruzada para garantir a melhor generalização dos modelos. Os melhores modelos foram selecionados com base em métricas de desempenho, como acurácia, *f1-score*, precisão e *recall*. Os resultados alcançaram uma taxa de 96,3% na detecção de falhas.

Ampliando a detecção de falhas para o rolamento principal, para o rolamento do eixo de alta velocidade da caixa de engrenagens e para o rolamento da extremidade com acionamento do gerador do mesmo parque eólico, Munguba *et al.* (2024) utilizaram de técnicas de

combinação entre as saídas de dezenove modelos de ML. A análise da importância dos atributos através da permutação revelou que os atributos do domínio do tempo são mais significativos para a correta classificação das classes. Além disso, os valores alcançados para a área sob a curva (AUC) foram de 98%, 89% e 95% para os casos do rolamento principal, do rolamento do eixo de alta velocidade e do rolamento da extremidade sem acionamento do gerador, respectivamente.

Em complemento, Vieira *et al.* (2024) exploraram o uso de aprendizagem de máquina para o prognóstico do rolamento principal, estudando o mesmo parque eólico. No entanto, a abordagem utilizou dados do sistema SCADA, como temperatura, velocidade do vento e potência ativa. Nesse estudo, o procedimento incluiu a filtragem dos dados e a criação de perfis de degradação, permitindo que os modelos de regressão fossem treinados utilizando uma estratégia de validação cruzada. Na etapa de validação, os valores médios das métricas MAE, MSE, RMSE e R^2 entre os modelos analisados foram 0,025, 0,004, 0,063 e 0,861, respectivamente. Ademais, os modelos apresentaram uma média de 20 dias para estimativa do tempo de vida útil restante.

A Tabela 1 apresenta os resultados de desempenho de diferentes estudos sobre a detecção de falhas em rolamentos de turbinas eólicas, publicados na literatura, utilizando as métricas de acurácia e F1 *score*.

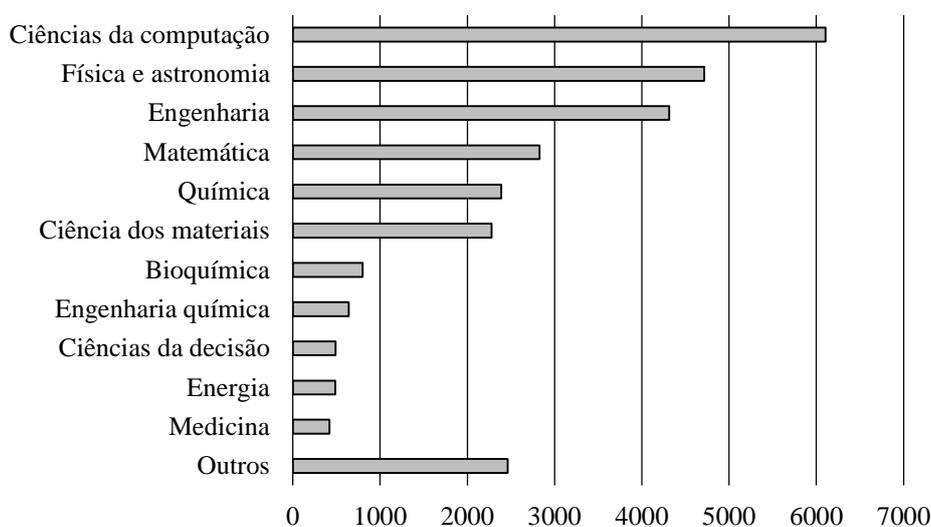
Tabela 1 - Métricas de performance de modelos para detecção de falhas de rolamentos em turbinas eólicas.

Referência	Método	Acur.	F1_score
Meng <i>et al.</i> (2022)	Extração de atributo baseada na decomposição de modo empírico (EMD) + KNN	1.000	-
Yang, S. <i>et al.</i> (2022)	CNN + RF	0.995	0.996
Cui, Weng e Zhang (2022)	Atributos dos domínio tempo e frequência + análise de componentes de vizinhança + KNN	0.875	-
Cui, Weng e Zhang (2022)	Atributos dos domínio tempo e frequência + ACV + RNA	0.833	-
Dong <i>et al.</i> (2022)	Entropia multiescala composta + seleção de atributo espectral robusta + máquina de aprendizagem extremo (ELM)	0.999	0.999
Leite <i>et al.</i> (2023)	Atributos do domínio do tempo, da frequência e mecânicos + validação cruzada 5x2 + <i>clustering-based local outlier</i>	0.904	0.937

Fonte: O autor (2024).

No campo da computação quântica, o estudo sobre o desenvolvimento e aplicação do QML tem sido uma área de interesse na academia, como demonstrado pelo número de publicações na base de dados Scopus na Figura 3. Até o ano de 2023, mais de 27 mil documentos foram publicados em diversas disciplinas, identificadas pela seguinte consulta “TITLE-ABS-KEY (qml OU *quantum machine learning* OU *variational quantum classifier* OU vqc OU *quantum support vector machine* OU qsvm OU *quantum neural network* OU qnn)”.

Figura 3 - Número de publicações por área sobre QML até o ano de 2023 na base de dados Scopus.



Fonte: O autor (2024).

Apesar do crescente número de publicações sobre QML, poucos trabalhos têm abordado a detecção de falhas. Um exemplo é o estudo de Silva e Droguett (2022) que realizaram a primeira abordagem híbrida de aprendizagem de máquina quântica-clássica para um problema de PHM, utilizando o conjunto de dados de rolamentos da *Machinery Failure Prevention Technology Society* (MFPT). O modelo constituído por um circuito quântico parametrizado com três portas de rotação e uma rede neural clássica foram utilizados para uma tarefa de classificação para dois tipos de falhas localizadas no anel externo e no anel interno, bem como a condição normal. Os resultados demonstraram que treinando o modelo 25 vezes, a acurácia foi de 97,6% para treinamento e 97,4% para teste.

Correa-Jullian *et al.* (2022) aplicaram QML para detecção de falhas de pitch em turbinas eólicas. A análise utilizou dados SCADA e os registros de alarmes de uma única turbina eólica para um problema de classificação. Duas técnicas para redução de dimensionalidade foram utilizadas, PCA e autoencoder (AE). Os algoritmos de *random forest* (RF), KNN e *support vector machines* (SVM) foram comparados em relação ao modelo de *quantum-support vector*

machines (QSVM). A maior acurácia obtida para o conjunto de teste foi 0,945 para o SVM e 0,924 para o QSVM, ambos modelos com a dimensão de 19 features pela técnica de AE.

MAIOR *et al.* (2023) propuseram o QML para máquinas rotativas utilizando dois conjuntos de dados de rolamentos, CWRU e *Jiangnan University* (JNU), avaliando diferentes configurações de circuitos quânticos parametrizados e analisando a variação do número de atributos e camadas do circuito. A configuração geral do modelo híbrido foi constituída pela codificação angular dos dados de entrada, processamento dos dados pelo circuito quântico e pela rede neural clássica. Os resultados evidenciaram que os modelos de QML superaram os modelos de rede do tipo MLP nas métricas calculadas.

De modo semelhante, Lins *et al.* (2024) investigaram o uso do QML para a detecção de sonolência, utilizando dados de eletroencefalograma (EEG) com o objetivo de prevenir acidentes em ambientes industriais. Nesse estudo, doze configurações de redes neurais quânticas (RNQs) foram consideradas, constituídas por composições das portas quânticas de rotação (R_y , R_z e R_y) e as portas CNOT, CZ, e iSWAP para cinco casos de estudo. As métricas de desempenho foram calculadas a partir de 10 execuções independentes dos modelos. A configuração R_y , R_z e $R_y + iSWAP$, com 10 camadas, aplicada ao caso 8, obteve a maior acurácia média de 99,35%, com um desvio padrão de 0,09.

2.1 CONTRIBUIÇÕES CIENTÍFICAS

Este estudo propõe o uso da computação quântica para a detecção de falhas em rolamentos de turbinas eólicas, utilizando dados de vibração, para que possa auxiliar na gestão estratégica da manutenção de parques eólicos. As principais contribuições da dissertação são:

- 1) Muitas aplicações de detecção de falhas em rolamentos são limitadas a dados de laboratório ou são implementadas em um número reduzido de turbinas eólicas. Neste estudo, foram utilizados dados de vibração de 28 turbinas eólicas de um parque eólico;
- 2) Muitos estudos anteriores sobre detecção de falhas utilizaram apenas um único domínio para a extração de atributos, o que pode não garantir a obtenção significativa das informações do sinal de vibração para caracterização da falha. Portanto, neste estudo, foram utilizados atributos dos domínios do tempo, frequência e mecânico como entradas para os modelos de detecção;
- 3) Para a detecção de falhas em rolamentos com modelos de QML, estudos anteriores concentraram-se em uma única abordagem aplicada a dados experimentais de vibração.

Neste estudo, foram empregadas estratégias de classificação supervisionada e não supervisionada para detecção de falha em três casos de estudo, utilizando dados operacionais reais de turbinas eólicas.

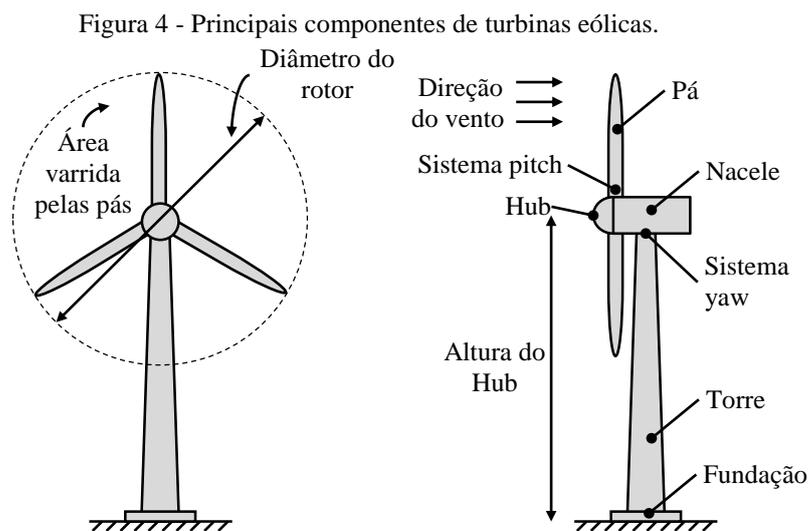
3 REFERENCIAL TEÓRICO

Este capítulo tem por finalidade introduzir os conceitos teóricos da literatura relacionados à detecção de falhas em turbinas eólicas. Primeiramente, o equipamento estudado neste trabalho é apresentado. Em seguida, são abordados os conceitos clássicos de aprendizado de máquina. Por fim, são expostas as definições da computação quântica aplicadas aos modelos de *Quantum Machine Learning*.

3.1 ENERGIA EÓLICA

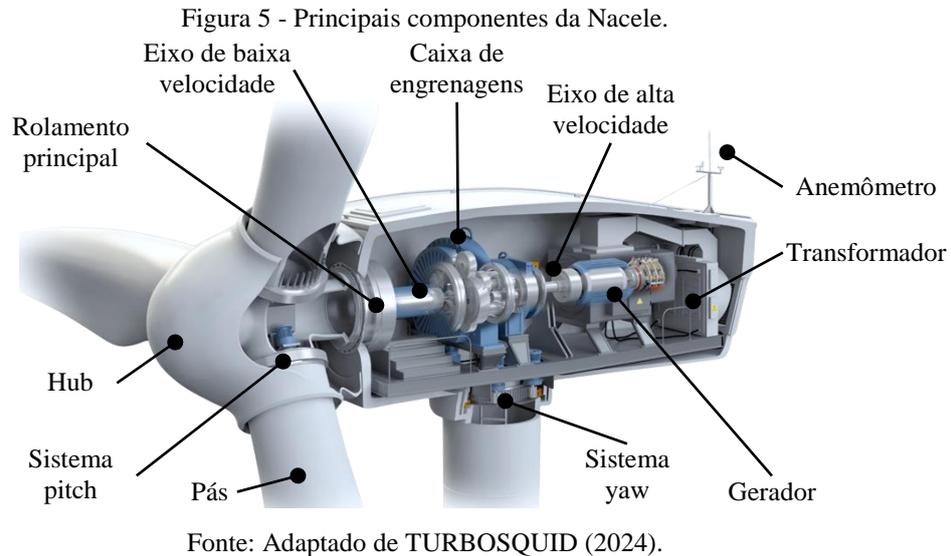
Turbinas eólicas (TEs) são equipamentos rotativos que transformam a energia cinética contida na velocidade do vento em energia elétrica, por meio da rotação de pás aerodinamicamente desenvolvidas. Essas pás estão conectadas a um gerador, que converte o movimento rotativo em eletricidade. Posteriormente, essa eletricidade é transferida para a rede elétrica para distribuição e uso pelos consumidores.

TEs modernas são constituídas por três pás com uma configuração padrão, projetadas para operar em uma ampla gama de velocidades e capazes de gerar eletricidade em uma escala de megawatts (MW) (BENBOUZID *et al.*, 2021). A Figura 4 apresenta os principais componentes de uma TE, conforme descrito nos trabalhos de Jiang *et al.* (2017) e Olabi *et al.* (2021).



Fonte: O autor (2024).

O movimento de rotação gerado pela ação do vento nas pás é transmitido pelo hub para o eixo de baixa velocidade, localizado na nacele. Este eixo transfere a energia mecânica para a caixa de engrenagens, que aumenta a velocidade de rotação para um nível adequado para o gerador elétrico, conforme mostrado na Figura 5.



Conforme destacado na Figura 5, alguns dos principais componentes das turbinas eólicas são:

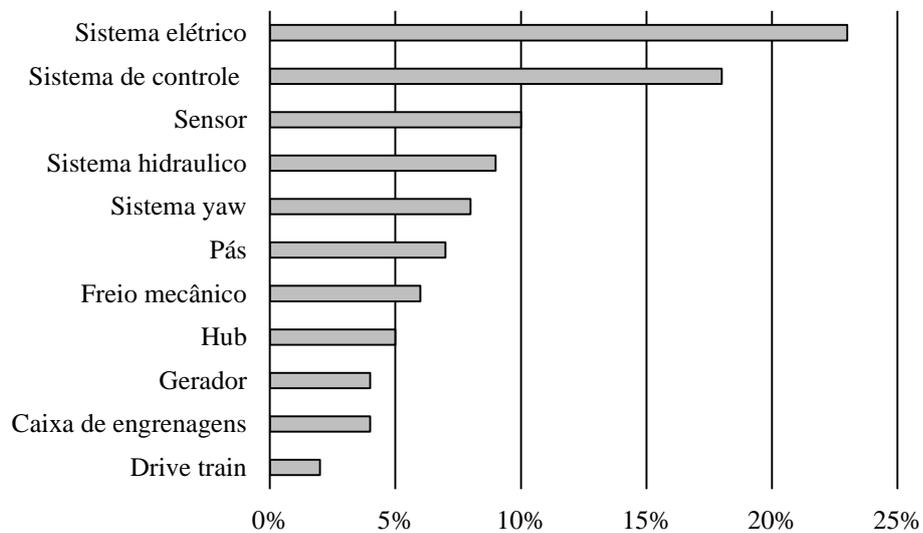
- a) Sistema de pitch: Controle individual do ângulo de cada pá, otimizando a captura de energia cinética. Além disso, pode atuar como freio em condições que excedem os limites operacionais;
- b) Hub: Transfere a energia mecânica para o eixo de baixa velocidade;
- c) Caixa de engrenagens: Aumenta a velocidade de rotação;
- d) Sistema yaw: Controle da rotação da nacele em relação ao eixo da torre, garantindo que a nacele permaneça perpendicular à direção do vento predominante;
- e) Gerador: Transforma a energia mecânica em energia elétrica.

Para a condição de maior eficiência teórica, a potência elétrica média gerada por uma TE aumenta exponencialmente com o aumento da velocidade do vento, conforme a Equação 1, onde P representa a potência elétrica gerada (W), ρ é a densidade do ar (kg/m^3), A é a área varrida pelas pás da TE (em metros quadrados), e V é a velocidade do vento (m/s) (MASSERAN, 2015).

$$P = \frac{1}{2} \rho A V^3 \quad (1)$$

O desempenho de TEs pode ser comprometido devido a uma série de falhas que podem ocorrer ao longo de sua vida útil. Essas falhas podem se manifestar de várias maneiras, abrangendo aspectos mecânicos, elétricos e de controle. A Figura 6 apresenta a distribuição da porcentagem de falha de cada componente para 1500 TEs em um período de 15 anos.

Figura 6 - Porcentagem de falha dos principais componentes de turbinas eólicas.



Fonte: Adaptado de Hahn, Durstewitz e Rohrig (2007)

Para o caso específico dos rolamentos em turbinas eólicas, diversos modos de falha podem surgir devido às condições operacionais e ao ambiente adverso em que operam. Entre os principais modos de falha estão deformação plástica, desgaste por atrito, erosão elétrica e fadiga (PENG *et al.*, 2022).

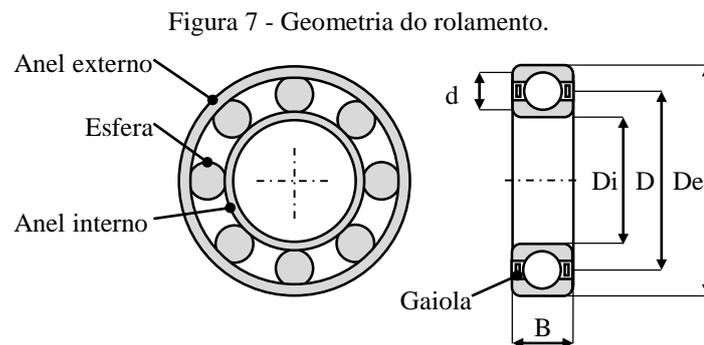
A fim de garantir a confiabilidade e maximizar a produção de energia eólica, as turbinas eólicas são submetidas a estratégia da manutenção pelo monitoramento por condição (MC). Sistemas de MC auxiliam na determinação do momento ótimo para manutenção, evitando ações de reparo desnecessárias e reduzindo o tempo de inatividade (PANDIT; ASTOLFI; DURAZO, 2023).

Os sistemas de MC aplicados na energia eólica coletam informações do equipamento em tempo real através de sensores distribuídos estrategicamente ao longo dos componentes da TE, auxiliando na identificação de desvios no comportamento normal (ENCALADA-DÁVILA *et al.*, 2021).

Para a operação e monitoramento da condição de TEs, dois sistemas de aquisição de dados são aplicados. O sistema SCADA (*Supervisory Control and Data Acquisition*) captura e armazena os valores de máximo, mínimo, média e desvio padrão de cada variável medida durante o período de dez minutos. Por outro lado, o sistema CMS (*Condition Monitoring System*) é projetado especificamente para monitorar o estado de saúde das turbinas eólicas (YANG, W.; JIANG, 2011).

3.2 ANÁLISE DE VIBRAÇÃO DE ROLAMENTOS

Rolamentos são dispositivos mecânicos aplicados na indústria em equipamentos rotativos para suportar cargas radiais e axiais, proporcionando uma interface de baixo atrito entre partes móveis. Eles são compostos por anéis internos e externos, juntamente com uma série de esferas, rolos ou elementos de rolamento que facilitam o movimento suave das peças adjacentes, como ilustrado na Figura 7.



Fonte: O autor (2024).

No monitoramento da condição de rolamentos, a análise de vibração de rolamentos é categorizada em três etapas: aquisição de dados, processamento de sinais e reconhecimento de falhas. Na análise de vibração, os sinais de vibração podem ser analisados em diferentes domínios, incluindo o domínio do tempo, da frequência e do tempo-frequência (GHAZALI; RAHIMAN, 2021).

No domínio do tempo, a análise de vibração se concentra em extrair os atributos diretamente da série do sinal. Esses atributos são então comparados com valores de referência recomendados ou utilizados para identificar padrões e tendências que possam indicar problemas no rolamento (GUPTA; PRADHAN, 2017). O Quadro 1 apresenta os atributos específicos do domínio do tempo que são frequentemente empregados na análise de vibração.

Quadro 1 - Atributos do domínio do tempo para análise de vibração de rolamentos.

Nome	Expressão	Nome	Expressão
Média	$\mu = \frac{1}{n} \sum_{i=1}^n x_i$	Pico a pico	$p2p$ $= \max(x) - \min(x)$
Variância	$\sigma = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2$	Fator de crista	$FC = \frac{\max(x)}{RMS}$
Desvio padrão	$std = (\sigma)^{\frac{1}{2}}$	Assimetria	$SK = \frac{1}{n} \sum_{i=1}^n \frac{(x_i - \mu)^3}{std^3}$
Raiz quadrada média (RMS)	$RMS = \left(\frac{1}{n} \sum_{i=1}^n x_i^2\right)^{\frac{1}{2}}$	Curtose	$K = \frac{\sum_{i=1}^n (x_i - \mu)^4}{n \cdot \sigma^2}$
Amplitude máxima	$\max_{amp} = \max(x)$	Energia	$E = \sum_{i=1}^n (x_i)^2$

Fonte: O autor (2024).

Na análise de vibração de rolamentos no domínio da frequência, os dados são convertidos de sinais do domínio do tempo para componentes de frequência por meio de técnicas como a transformada rápida de Fourier (FFT) (LIN; YE, 2019). A análise tradicional pela FFT proporciona uma decomposição do sinal em funções senoidais, que são caracterizadas por suas amplitudes e frequências. Essa relação é formalizada pela Equação 2, na qual $x(t)$ representa o sinal no domínio do tempo, $\omega = 2\pi f$ é a frequência angular e $F(\omega)$ é a transformada de $x(t)$ (SINGRU *et al.*, 2018).

$$F(\omega) = \int_{-\infty}^{+\infty} x(t)e^{-i\omega t} dt \quad (2)$$

Esse processo revela as características espectrais das vibrações, permitindo a identificação de frequências associadas a defeitos específicos nos rolamentos, como passagem de rolo, falhas de gaiola e danos nas pistas. Essas falhas resultam em uma série de efeitos recorrentes que ocorrem em intervalos regulares, denominadas por frequências características:

- a) BPFO (*Ball Pass Frequency Outer*): Falhas na pista externa do rolamento, como definido pela Equação 3, onde n é o número de elementos rolantes, Ω é a velocidade de

rotação, d é o diâmetro do elemento rolante, D é o diâmetro de giro e ϕ é o ângulo de contato;

- b) BPF_I (*Ball Pass Frequency Inner*): Falhas na pista interna do rolamento, definida pela Equação 4;
- c) FTF (*Fundamental Train Frequency*): Defeitos na gaiola do rolamento, definida pela Equação 5;
- d) BSF (*Ball Spin Frequency*): Defeitos nos elementos rolantes (esferas, roletes), definida pela Equação 6.

$$BPFO = \frac{n\Omega}{2} \left(1 - \frac{d}{D} \cos\phi\right) \quad (3)$$

$$BPF_I = \frac{n\Omega}{2} \left(1 + \frac{d}{D} \cos\phi\right) \quad (4)$$

$$FTF = \frac{\Omega}{2} \left(1 - \frac{d}{D} \cos\phi\right) \quad (5)$$

$$BSF = \frac{D}{2d} \left(1 + \left(\frac{d}{D} \cos\phi\right)^2\right) \quad (6)$$

No entanto, estas são frequências teóricas que não consideram o deslizamento dos elementos rolantes. O deslizamento relativo entre os elementos rolantes e as pistas interna/externa ocorre quando há um aumento na força centrífuga sobre os elementos rolantes. Isso faz com que a força de atrito na pista externa supere a força de tração na pista interna, resultando no deslizamento da gaiola do rolamento e provocando desvios no espectro de frequência dos sinais de vibração (LUO *et al.*, 2022).

Uma das técnicas mais utilizadas na literatura para o monitoramento e diagnóstico de falhas em rolamentos é a análise de envelope, que permite uma distinção mais clara das frequências características de falha no espectro de envelope em relação ao espectro do sinal bruto de vibração (SIEGEL *et al.*, 2014). A análise de envelope consiste nos seguintes passos (RANDALL; ANTONI, 2011):

- 1) Aplicação do filtro passa banda no sinal de vibração, para amplificar o sinal de impacto produzido pelos elementos rolantes;
- 2) Obtenção do envelope do sinal filtrado pela transformada de Hilbert;
- 3) Utilização da FFT no envelope, para converter o sinal do domínio do tempo para o domínio da frequência;

- 4) Análise do espectro do envelope para identificar picos correspondentes às frequências características de falha.

No espectro do envelope, é possível observar amplitudes de frequências próximas às frequências características de falhas e seus harmônicos, conhecidas como *sidebands*. A presença de *sidebands* ocorre devido à periodicidade da carga e à transmissão das vibrações até o sensor (YANG, Y. *et al.*, 2020). O Quadro 2 descreve o comportamento típico dos tipos de falhas em rolamentos observados na análise do espectro do envelope.

Quadro 2 - Comportamento característico de falha em rolamentos na análise de envelope.

Tipo de falha	Descrição
Pista externa	BPFO e harmônicos
Pista interna	BPMFI e harmônicos com <i>sidebands</i> espaçados pela velocidade de rotação e harmônicos da velocidade de rotação
Elementos rolantes	BSF e harmônicos com <i>sidebands</i> espaçados pela FTF, harmônicos da FTF

Fonte: Traduzido de Smith e Randall (2015).

A evolução da falha em rolamentos é geralmente descrita em quatro estágios, os quais podem ser analisados através do espectro de vibração, como ilustrado na Figura 8. No estágio inicial, pequenos defeitos geram sinais na região de alta frequência, sendo uma fase considerada de operação normal. No estágio intermediário, os defeitos começam a produzir impulsos que ativam o sinal na faixa de frequência natural do rolamento. No estágio avançado, as frequências características de falha e seus harmônicos começam a se tornar visíveis no espectro. À medida que o defeito progride, os harmônicos e os *sidebands* aumentam. No estágio crítico, as frequências características do rolamento diminuem em amplitude, e o espectro é afetado por um aumento no ruído e na presença de frequências aleatórias (IUNUSOVA *et al.*, 2024; LEITE *et al.*, 2021).

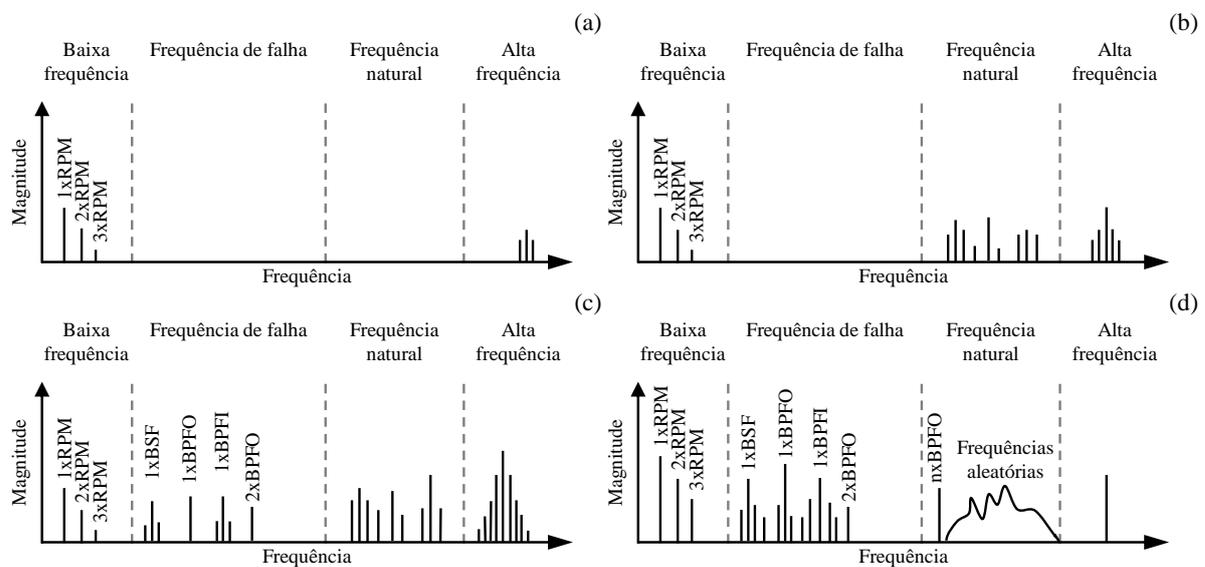
No entanto, os principais desafios na análise de vibração de rolamentos nas TEs são:

- a) As TEs operam sob condições ambientais adversas, onde os sinais característicos dos rolamentos podem ser obstruídos por ruídos adicionais (PENG *et al.*, 2022);
- b) As TEs são compostas por diversos componentes, o que torna a análise de vibração ainda mais desafiadora devido ao cruzamento de frequências e à presença de

componentes de baixa frequência, especialmente na caixa de engrenagens (AZEVEDO; ARAÚJO; BOUCHONNEAU, 2016);

- c) A análise de vibração é geralmente conduzida com base na expertise de um especialista, o que pode ser um desafio, dado o volume e a complexidade de detecção da falha (EL-THALJI; JANTUNEN, 2015; STRÖMBERGSSON; MARKLUND; BERGLUND, 2021).

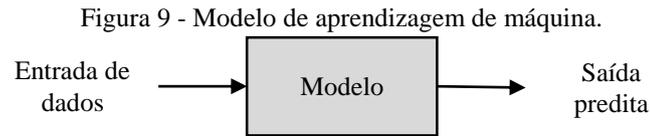
Figura 8 - Estágios de degradação do rolamento no espectro de vibração. (a) Estágio inicial; (b) Estágio intermediário; (c) Estágio avançado; (d) Estágio crítico.



Fonte: Adaptado de Iunusova *et al.* (2024).

3.3 APRENDIZAGEM DE MÁQUINA

Algoritmos de aprendizagem de máquina, ou *machine learning* (ML), são modelos capazes de aprender com informações passadas e reconhecer padrões de relacionamentos presentes nos dados (REZAMAND *et al.*, 2020). Modelos de aprendizagem de máquinas podem ser expressos como uma função $f(x)$ que recebe um conjunto de informações x , chamadas de conjunto de treinamento, para identificar padrões e determinar uma saída predita \tilde{y} , conforme apresentado na Figura 9. Dessa maneira, o modelo treinado é capaz de receber novas informações x e determinar as saídas previstas para os dados que não foram utilizadas no treinamento.



Fonte: O autor (2024).

Modelos de aprendizagem de máquina são técnicas que podem auxiliar em tomadas de decisões confiáveis e repetitivas, principalmente quando a modelagem física se torna difícil, devido à complexidade física do sistema e/ou quando o conjunto de dados apresenta alta dimensionalidade (BENBOUZID *et al.*, 2021; JANIÉSCH; ZSCHECH; HEINRICH, 2021).

Os modelos de aprendizagem de máquina são divididos em três principais categorias:

- a) **Aprendizagem supervisionada:** É a classe de algoritmos em que o modelo é treinado com informações rotuladas. A aprendizagem supervisionada é caracterizada pela ideia do supervisor que fornece uma saída esperada para cada ponto da base de dados e pelo modelo que se ajusta para determinar uma saída. O modelo pode ter os parâmetros ajustados, a fim de diminuir o erro entre a saída prevista e a saída esperada, durante a fase de treinamento (BONACCORSO, 2017);
- b) **Aprendizagem não supervisionada:** É a classe de modelos treinados com o conjunto de dados não rotulados, com o objetivo de encontrar as regularidades e padrões na estrutura dos dados, tal que possa distinguir certos padrões que ocorrem com mais frequência do que outros (ALPAYDIN, 2020). A aprendizagem não supervisionada é recomendada quando é necessário aprender como um conjunto de informações pode ser agrupado, de acordo com sua similaridade (BONACCORSO, 2017);
- c) **Aprendizagem por reforço:** É baseado no conceito de um sinal de reforço ou recompensa, onde o modelo interage com um ambiente e aprende por tentativa e erro. Na aprendizagem por reforço, o objetivo é determinar a ação apropriada, que deve ser aquela que maximiza a recompensa (LADOSZ *et al.*, 2022).

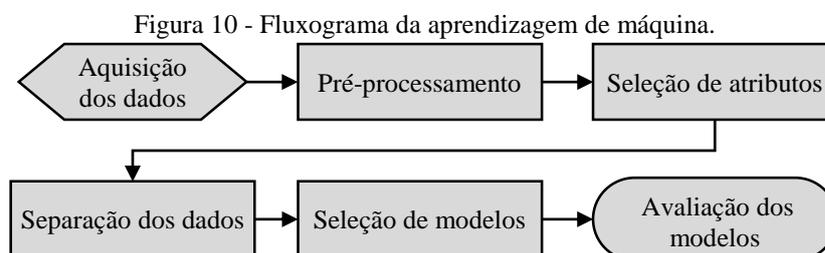
Além disso, modelos de aprendizagem de máquina podem ser classificados nos seguintes tipos:

- a) **Modelos de classificação:** São modelos utilizados quando existe somente um número discreto de resultados possíveis. Modelos de classificação pertencem a classe dos modelos de aprendizagem supervisionada, em que se baseiam em identificar padrões

nas classes predefinidas, com o objetivo de definir a classe pertencente de um novo conjunto de dados (NGUYEN *et al.*, 2021);

- b) Modelos de regressão: São modelos empregados para prever um valor de saída contínuo. Modelos de regressão pertencem a classe dos modelos de aprendizagem supervisionada, com o objetivo de mapear os valores de entrada para o valor de saída predita, construindo uma função $f(x)$ desconhecida que se ajuste de maneira melhor para prever valores futuros (ALPAYDIN, 2020);
- c) Modelos de agrupamento: São modelos da classe não supervisionada. Modelos de agrupamento são capazes de descobrir propriedades e padrões no conjunto de amostras de treinamento não rotuladas para categorizar os dados em grupos (ZHOU, 2021). Normalmente, modelos de agrupamento particionam o conjunto de dados em subgrupos, chamados de *clusters*, onde a atribuição para um determinado *cluster* depende dos atributos do conjunto de dados e/ou da posição de um conjunto de dados em relação à outro (BONACCORSO, 2017).

As etapas da aprendizagem de máquina, apresentadas na Figura 10, podem ser sintetizadas nos seguintes passos: a) aquisição de dados; b) pré-processamento, que consiste na remoção de dados faltosos e *outliers*, extração de atributos e normalização dos dados; c) seleção de atributos, que é composto pela seleção das variáveis com maior importância para determinar a saída predita e/ou a redução da dimensionalidade do conjunto de dados; d) separação dos dados para treinamento e validação; e) seleção dos modelos mais adequados para o treinamento; f) avaliação dos modelos com base em métricas de performance.



Fonte: O autor (2024).

3.3.1 Análise de componentes principais

No treinamento de modelos de ML, a redução de dimensionalidade desempenha um papel crucial na simplificação e otimização do processo de aprendizagem. Em geral, o

desempenho dos modelos melhora quando a dimensionalidade dos dados aumenta, mas quando a dimensionalidade continua a aumentar, o desempenho pode ser prejudicado pela redução da capacidade de generalização do modelo (JIA *et al.*, 2022).

A redução de dimensionalidade visa transferir o conjunto de dados originais para um espaço de menor dimensão, preservando ao máximo possível as características essenciais enquanto elimina os atributos redundantes, ruídos ou dados irrelevantes (HUANG; WU; YE, 2019).

As técnicas de redução de dimensionalidade podem ser conduzidas através da seleção de atributos e/ou da transformação dos mesmos. A seleção de atributos envolve o processo de identificar e escolher as características mais relevantes para o problema em questão. Por outro lado, a transformação de atributos utiliza operações matemáticas para diminuir o espaço dimensional (AYESHA; HANIF; TALIB, 2020).

A análise de componentes principais (PCA) é um método da redução de dimensionalidade que busca identificar novas variáveis não correlacionadas, chamadas de componentes principais, pela combinação linear das variáveis originais. Essa abordagem visa capturar as direções ao longo das quais a variação nos dados é máxima (RINGNÉR, 2008).

Conforme destacado por Jolliffe e Cadima (2016) e Greenacre *et al.* (2022), os procedimentos para a determinação das componentes principais para um conjunto de dados $\mathcal{D} = \{x^i\}_{i=1}^p = X_{n \times p}$, com n amostras e p variáveis, compreendem:

- 1) Normalizar os dados em relação à média \bar{x}^i igual a zero e o desvio padrão σ_{x^i} igual a 1;
- 2) Determinar a matriz de covariância com o conjunto de dados normalizados. Onde a matriz de covariância S é uma matriz simétrica de dimensão $p \times p$, onde cada elemento de S representa a covariância entre as variáveis;
- 3) Calcular os autovalores λ_i e os autovetores v_i da matriz S e formar a matriz de autovetores $V_{p \times p}$ ordenados de forma decrescente com base nos autovalores. Os autovalores representam a variância explicada por cada componente principal, sendo a soma de todos os autovalores equivalente à variância total dos dados;
- 4) Determinar o número k de componentes principais que atinja o limite t da razão entre a variância explicada acumulada das k componentes e a variância total das p componentes;

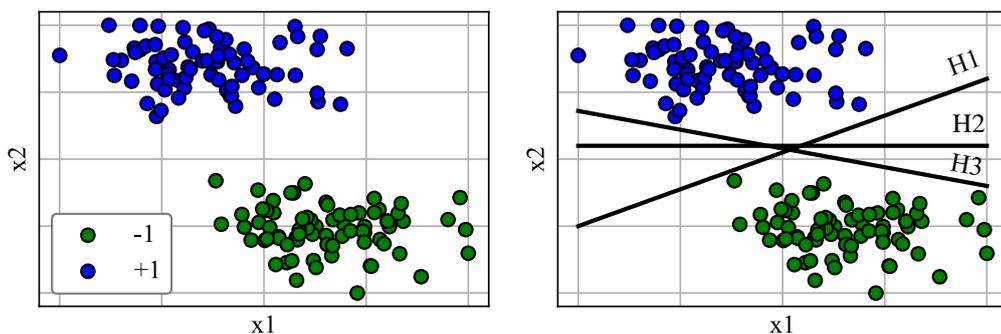
- 5) Definir o novo conjunto de dados $X_{n \times k} = X_{n \times p} V_{p \times k}$ formado pelas k componentes principais.

Embora as novas k variáveis sejam funções lineares de todas as p variáveis do conjunto original, a PCA enfrenta a desvantagem da perda de interpretabilidade. À medida que os dados são projetados em um novo espaço de menor dimensão, a relação entre os componentes principais e as características originais podem torna-se menos intuitiva.

3.3.2 Support vector machine

Support vector machine (SVM) é uma das técnicas de aprendizado de máquina supervisionada mais utilizadas, que pode ser aplicada em tarefas de classificação ou regressão, mas é amplamente utilizado em problemas de classificação, conforme apresenta a Figura 11a (LEI *et al.*, 2020). O objetivo do classificador SVM é de criar e ajustar o melhor hiperplano de separação de dimensão $n - 1$, para o conjunto de dados de dimensão n , de modo que possa isolar o espaço dos atributos em classes (REBALA; RAVI; CHURIWALA, 2019). A Figura 11b apresenta possíveis hiperplanos de separação para o caso de um classificador linear.

Figura 11 - Classificação linear pela técnica de SVM. (a) Conjunto de dados; (b) Hiperplanos de separação.



Fonte: O autor (2024).

Modelos SVMs podem ser classificados em duas categorias (KURANI *et al.*, 2023):

- SVM linear: Quando o conjunto de dados pode ser linearmente separado, que implica que o conjunto de dados pode ser classificado por meio de uma única linha reta;
- SVM não-linear: Quando o conjunto de dados não admite uma separação linear, ou seja, a classificação não pode ser feita por meio de uma linha reta. Nesse caso, o conjunto de

dados é transformado para um espaço de dimensão superior, de modo que possa ser separado por um hiperplano de dimensão $n - 1$.

Para o caso do SVM linear, onde $\mathcal{D} = \{\vec{x}_i, y_i\}_{i=1}^n$ é o conjunto de dados com $y_i \in \{+1, -1\}$ e $\vec{x}_i \in \mathbb{R}^d$, o hiperplano de separação pode ser definido pela Equação 7, em que $\vec{w}^T = [w_1 \ w_2 \ \dots \ w_d] \in \mathbb{R}^d$ é o vetor normal que controla a direção do hiperplano e $b \in \mathbb{R}$ é o parâmetro que controla a distância entre o hiperplano e a origem por $b/\|\vec{w}\|$ (ZHOU, 2021).

$$\vec{w}^T \cdot \vec{x} + b = 0 \quad (7)$$

O hiperplano definido pela Equação 7 é posicionado de tal forma que a distância entre os pontos mais próximos das diferentes classes seja maximizada. Desse modo, supondo que todos os dados de treinamento estejam restritos pela Equação 8, o conjunto de dados pode ser linearmente separado pela margem $2/\|\vec{w}\|$, conforme o hiperplano plano descrito pela Equação 9 (BURGES, 1998).

$$\begin{cases} \vec{w}^T \cdot \vec{x}_i + b \geq +1, & y_i = +1 \\ \vec{w}^T \cdot \vec{x}_i + b \leq -1, & y_i = -1 \end{cases} \quad (8)$$

$$y_i(\vec{w}^T \cdot \vec{x}_i + b) - 1 \geq 0, \forall i \text{ e } y_i \in \{+1, -1\} \quad (9)$$

Portanto, o objetivo do SVM é a maximização da região de separação dos dados em relação ao hiperplano $\vec{w}^T \cdot \vec{x} + b = 0$, obtida pela minimização de $\frac{1}{2} \|\vec{w}\|^2$ e restrita pela Equação 9. A solução do problema do SVM linear é direcionada pela forma dual, alcançada pela função de Lagrange, definida pela Equação 10, onde α são os multiplicadores de Lagrange (ZHOU, 2021).

$$L(\vec{w}, b, \alpha) = \frac{1}{2} \|\vec{w}\|^2 + \sum_{i=1}^n \alpha_i (1 - y_i(\vec{w}^T \cdot \vec{x}_i + b)) \quad (10)$$

O problema de otimização expresso pela Equação 10 pode ser reformulado pela derivada parcial em relação \vec{w} e b , conforme apresentado na Equação 11. Através desse processo, o problema de maximização é transformado na Equação 12, denominada de forma dual.

$$\frac{\partial L}{\partial \vec{w}} = 0; \frac{\partial L}{\partial b} = 0 \quad (11)$$

$$L = \sum_i \alpha_i + \sum_{i,j} \alpha_i \alpha_j y_i y_j (\vec{x}_i \cdot \vec{x}_j); \begin{cases} \alpha_i \geq 0, \forall i \\ \sum_{i=1}^n \alpha_i y_i = 0 \end{cases} \quad (12)$$

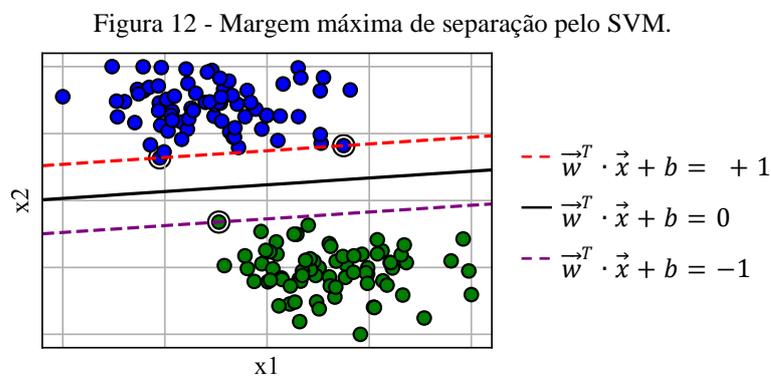
Ao solucionar o problema de maximização da Equação 12, a função que define o hiperplano com margem máxima é dada pela Equação 13. Os multiplicadores de Lagrange são determinados pelas condições de Karush-Kuhn-Tucker (KKT) $\alpha_i \geq 0, \forall i, \sum_{i=1}^n \alpha_i y_i = 0$ e $\vec{w} - \sum_{i=1}^n \alpha_i y_i \vec{x}_i = 0$ (NALEPA; KAWULOK, 2019).

$$f(x) = \sum_{i=1}^n \alpha_i y_i \vec{x}_i^T \cdot \vec{x} + b \quad (13)$$

De acordo com Zhou (2021), o parâmetro b pode ser definido no caso de existir $y_s f(x_s) = 1$ para cada vetor de suporte (x_s, y_s) , conforme a Equação 14. Na Equação 14, n_{sv} denota o número de vetores de suporte, definido por $n_{sv} = \{i \mid \alpha_i > 0, i = 1, 2, \dots, n\}$, e vs representa o conjunto de vetores de suporte.

$$b = \frac{1}{n_{sv}} \sum_{x_s \in vs} \left(\frac{1}{y_s} - \sum_{x_i \in vs} \alpha_i y_i \vec{x}_i^T \cdot \vec{x}_s \right) \quad (14)$$

A Figura 12 ilustra a margem máxima de separação para o problema de classificação da Figura 11a, em que os pontos circulos são denominados como vetores de suporte e a margem ótima é definida como a distância entre o hiperplano ideal e os pontos de dados mais próximos de cada classe.



Fonte: O autor (2024).

Uma extensão do modelo clássico de SVM, definida para lidar com casos em que o conjunto de dados não é linearmente separável, é o SVM com margens suaves. Cortes e Vapnik (1995) introduziram o SVM com margens suaves, que define a margem de separação por um número mínimo de erro. Ou seja, permite que alguns pontos de dados sejam classificados incorretamente, mas que resulte em uma margem mais ampla e generalizada (CHAUHAN; DAHIYA; SHARMA, 2019).

O problema de otimização do modelo clássico de SVM pode ser reformulado de forma que as restrições impostas na Equação 8 possam ser reescritas pela Equação 15, no contexto do algoritmo de SVM com margens suaves (NALEPA; KAWULOK, 2019). Essa reformulação introduz a variável de folga positiva ξ_i , definida por $\xi_i \geq 0 \forall i$.

$$\begin{cases} \vec{w}^T \cdot \vec{x}_i + b \geq +1 - \xi_i, & y_i = +1 \\ \vec{w}^T \cdot \vec{x}_i + b \leq -1 + \xi_i, & y_i = -1 \end{cases} \quad (15)$$

Dessa maneira, o algoritmo de SVM com margens suaves busca encontrar o hiperplano de separação que maximiza a margem entre as classes pela função objetivo, definida pela Equação 16 (CERVANTES *et al.*, 2020). Onde C é um termo de regularização que controla o equilíbrio entre maximizar a margem e minimizar o erro de classificação.

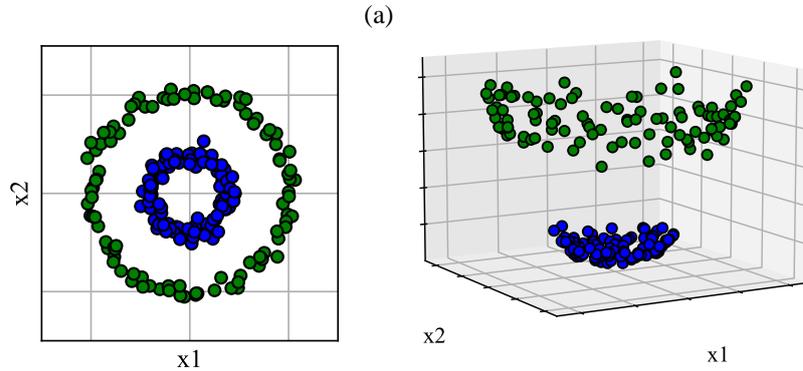
$$\min_{\vec{w}, b, \xi_i} \frac{\|\vec{w}\|^2}{2} + C \sum_{i=1}^n \xi_i; \begin{cases} \xi_i \geq 0, \forall i \\ y_i(\vec{w}^T \cdot \vec{x}_i + b) \geq 1 - \xi_i \end{cases} \quad (16)$$

Conforme já mencionado, quando a separação dos dados por um hiperplano é inviável, o SVM não linear mapeia os dados para um novo espaço de maior dimensão por $\phi(\vec{x})$, denominado de espaço de características. Nesse novo espaço, um SVM linear, com o hiperplano definido pela Equação 17, pode ser utilizado para separar o conjunto de dados de forma eficiente.

$$f(\vec{x}) = \vec{w}^T \cdot \phi(\vec{x}) + b \quad (17)$$

Um exemplo do SVM não linear é ilustrado na Figura 13. Onde o conjunto de dados da Figura 13a com 2-dimensões é mapeado para um espaço de maior dimensão, conforme ilustrado na Figura 13b, de forma que possa ser definido um hiperplano de separação entre as classes, e possa ser utilizado para a tarefa de classificação.

Figura 13 - SVM para classificação de dados não separáveis linearmente. (a) Baixa dimensão; (b) Alta dimensão.



Fonte: O autor (2024).

Para superar a dificuldade de determinar a função $\phi(\vec{x})$ que mapeia o conjunto de dados para o espaço de características de alta dimensão, o método de SVM não linear utiliza do truque de kernel para determinar a similaridade entre os pontos do conjunto de dados, sem determinar o mapeamento dos dados explicitamente (ZHOU, 2021).

A Equação 18 define a função kernel, que determina o produto interno entre dois vetores de características, sem avaliar explicitamente o novo espaço de maior dimensão. E a Equação 19 especifica o hiperplano com margem máxima de separação para o problema de separação não linear.

$$k(\vec{x}_i, \vec{x}_j) = \langle \phi(\vec{x}_i), \phi(\vec{x}_j) \rangle = \phi(\vec{x}_i)^T \phi(\vec{x}_j) \quad (18)$$

$$f(x) = \sum_{i=1}^n \alpha_i y_i k(\vec{x}, \vec{x}_i) + b \quad (19)$$

No algoritmo de SVM, as funções de kernel são utilizadas para definir o grau de similaridade entre cada par de pontos do conjunto de dados originais. Essa informação é usada para construir a matriz de kernel, a partir da qual o hiperplano de separação é definido. O Quadro 3 apresenta as funções de kernel comuns aplicados na técnica de SVM.

Quadro 3 - Funções de kernel comuns.

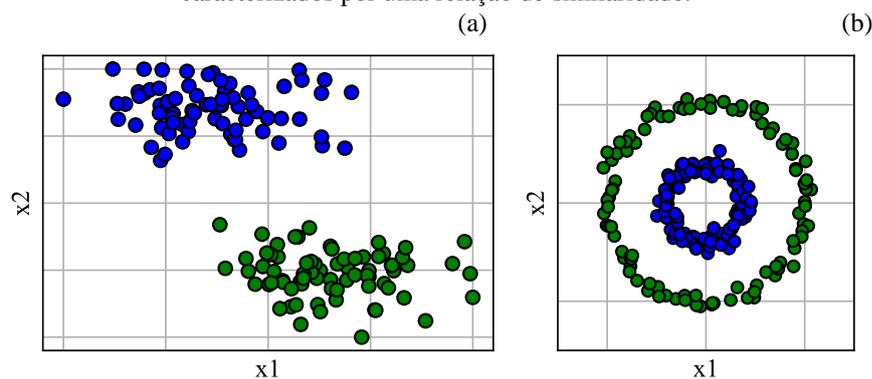
Nome	Expressão
Linear	$k(\vec{x}_i, \vec{x}_j) = \vec{x}_i^T \vec{x}_j$
Polinomial	$k(\vec{x}_i, \vec{x}_j) = (\vec{x}_i^T \vec{x}_j)^d$
Sigmoidal	$k(\vec{x}_i, \vec{x}_j) = (\beta \vec{x}_i^T \vec{x}_j + \theta), \beta > 0$ $\text{e } \theta < 0$
Função de base radial (RBF)	$k(\vec{x}_i, \vec{x}_j) = \left(-\frac{\ \vec{x}_i - \vec{x}_j\ ^2}{2\sigma^2}\right)$

Fonte: Adaptado de Zhou (2021).

3.3.3 Agrupamento espectral

O algoritmo de agrupamento espectral é uma abordagem da aprendizagem de máquina não supervisionada aplicada para o agrupamento dos dados em subconjuntos, chamados de *clusters*. Diferentemente de outros modelos de agrupamento, que se baseiam na compacidade espacial entre os dados, o agrupamento espectral é baseado no grau de conectividade ou similaridade entre cada ponto do conjunto de dados. A Figura 14 apresenta a diferença entre conjuntos de dados compactos e conjuntos de dados caracterizados por uma relação de similaridade.

Figura 14 - Comparação entre conjuntos de dados para agrupamento. (a) *clusters* bem agrupados; (b) *clusters* caracterizados por uma relação de similaridade.



Fonte: O autor (2024).

O agrupamento espectral utiliza o formalismo da teoria de grafos $G = (V, E)$, onde $V = \{v_1, v_2, \dots, v_n\}$ é o conjunto de vértices que representa um ponto x_i do conjunto de dados, e E define o conjunto de arestas. Caso dois vértices x_i e x_j , sejam conectados, um peso s_{ij} é atribuído para a aresta. Dois vértices são conectados quando a similaridade s_{ij} é positiva ou

maior do que um valor limite, e se $s_{ij} = 0$ significa que x_i e x_j não são conectados por uma aresta.

Conforme Von Luxburg (2007) o agrupamento espectral pode ser formulado em:

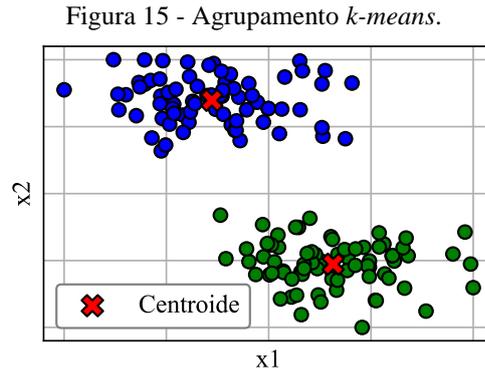
- a) As arestas que conectam grupos distintos devem ter pesos extremamente baixos;
- b) As arestas que conectam pontos dentro do mesmo *cluster* devem ter pesos altos, ou seja, os pontos dentro de um mesmo grupo são altamente similares entre si.

A aplicação da técnica de agrupamento espectral envolve seguir os passos descritos por Ng, Jordan e Weiss (2001):

- 1) Dado o conjunto de dados $\mathcal{D} = \{x_i\}_{i=1}^n$, construa a matriz de afinidade $A \in \mathbb{R}^{n \times n}$ definida por $a_{ij} = S(x_i, x_j)$. Onde $S(x_i, x_j)$ é uma função de similaridade;
- 2) Determine a matriz Laplaciana normalizada $\mathcal{L} = D^{-1/2} L D^{-1/2}$. Onde $L = D - A$ é a matriz laplaciana e D é a matriz de grau, definida pela matriz diagonal formada por $d_i = \sum_{j=0}^n a_{ij}$. E calcule os primeiros k autovetores u_1, u_2, \dots, u_k da matriz \mathcal{L} ;
- 3) Construa a matriz $U \in \mathbb{R}^{n \times k}$, com os autovetores u_1, u_2, \dots, u_k como colunas;
- 4) Determine a matriz $Y \in \mathbb{R}^{n \times k}$, com a normalização das linhas da matriz U à norma 1 por $y_{ij} = u_{ij} / (\sum_j u_{ij}^2)^{1/2}$;
- 5) Por fim, considere cada linha de Y como um ponto no \mathbb{R}^k , e agrupe em k *clusters* pelo método de *k-means*.

O princípio fundamental do agrupamento espectral baseia-se na propriedade da matriz \mathcal{L} , onde o número de autovalores com valor zero corresponde ao número de *clusters* distintos. A matriz \mathcal{L} é simétrica e positiva definida, com m autovalores $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_m$ e $\lambda_1 = 0$ definido como o menor autovalor de \mathcal{L} (VON LUXBURG, 2007; WIERZCHOŃ *et al.*, 2018)

O algoritmo de *k-means* é uma técnica de aprendizagem de máquina não supervisionada para o agrupamento de dados em um número de centroides predefinidos. A Figura 15 apresenta o agrupamento dos dados da Figura 14a pelo método de *k-means*.



Fonte: O autor (2024).

A técnica de agrupamento *k-means* consiste em dividir o conjunto de dados em *cluster* associados a um centroide, inicializado de maneira aleatória. Para cada ponto do conjunto de dados é avaliada a distância entre o centroide e o ponto, sendo cada ponto associado ao cluster mais próximo. O objetivo do *k-means* é de minimizar a soma das distâncias quadradas entre cada ponto e o centroide do seu *cluster* (IKOTUN *et al.*, 2023).

A Figura 16 demonstra o agrupamento do conjunto de dados pelo pseudocódigo do modelo *k-means*.

Figura 16 - Pseudocódigo *k-means*.

Pseudocódigo *k-means*

Entrada: Conjunto de dados $\mathcal{D} = \{x_i\}_{i=1}^n$ e o número de *clusters* k , com $k \geq 2$

Saída: Subconjuntos de dados C_1, \dots, C_k

- 1: Inicialize os *clusters* c_1, \dots, c_k aleatoriamente
 - 2: **Repetir até o número máximo de iterações ou até a tolerância**
 - 3: **Para cada** x em $\mathcal{D} = \{x_i\}_{i=1}^n$ **faça**
 - 4: **Para todo** centroide c em $\{c_1, \dots, c_k\}$ **faça**
 - 5: **Se** $\|x - c_i\| < \|x - c_j\|$ com $i \neq j, j \leq k$ e $1 \leq i$
 - 6: **Defina** x para C_i
 - 7: Calcule a posição do centroide como:
 - 8:
$$c_i = \frac{1}{n_i} \sum_{x \in C_i} x$$
 - 9: **Fim**
-

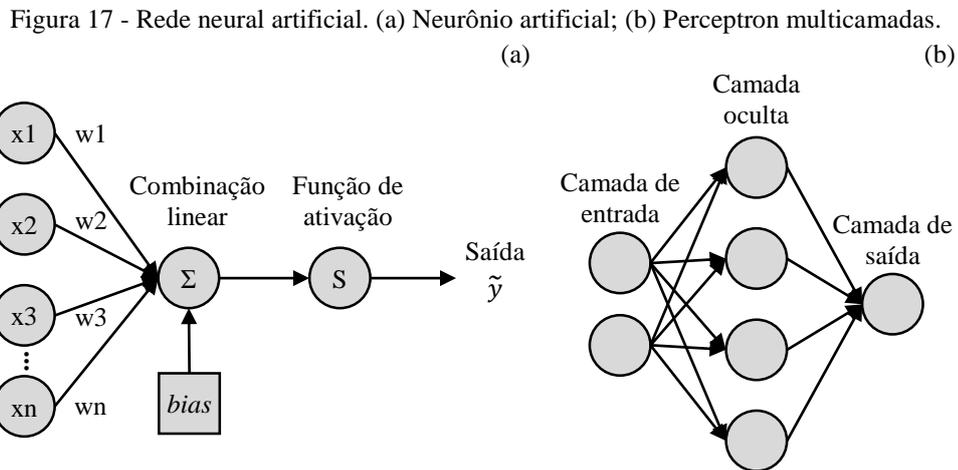
Fonte: O autor (2024).

3.3.4 Redes neurais artificiais

As redes neurais artificiais (RNAs) são modelos de aprendizagem de máquina inspirados no sistema nervoso biológico, principalmente no cérebro humano, desenvolvido para

imitar como este sistema processa os dados. De acordo com Abiodun *et al.* (2018), entre as aplicações das RNAs, destacam-se: reconhecimento de imagens, processamento de linguagem natural, otimização, classificação e previsão de séries temporais.

Estes modelos são constituídos por várias unidades de processamento interconectados entre si, chamados de neurônios artificiais. O modelo do neurônio artificial foi estabelecido por McCulloch e Pitts (1943), apresentado na Figura 17a, em que recebe um conjunto de parâmetros de entrada $x = [x_1, x_2, x_3, \dots, x_n]$ e associa-os a um peso correspondente w_i , por combinação linear. Além disso, um termo adicional pode ser considerado no neurônio artificial para garantir a generalização da RNA, chamado de *bias*. A saída do neurônio artificial \tilde{y} é determinada pela aplicação de uma função de ativação no resultado da combinação linear.



Fonte: O autor (2024).

Baseado nisso, Rosenblatt (1958) propôs o Perceptron, um sistema para classificação binária. No modelo Perceptron, o resultado da combinação das entradas e pesos é comparado ao um valor limite através da função de ativação, conforme apresenta a Equação 20. Se o resultado for superior ao valor limite, a saída do Perceptron é 1, caso contrário, a saída é 0.

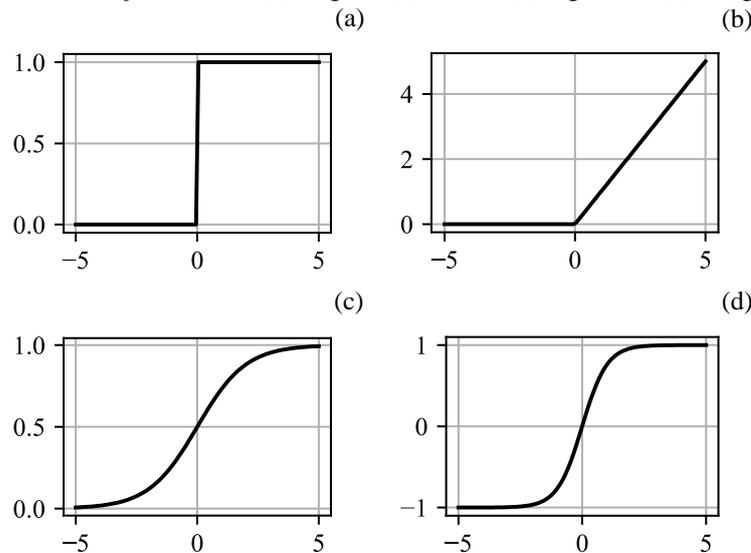
$$\tilde{y} = \begin{cases} 1, & \text{se } \sum_{i=1}^n w_i x_i + bias \geq limiar \\ 0, & \text{se } \sum_{i=1}^n w_i x_i + bias < limiar \end{cases} \quad (20)$$

O modelo de RNA do tipo Perceptron multicamadas (MLP), apresentado na Figura 17b, é constituído por camadas conectadas, denominadas por: camada de entrada, uma ou mais camadas ocultas e a camada de saída. No modelo MLP, as camadas adicionais são responsáveis

por garantir que RNAs possam lidar com problemas não lineares (ALPAYDIN, 2020; JANIESCH; ZSCHECH; HEINRICH, 2021).

Em geral, dependendo do objetivo da RNA, as funções de ativação podem assumir outros formatos para determinar a ativação do neurônio e sua intensidade. As funções de ativação, apresentadas pela Figura 18, mais utilizadas são: degrau, ReLU, sigmoide e tangente hiperbólica (APICELLA *et al.*, 2021).

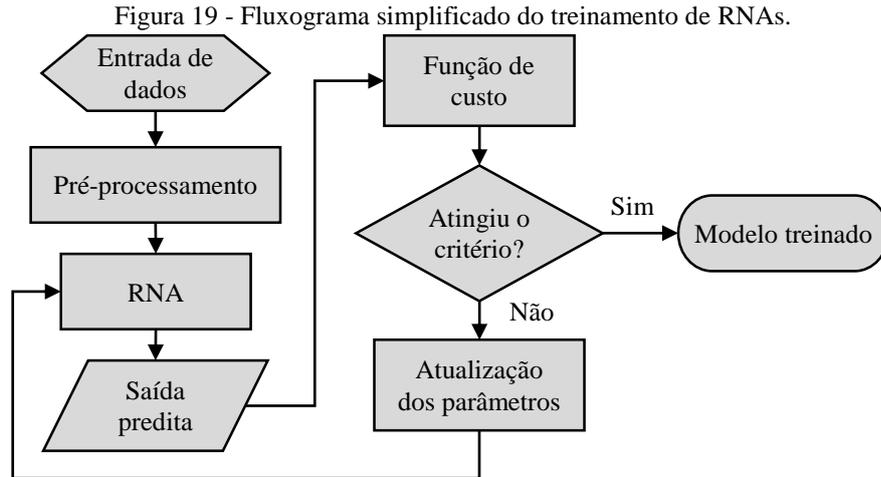
Figura 18 - Funções de ativação comuns. (a) Degrau; (b) ReLU; (c) Sigmoide; (d) Tangente hiperbólica.



Fonte: O autor (2024).

O treinamento das RNAs é um processo que envolve a otimização dos pesos e parâmetros para minimizar o erro entre a saída observada e a saída prevista. O método mais utilizado na aprendizagem de máquina para atualizações dos parâmetros, durante o treinamento, é o algoritmo de retropropagação, ou *backpropagation* (LILLICRAP *et al.*, 2020).

A Figura 19 apresenta o fluxograma para treinamento da RNA pelo método de *backpropagation*. No algoritmo de *backpropagation*, o erro entre a saída prevista e a saída observada é determinado por uma função de custo. Por meio da regra da cadeia do cálculo diferencial, o erro é propagado da camada de saída em direção à camada de entrada. Durante esse processo, os pesos das conexões entre os neurônios são ajustados gradativamente usando uma técnica de otimização.



Fonte: O autor (2024).

Um dos métodos de otimização amplamente utilizado em aprendizagem de máquina é o gradiente descendente (GD), definido pela Equação 21. O método do GD é um algoritmo iterativo para encontrar o mínimo de uma função, ajustando os parâmetros da RNA na direção oposta ao gradiente da função de perda, por meio da taxa de aprendizagem α .

$$w_{i,j+1} = w_{i,j} - \alpha \frac{\partial}{\partial w_i} l(w_{i,j}) \quad (21)$$

Onde w corresponde ao parâmetro atualizável, $l(w)$ é a função de custo, $\frac{\partial}{\partial w_i} l(w_i)$ é a derivada parcial da função de custo em relação ao parâmetro w_i , i corresponde ao i -ésimo parâmetro e j está relacionada com a j -ésima iteração.

As funções de custo, ou funções de perda, desempenham papel fundamental no treinamento das RNAs e na melhoria de sua performance. As funções de custo podem ser compreendidas como a distância entre o valor predito e o valor observado. Assim, a seleção de uma função de custo inadequada afetará de certo ponto a eficácia do algoritmo (WANG, Q. *et al.*, 2020). As funções de custo são classificadas de acordo com a aplicação do modelo de aprendizagem de máquina. O Quadro 4 apresenta alguns exemplos de funções de custos para cada tipo aplicação das RNAs.

Quadro 4 - Exemplos de funções de custo.

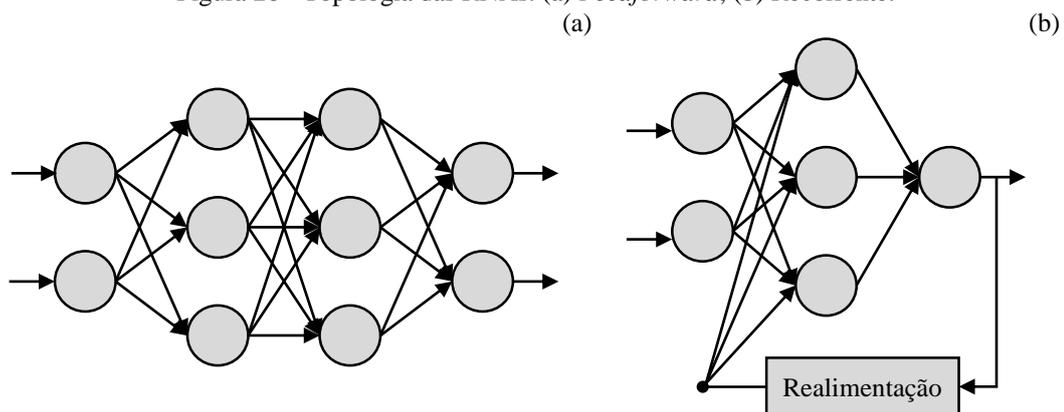
Nome	Aplicação	Expressão
Erro absoluto	Regressão	$l(y, \tilde{y}) = y_i - \tilde{y}_i $
Erro quadrático		$l(y, \tilde{y}) = (y_i - \tilde{y}_i)^2$
Erro médio quadrático		$l(y, \tilde{y}) = \frac{1}{N} \sum_{i=0}^N (y_i - \tilde{y}_i)^2$
Entropia cruzada	Classificação	$l(y, \tilde{y}) = -\frac{1}{N} \sum_{i=0}^N [y_i \log(\tilde{y}_i) + (1 - y_i) \log(1 - \tilde{y}_i)]$
Entropia cruzada sigmoide		$l(y, \tilde{y}) = -[y_i \log(\tilde{y}_i) + (1 - y_i) \log(1 - \tilde{y}_i)]$

Fonte: Adaptado de WANG *et al.* (2020).

Ainda no contexto do treinamento de RNAs, o número de vezes do ciclo de treinamento é denominado de época. Uma época corresponde a uma única passagem completa por todo o conjunto de dados de treinamento ao longo do treinamento do modelo. Durante uma época, os dados de treinamento podem ser divididos em subconjunto, chamados de *batches*. O tamanho do *batch* determina quantos exemplos de treinamento são processados simultaneamente antes de atualizar os pesos da RNA.

Além disso, as RNAs podem ser categorizadas em dois principais tipos de topologias: *feedforward* e recorrente. A Figura 20 mostra a representação esquemática dessas topologias.

Figura 20 - Topologia das RNAs. (a) *Feedforward*; (b) Recorrente.



Fonte: O autor (2024).

Em redes *feedforward*, a informação flui em uma única direção, desde a camada de entrada até a camada de saída. Ou seja, o resultado de uma camada é usado como entrada para a próxima. Em contraste, as redes recorrentes RNN são caracterizadas por um comportamento

dinâmico. As RNNs permitem que o resultado de uma camada pode ser usado como entrada para si mesma ou para outras camadas anteriores (BREZAK *et al.*, 2012).

3.3.5 Métricas de performance

As métricas de performance são medidas quantitativas que avaliam o desempenho dos modelos de ML e possibilitam a comparação entre diferentes modelos, a avaliação da generalidade para novos conjuntos de dados e o monitoramento da evolução do desempenho ao longo do tempo.

Para problemas de classificação binária, as métricas de performance são derivadas da matriz de confusão. Essa matriz, ilustrada na Figura 21, define a frequência de previsão para cada classe do modelo em relação às classes reais dos dados. Pela matriz de confusão, verdadeiro positivo (VP) e verdadeiro negativo (VN) correspondem ao número de instâncias que foram corretamente classificadas. Enquanto, falso positivo (FP) e falso negativo (FN) equivalem ao número de instâncias que foram classificadas incorretamente (HOSSIN; SULAIMAN, 2015).

Figura 21 - Representação da matriz de confusão.

		Classe predita	
		Normal	Falha
Classe real	Normal	VP	FN
	Falha	FP	VN

Fonte: O autor (2024).

As métricas baseadas na matriz de confusão são definidas por:

- a) Acurácia: Define a porcentagem das previsões corretas em relação ao número total de amostras, definida pela Equação 22;
- b) Acurácia balanceada: Reflete o grau de precisão das previsões corretas, sendo especialmente útil em problemas de classificação com classes desbalanceadas. A acurácia balanceada é calculada conforme descrito na Equação 23;

- c) *Precisão*: Indica a proporção de previsões positivas feitas pelo modelo que foram corretamente classificadas em relação ao total de previsões da classe positiva, conforme a Equação 24;
- d) *Especificidade*: Mede a fração das previsões negativas que foram corretamente classificadas, conforme estabelecido pela Equação 25;
- e) *Recall*: Determina a fração das previsões positivas que foram corretamente classificadas, como definido pela Equação 26;
- f) *F1-score*: Representa uma média harmônica entre as métricas de precisão e *recall*, definida pela Equação 27;

$$Acurácia = \frac{VP+VN}{VP+VN+FP+FN} \quad (22)$$

$$Acurácia\ balanceada = \frac{1}{2} \left(\frac{VP}{VP+FN} + \frac{VN}{VN+FP} \right) \quad (23)$$

$$Precisão = \frac{VP}{VP+FP} \quad (24)$$

$$Especificidade = \frac{VN}{VN+FP} \quad (25)$$

$$Recall = \frac{VP}{VP+FN} \quad (26)$$

$$f1_{score} = \frac{2 \times Precisão \times Recall}{Precisão + Recall} \quad (27)$$

Para problemas de agrupamento, é comum avaliar a eficácia em relação às classes reais utilizando a média de acurácia, calculada através da informação mútua normalizada (IMN). Essa métrica quantifica a capacidade do algoritmo em concordar com as classes reais, oferecendo uma medida confiável da qualidade dos agrupamentos gerados (NEWMAN; CANTWELL; YOUNG, 2020). A IMN é descrita por meio da teoria da informação, pelos conceitos de entropia de Shannon e informação mútua.

A entropia de Shannon (H) é uma medida de incerteza atribuída a uma distribuição de probabilidade. Em situações onde a incerteza é maior, a entropia é maior, enquanto em cenários mais previsíveis, a entropia possui valores mais baixos (TSALLIS, 2022). Para dois *clusters* N e F , a entropia de Shannon é determinada pela Equação 28, onde $P(i)$ e $P(j)$ correspondem às probabilidades de um elemento escolhido aleatoriamente cair na classe N e F , respectivamente, e M corresponde ao número de elementos (KNOPS *et al.*, 2006).

$$\begin{aligned}
H(N) &= - \sum_{i=1}^M P(i) \log(P(i)) \\
H(F) &= - \sum_{j=1}^M P(j) \log(P(j)) \\
P(i) &= \frac{M_i}{M}; P(j) = \frac{M_j}{M}
\end{aligned} \tag{28}$$

A informação mútua (IM), descrita pela Equação 29, entre N e F pode ser compreendida como uma medida de similaridade que é maximizada quando as rotulagens são iguais. Para o caso dos rótulos não correlacionados, a IM é zero, enquanto para rótulos idênticos, a IM é igual a $H(N) = H(F)$ (NEWMAN; CANTWELL; YOUNG, 2020).

$$IM(N, F) = \sum_{i=1}^{|N|} \sum_{j=1}^{|F|} P(i, j) \log \left(\frac{P(i, j)}{P(i)P^*(j)} \right) \tag{29}$$

Dessa maneira, por conveniência, a IM é normalizada resultando em uma métrica com valores compreendidos entre 0 e 1, conforme descrito na Equação 30.

$$IMN(N, F) = \frac{IM(N, F)}{\text{média}(H(N), H(F))} \tag{30}$$

3.4 COMPUTAÇÃO QUÂNTICA

A computação quântica é o campo da ciência da computação que utiliza os fundamentos da mecânica quântica para o processamento de informações. Em contraste com a computação clássica, onde a unidade de informação mais básica é representada pelo *bit*, que opera em dois estados possíveis 0 ou 1, a computação quântica opera com *qubits*.

O *qubit*, a unidade fundamental da computação quântica, representa o estado quântico de um sistema. Ao contrário do *bit*, o *qubit* não é restrito a uma representação binária, mas pode estar em uma combinação linear dos estados 0 e 1 (MCMAHON, 2007).

A Equação 31 define matematicamente o estado de um *qubit* pela notação de Dirac, chamada de *ket*. O *qubit* $|\psi\rangle$ é descrito por um vetor complexo bidimensional dentro do espaço de Hilbert \mathbb{C}^2 . As amplitudes complexas α e β , definem a probabilidade de ocorrência do estado quântico e obedecem a relação $|\alpha|^2 + |\beta|^2 = 1$ (MCMAHON, 2007; RIEFFEL; POLAK, 2011).

$$|\psi\rangle = [\alpha \ \beta]^T; \{\alpha, \beta\} \in \mathbb{C} \tag{31}$$

Na computação quântica, o *qubit* é decomposto em relação aos vetores bases $|0\rangle$ e $|1\rangle$ do \mathbb{C}^2 , conforme a Equação 32 (MCMAHON, 2007). Após a medição, o *qubit* assume o estado $|0\rangle$ com probabilidade $|\alpha|^2$ ou o estado $|1\rangle$ com probabilidade $|\beta|^2$.

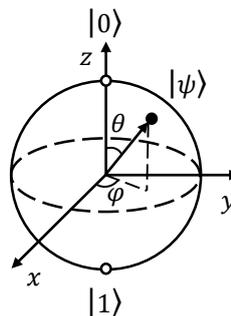
$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle; |0\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; |1\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (32)$$

Conforme apresentado por Yanofsky e Mannucci (2008), a Equação 32 pode ser reescrita para a forma polar, dependendo somente de dois parâmetros reais: θ e φ , com $\theta \in [0, \pi]$ e $\varphi \in [0, 2\pi)$. O estado do *qubit* na forma polar parametrizada é descrito pela Equação 33.

$$|\psi\rangle = \cos(\theta)|0\rangle + e^{i\varphi}\sin(\theta)|1\rangle \quad (33)$$

O estado de um *qubit* é representado graficamente por um vetor \mathbb{R}^3 que aponta da origem até a superfície de uma esfera tridimensional de raio unitário, chamada de *Bloch sphere*. A Figura 22 apresenta a *Bloch sphere*, onde cada ponto da superfície da esfera corresponde a um estado quântico possível do *qubit*. A parte norte da *Bloch sphere* corresponde ao estado $|0\rangle$ e a parte sul corresponde ao estado $|1\rangle$ (YANOFSKY; MANNUCCI, 2008).

Figura 22 - Representação geométrica de um *qubit*.



Fonte: O autor (2024).

Na mecânica quântica, a notação *bra-ket* é a forma padrão utilizada para representar um estado quântico e as operações realizadas na transformação do sistema quântico, onde $|\cdot\rangle$ corresponde ao *ket* e $\langle\cdot|$ é chamado de *bra*. Além disso, duas operações fundamentais são utilizadas na mecânica quântica, o produto interno e o produto tensorial (SCHULD; PETRUCCIONE, F., 2018).

O produto interno combina dois estados quânticos do \mathbb{C}^n em um único número complexo, sendo utilizado na mecânica quântica para determinar a probabilidade de encontrar um sistema em um determinado estado. O produto interno entre os estados quânticos $|\psi_1\rangle = \alpha_1|0\rangle + \beta_1|1\rangle$ e $|\psi_2\rangle = \alpha_2|0\rangle + \beta_2|1\rangle$ é definido pela Equação 34 (MCMAHON, 2007; SCHULD; PETRUCCIONE, F., 2018).

$$\langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_1\rangle^* = [\alpha_1^* \ \beta_1^*] \begin{bmatrix} \alpha_2 \\ \beta_2 \end{bmatrix} = \alpha_1^*\alpha_2 + \beta_1^*\beta_2 \quad (34)$$

Um estado quântico $|\psi_1\rangle$ é considerado normalizado, quando o produto interno entre ele mesmo é igual a um. E dois estados quânticos $|\psi_1\rangle$ e $|\psi_2\rangle$, são considerados ortogonais quando o produto interno entre eles é zero e ortonormais quando cada estado quântico é normalizado e os estados são ortogonais (MCMAHON, 2007). A Equação 35 descreve as propriedades de normalidade e ortogonalidade dos estados quânticos.

$$\begin{aligned} \langle\psi_1|\psi_1\rangle &= 1, |\psi_1\rangle \text{ é dito normalizado;} \\ \langle\psi_1|\psi_2\rangle &= 0, |\psi_1\rangle \text{ e } |\psi_2\rangle \text{ são ortogonais.} \end{aligned} \quad (35)$$

A propriedade de ortogonalidade, define que os estados quânticos são distintos e independentes. Para o caso de dois estados ortonormais $|\psi_1\rangle$ e $|\psi_2\rangle$, a probabilidade de encontrar um sistema no estado $|\psi_2\rangle$ se ele estiver inicialmente no estado $|\psi_1\rangle$ é nula.

De mesmo modo, o produto tensorial ou *Kronecker* é um método de combinação de espaços vetoriais, que gera um novo espaço vetorial maior, para representar estados quânticos e para compreender a mecânica quântica de sistemas com mais de um *qubit*. O produto tensorial entre os estados quânticos $|\psi_1\rangle = \alpha_1|0\rangle + \beta_1|1\rangle$ e $|\psi_2\rangle = \alpha_2|0\rangle + \beta_2|1\rangle$ é definido pela Equação 36 (MCMAHON, 2007; YANOFSKY; MANNUCCI, 2008).

$$|\psi_1\rangle|\psi_2\rangle = |\psi_1\psi_2\rangle = |\psi_1\rangle\langle\psi_2| = |\psi_1\rangle \otimes |\psi_2\rangle = \begin{bmatrix} \alpha_1\alpha_2^* & \alpha_1\beta_2^* \\ \beta_1\alpha_2^* & \beta_1\beta_2^* \end{bmatrix} \in \mathbb{C}^4 \quad (36)$$

O Quadro 5 resume algumas notações utilizadas na mecânica quântica, de acordo com a notação de Dirac.

Quadro 5 - Resumo de algumas notações da mecânica quântica.

Notação	Descrição
$ \psi\rangle$	Vetor, chamado de <i>ket</i> .
$\langle\psi $	Vetor, chamado de <i>bra</i> .
$\langle\psi_1 \psi_2\rangle$	Produto interno entre os vetores $ \psi_1\rangle$ e $ \psi_2\rangle$.
$ \psi_1\rangle \otimes \psi_2\rangle$	Produto tensorial entre os vetores $ \psi_1\rangle$ e $ \psi_2\rangle$.
A^*	Conjugado complexo da matriz A , formada pela troca do sinal da parte imaginária.
A^T	Transposta da matriz A .
A^\dagger	Conjugado Hermitiano ou operador adjunto da matriz A , $A^\dagger = (A^T)^*$.
$\langle\psi_1 A \psi_2\rangle$	Produto interno entre $ \psi_1\rangle$ e $A \psi_2\rangle$.

Fonte: Adaptado e traduzido de Nielsen e Chuang (2010).

O estado quântico de um *qubit* descrito em uma base ortogonal é determinado pela regra de Born, por meio de uma medição projetiva. Como descrito anteriormente, o estado quântico irá colapsar em um dos estados $|0\rangle$ ou $|1\rangle$, para a base ortogonal $\{|0\rangle, |1\rangle\}$, quando o estado quântico for medido. Onde, as probabilidades são descritas pelas amplitudes complexas α e β , com probabilidade de $|\alpha|^2$ para o estado $|0\rangle$ e probabilidade de $|\beta|^2$ para o estado $|1\rangle$.

A regra de Born definida pela Equação 37, corresponde à probabilidade de obter um estado quântico específico $|x\rangle$ ao medir um sistema quântico em uma medição projetiva em uma base ortogonal $\{|x\rangle, |x^\perp\rangle\}$, determinada pelo quadrado do produto interno dos dois estados (SCHULD; PETRUCCIONE, F., 2018; YANOFKY; MANNUCCI, 2008).

$$P(x) = |\langle x|\psi\rangle|^2; \sum P(x_i) = 1 \quad (37)$$

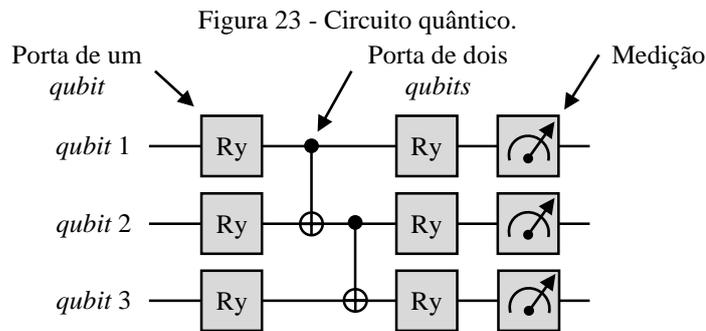
A Equação 38 descreve a aplicação da regra de Born para o *qubit* $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, descrito na base ortogonal $\{|0\rangle, |1\rangle\}$.

$$\begin{aligned} P(0) &= |\langle 0|(\alpha|0\rangle + \beta|1\rangle)|^2 = |\alpha\langle 0|0\rangle + \beta\langle 0|1\rangle|^2 = |\alpha|^2; \\ P(1) &= |\langle 1|(\alpha|0\rangle + \beta|1\rangle)|^2 = |\alpha\langle 1|0\rangle + \beta\langle 1|1\rangle|^2 = |\beta|^2. \end{aligned} \quad (38)$$

De acordo com Garg e Ramakrishnan (2020), o potencial da computação quântica sobre a computação clássica está relacionada na capacidade dos *qubits* exibirem os fenômenos de superposição e emaranhamento da mecânica quântica.

As transformações no estado do *qubit* são realizadas por operadores representados por matrizes unitárias e reversíveis, chamadas de portas quânticas. Em dispositivos quânticos, as portas quânticas são utilizadas em sequência para a manipulação da informação quântica, essa estrutura apresentada na Figura 23 é chamada de circuito quântico (RIEFFEL; POLAK, 2011; SCHULD; PETRUCCIONE, F., 2018). A Equação 39 descreve a transformação do estado do *qubit* $|\psi\rangle$ por meio da porta quântica U .

$$|\psi_{i+1}\rangle = U|\psi_i\rangle; U^\dagger U = U U^\dagger = I; U^{-1} = U^\dagger \quad (39)$$



Fonte: O autor (2024).

Em circuitos quânticos, como apresentado na Figura 23, cada linha representa um *qubit*, que é modificado por meio da sequência de operações aplicadas pelas portas quânticas (esquerda para direita), até a realização da medição do estado quântico. As portas quânticas são classificadas em portas que operam em um único *qubit* (*single-qubit gates*) e portas que atuam em múltiplos *qubits* (*multi-qubit gates*).

3.4.1 Superposição quântica

A superposição quântica, definida nas Equações 32 e 40, é o fenômeno da mecânica quântica que define o *qubit* estar simultaneamente em mais de um estado, ou de ser formado pela combinação linear de estados quânticos. Para o espaço vetorial complexo bidimensional, formado pelos vetores $|0\rangle$ e $|1\rangle$, o *qubit* é considerado em superposição quântica com possibilidade de assumir não somente os estados $|0\rangle$ e $|1\rangle$, mas qualquer outro formado pela combinação desses dois estados (RIEFFEL; POLAK, 2011).

$$|\psi\rangle = \sum_{i=1}^n \alpha_i |\psi_i\rangle \quad (40)$$

Nielsen e Chuang (2010) definem um *qubit* como aquele que pode existir em um *continuum* de estados entre o vetor base, até que seja medido.

A vantagem da superposição quântica está no paralelismo quântico, que se refere à capacidade dos computadores quânticos de realizar cálculos em paralelo em um espaço de estados quânticos probabilísticos (GARG; RAMAKRISHNAN, 2020).

3.4.2 Emaranhamento quântico

O emaranhamento quântico é outro fenômeno fundamental da mecânica quântica, descrito por sistemas compostos, que define uma relação condicional entre os *qubits*. McMahon (2007) define que, se dois sistemas *A* e *B* são emaranhados, certos valores das propriedades do sistema *A* são correlacionados com os valores das propriedades do sistema *B*, mesmo que os dois sistemas estejam espacialmente separados.

O estado quântico para um sistema formado por um conjunto de *qubits*, associado ao espaço vetorial *V*, pode ser determinado pelo produto tensorial em que $|v_i\rangle$ está contido em *V*, conforme apresentado pela Equação 41 (RIEFFEL; POLAK, 2011).

$$|\psi\rangle = |v_1\rangle \otimes |v_2\rangle \otimes \cdots \otimes |v_n\rangle \quad (41)$$

Horodecki *et al.* (2009) define que sistemas quânticos formados por um conjunto de *qubits*, são classificados em dois estados:

- a) Estado quântico separável: Quando o estado quântico de um sistema composto é descrito pelo produto tensorial da Equação 41;
- b) Estado quântico emaranhado: Quando a Equação 41 não consegue descrever o estado quântico de um sistema composto. Estados quânticos emaranhados são descritos pela soma não trivial de produtos tensoriais.

Em sistemas quânticos compostos emaranhados, os estados individuais estão intimamente relacionados entre si, em que o resultado de uma medição para um estado quântico determinará sempre o resultado da medição para um outro estado (YANOFSKY;

MANNUCCI, 2008). Conforme Horodecki *et al.* (2009), os estados quânticos emaranhados exibem correlações entre sistemas quânticos que as correlações clássicas não conseguem capturar.

3.4.3 Interferência quântica

Interferência quântica, outro princípio da mecânica quântica presente na computação quântica, se assemelha ao comportamento da interferência em ondas mecânicas. Quando duas ondas mecânicas se sobrepõem em um mesmo ponto do espaço, origina-se uma nova onda com amplitude diferente, devido ao fenômeno da interferência.

Na computação quântica, as amplitudes complexas podem ser interferidas e por consequência as probabilidades de ocorrência diminuídas, devido ao fato que números complexos podem se anular quando somados (YANOFSKY; MANNUCCI, 2008). McMahon (2007) define os dois tipos de interferência:

- a) Interferência positiva: Quando as amplitudes complexas adicionam construtivamente para aumentar a probabilidade do estado quântico;
- b) Interferência negativa: Quando as amplitudes complexas se somam destrutivamente para diminuir a probabilidade do estado quântico.

3.4.4 Portas quânticas

Na computação clássica, um conjunto de portas lógicas são utilizadas para formar um circuito digital que serve para manipular ou para processar informações. Em circuitos quânticos, as portas quânticas são utilizadas para modificar o estado quântico de um sistema, por meio de operações com matrizes unitárias, descritas pela Equação 39.

Sistemas quânticos admitem uma característica probabilística. Isso significa que o estado de um sistema quântico $|\psi\rangle$ não pode ser determinado com certeza, apenas com uma probabilidade associada a uma amplitude complexa (HIRVENSALO, 2013). Dessa maneira, as portas quânticas servem para modificar a distribuição de probabilidade associada a um estado quântico e/ou para definir correlação entre os estados quânticos, por meio do emaranhamento.

A diferença da computação clássica em relação a computação quântica está no fato que o espaço vetorial clássico se combina pela soma direta, enquanto sistemas quânticos são combinados pelo produto tensorial. A principal diferença reside no tamanho dos espaços

vetoriais: o sistema clássico apresenta dimensão de $2n$, enquanto o sistema quântico, tem o espaço vetorial resultante de 2^n (RIEFFEL; POLAK, 2011).

3.4.4.1 Portas quânticas para um *qubit*

Portas quânticas para um *qubit* são operadores unitários que operam individualmente em *qubit* para modificar o estado quântico, controlando a probabilidade de estar em um determinado estado. As portas quânticas para um *qubit* são formalmente descritas por matrizes unitárias de dimensão 2×2 (SCHULD; PETRUCCIONE, F., 2018).

Uma das portas quânticas mais utilizadas na computação quântica é a porta Hadamard, definida pela Equação 42, sendo responsável por induzir superposição balanceada no estado quântico.

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad (42)$$

A finalidade da porta Hadamard está no fato de que, para um sistema quântico com n *qubits*, com cada *qubit* inicialmente no estado $|0\rangle$ e aplicando em paralelo para cada *qubit* a porta Hadamard, o estado produzido é uma superposição igual de todos os estados (WILLIAMS, 2011).

Adicionalmente, outras portas para um único *qubit* comumente utilizadas são as matrizes de Pauli, definidas na Equação 43. Onde I corresponde à matriz identidade e as portas X , Y e Z são matrizes que rotacional a *Bloch sphere* em 180° em torno dos eixos x , y e z , respectivamente.

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}; \quad X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (43)$$

A porta Pauli- X , também conhecida como *bit flip*, é análoga à porta lógica clássica NOT, pois inverte o estado inicial do *qubit*, conforme demonstrado na Equação 44. Já a porta Pauli- Z é referida como *phase flip*, pois transforma o *qubit* $|\psi_i\rangle = \alpha|0\rangle + \beta|1\rangle$ para um estado $|\psi_{i+1}\rangle = \alpha|0\rangle - \beta|1\rangle$, como evidenciado na Equação 44.

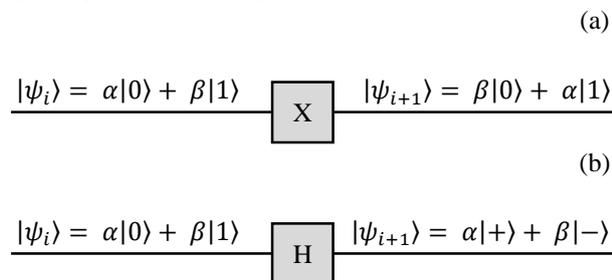
$$\begin{aligned} X|0\rangle &= |1\rangle \text{ e } X|1\rangle = |0\rangle; \\ Z|0\rangle &= |0\rangle \text{ e } Z|1\rangle = -|1\rangle. \end{aligned} \quad (44)$$

Além disso, existem as portas quânticas que atuam em um único *qubit* que são controladas por parâmetros. Como as portas de rotação em relação aos eixos x , y e z da *Bloch sphere*, denominadas como R_x , R_y e R_z , que são utilizadas quando não se pretende rotacionar em 180° , mas em um ângulo θ específico em relação a uma direção em particular (YANOFSKY; MANNUCCI, 2008). A Equação 45 define as portas de rotação paramétricas, em que para $\theta = \pi$, as portas de rotação R_x , R_y e R_z são equivalentes as portas X , Y e Z de Pauli, respectivamente.

$$\begin{aligned} R_x(\theta) &\equiv e^{-i\theta X/2} = \cos\frac{\theta}{2}I - i\text{sen}\frac{\theta}{2}X = \begin{bmatrix} \cos\frac{\theta}{2} & -i\text{sen}\frac{\theta}{2} \\ -i\text{sen}\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}; \\ R_y(\theta) &\equiv e^{-i\theta Y/2} = \cos\frac{\theta}{2}I - i\text{sen}\frac{\theta}{2}Y = \begin{bmatrix} \cos\frac{\theta}{2} & -i\text{sen}\frac{\theta}{2} \\ i\text{sen}\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix}; \\ R_z(\theta) &\equiv e^{-i\theta Z/2} = \cos\frac{\theta}{2}I - i\text{sen}\frac{\theta}{2}Z = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix}. \end{aligned} \quad (45)$$

As ações das portas quânticas sobre o *qubit* $|\psi\rangle$, podem ser representadas por circuitos quânticos, onde cada porta quântica é representada por um bloco e cada linha representa um *qubit*. A Figura 24 apresenta a ação das portas Pauli-X e Hadamard para o *qubit* na base $\{|0\rangle, |1\rangle\}$.

Figura 24 - Ação da porta quântica em um *qubit*. (a) Porta Pauli-X; (b) Porta Hadamard.



Fonte: Adaptado de McMahon (2007).

3.4.4.2 Portas quânticas para múltiplos *qubits*

Portas quânticas para múltiplos *qubits* são a classe de operadores unitários que atuam em dois ou mais *qubits* simultaneamente. Nesta classe existem as portas chamadas de portas de controle quânticas, que utilizam um *qubit* de controle para determinar se uma operação unitária específica será aplicada para um *qubit* alvo.

A porta de controle NOT, ou CNOT, definida pela Equação 46, é um exemplo de uma porta de controle quântica. Onde a porta atua em dois *qubits*, sendo o primeiro *qubit* considerado o de controle e o segundo *qubit* o alvo (SCHULD; PETRUCCIONE, F., 2018).

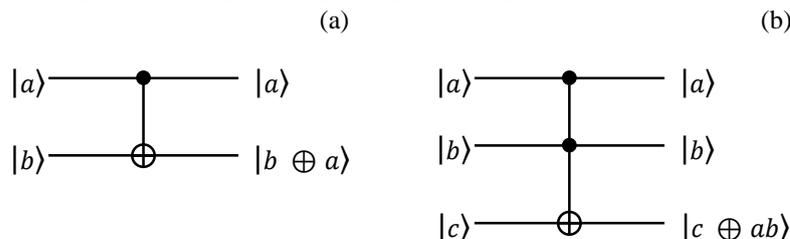
$$CNOT = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (46)$$

A porta CNOT atua de modo que se o *qubit* de controle for definido como $|0\rangle$, o *qubit* alvo não será modificado. Se o *qubit* de controle for definido como $|1\rangle$, o *qubit* de destino será invertido, conforme demonstrado na Equação 47.

$$|00\rangle \rightarrow |00\rangle; |01\rangle \rightarrow |01\rangle; |10\rangle \rightarrow |11\rangle; |11\rangle \rightarrow |10\rangle. \quad (47)$$

Similarmente, a porta Toffoli é uma porta quântica que atua em três *qubits*. Onde dois *qubits* são definidos como de controle, que afetam o estado de um terceiro *qubit*, denominado com *qubit* alvo ou de destino, onde o estado do *qubit* $|c\rangle$ é modificado se e somente se $|a\rangle = |b\rangle = |1\rangle$. A Figura 25 apresenta uma representação típica das portas CNOT e Toffoli no circuito quântico, onde \oplus corresponde ao símbolo da operação OU exclusivo.

Figura 25 - Representação típica das portas para múltiplos *qubits*. (a) CNOT; (b) Toffoli.



Fonte: Adaptado de McMahon (2007) e Nielsen e Chuang (2010).

3.5 QUANTUM MACHINE LEARNING

O *Quantum Machine Learning* (QML) é um campo emergente da ciência da computação que combina os princípios da computação quântica com as técnicas de aprendizado de máquina tradicionais.

Os algoritmos de QML buscam explorar os fenômenos de superposição e emaranhamento para solução de problemas complexos de forma eficiente e mais rápida. Teoricamente, essa abordagem possui o potencial de superar os algoritmos clássicos, sendo aplicável em áreas como SVM quântico, PCA quântico e até mesmo aprendizado por reforço quântico (BIAMONTE *et al.*, 2017).

A Figura 26 ilustra quatro diferentes possibilidades de como a computação quântica pode ser aplicada à aprendizagem de máquina. A computação quântica e a aprendizagem de máquina podem ser combinadas dependendo se os dados de treinamento são quânticos (Q) ou clássicos (C) e se o computador é quântico (Q) ou clássico (C):

- a) CC (Clássico-Clássico): É a situação que opera inteiramente no domínio clássico da computação;
- b) CQ (Clássico-Quântico): Utiliza dados clássicos para o processo de aprendizagem de máquina em *hardwares* quânticos;
- c) QC (Quântico-Clássico): Aplicação dos modelos de ML utilizadas na aprendizagem dos estados quânticos;
- d) QQ (Quântico-Quântico): Abordagem puramente quântica;

Figura 26 - Tipos de abordagens da aprendizagem de máquina e computação quântica.

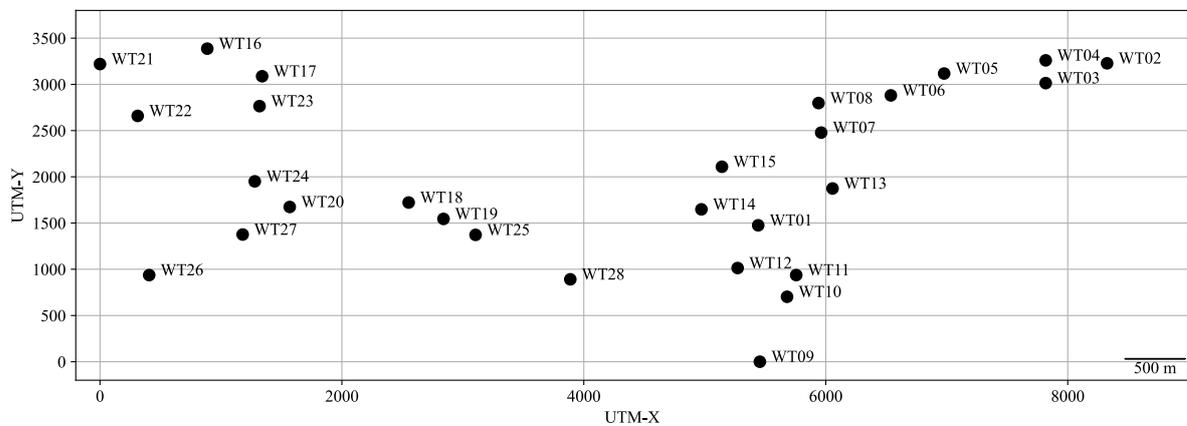
		Tipo de algoritmo	
		CC	CQ
Tipo de dado	Clássico	QC	QQ
	Quântico		
		Clássico	Quântico

Fonte: O autor (2024).

4 METODOLOGIA

Neste estudo, foi explorado o uso do QML na detecção de falhas de rolamentos, utilizando dados de CMS de vibração de 28 turbinas eólicas de um parque eólico localizado no Brasil. A Figura 27 apresenta a distribuição espacial, que representa a distância relativa entre cada turbina eólica, com as identificações codificadas para preservar a confidencialidade das informações do parque eólico.

Figura 27 - Distribuição espacial das turbinas eólicas analisadas.



Fonte: O autor (2024).

Para avaliar a eficácia dos modelos de QML na detecção de falhas, foram considerados três casos distintos de falhas em rolamentos: rolamento principal (MnBrg), rolamento do eixo de alta velocidade da caixa de engrenagens (GbxHss) e rolamento da extremidade sem acionamento do gerador (GnNDE). A Tabela 2 detalha a distribuição de amostras normais e de falhas para cada caso de estudo, incluindo a quantidade de turbinas eólicas e atributos. O Apêndice A fornece uma descrição expandida de cada conjunto de dados, incluindo informações sobre as turbinas eólicas incluídas, bem como o percentual de amostras de falhas registradas em cada uma delas.

Tabela 2 - Descrição resumida do conjunto de dados de cada caso de estudo.

Identificação	TEs	Amostras normais	Amostras de falha	Atributos
MnBrg	10	661	489	41
GbxHss	15	770	543	100
GnNDE	15	1209	257	26

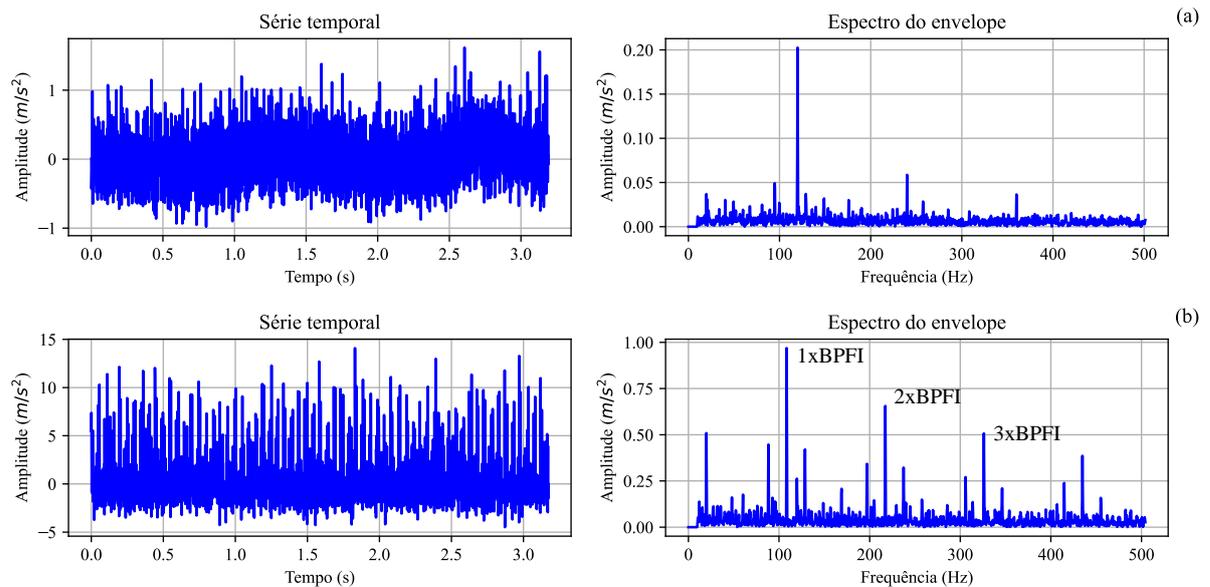
Fonte: O autor (2024).

Os conjuntos de dados apresentados na Tabela 2 foram desenvolvidos como parte do estudo realizado por Leite *et al.* (2023). Neste estudo, as amostras foram rotuladas com o auxílio dos relatórios de O&M e utilizando a análise direta do sinal de vibração, e posteriormente, atributos foram extraídos para uma tarefa de classificação binária (normal = 0 e falha = 1). Dessa maneira, os atributos do domínio do tempo, frequência e mecânico utilizados para o treinamento dos modelos foram:

- Domínio do tempo: Raiz quadrada média, variância, curtose, pico a pico, simetria e fator de crista;
- Domínio da frequência: Frequência média, frequência central, frequência RMS e frequência de desvio padrão;
- Domínio mecânico: Harmônicos da frequência de BPFO, BPFI e BSF, e velocidade do eixo.

A Figura 28 apresenta dois exemplos da série temporal e do espectro do envelope para o caso GnNDE, comparando uma TE normal e uma TE com falha. A análise de envelope foi realizada utilizando uma faixa de filtro de 500 Hz a 10 kHz. Para o comportamento de falha, o espectro apresenta harmônicos e *sidebands* do BPFI.

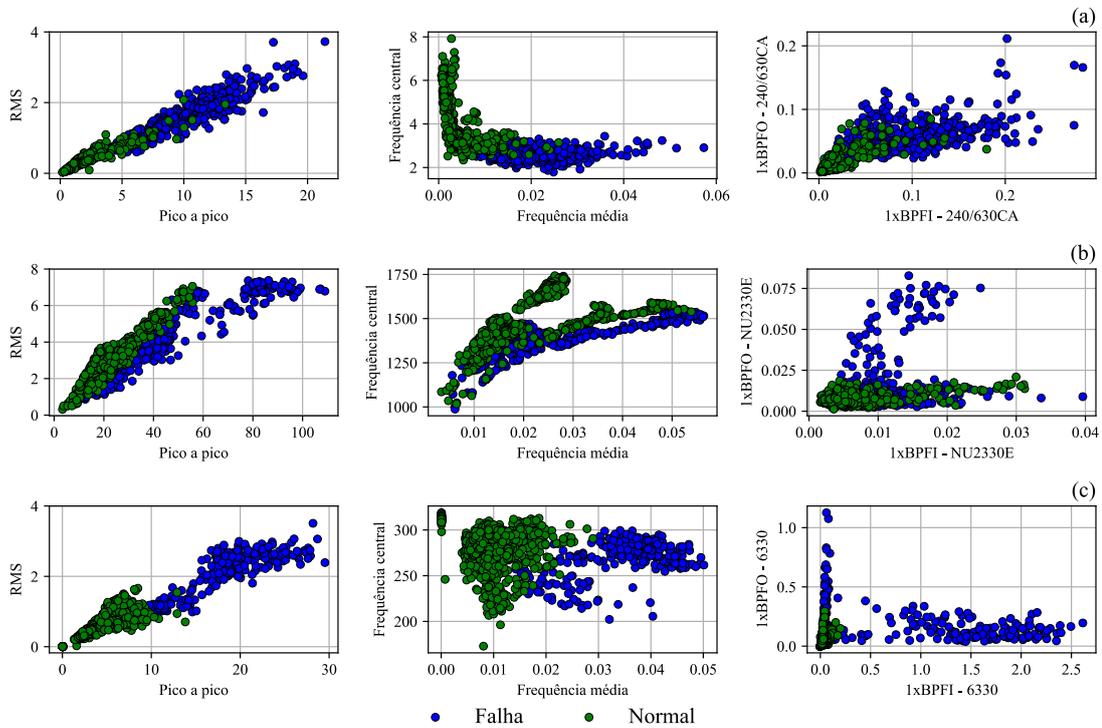
Figura 28 - Série temporal e espectro do envelope para o caso GnNDE. (a) TE normal; (b) TE com falha.



Fonte: O autor (2024).

A Figura 29 mostra a dispersão de alguns atributos extraídos do sinal de vibração por caso de estudo. Observa-se que, no cenário de falha, os dados apresentam uma dispersão maior em comparação com o comportamento normal.

Figura 29 - Dispersão dos dados de cada domínio por caso de estudo. (a) MnBrg; (b) GbxHss; (c) GnNDE.



Fonte: O autor (2024).

O Quadro 6 apresenta as principais bibliotecas utilizadas para aplicação da metodologia. Neste estudo, foram utilizadas exclusivamente bibliotecas de simulação quântica em Python, devido aos seguintes fatores:

- a) São estruturas de código *open-source* que permitem a integração com outras bibliotecas de *Machine Learning*;
- b) Os fenômenos quânticos são simulados em um ambiente livre de erros, em contraste com as limitações dos computadores quânticos atuais (RESCH; KARPUZCU, 2021);
- c) Permitem o desenvolvimento e validação de algoritmos quânticos antes que os *hardwares* quânticos reais estejam amplamente disponíveis.

Quadro 6 - Principais bibliotecas do Python utilizadas na implementação dos modelos.

Biblioteca	Descrição
Pandas	Leitura e manipulação de dados.
Scikit-Learn	Normalização, PCA, separação dos dados para treinamento e teste, e treinamento de modelos de ML.
Matplotlib	Criação de gráficos.
PennyLane	Construção, treinamento e análise de algoritmos quânticos.
Qiskit	Construção, treinamento e análise de algoritmos quânticos.

Fonte: O autor (2024).

A metodologia foi implementada no cluster disponibilizado pelo Centro de Energias Renováveis (CER) da Universidade Federal de Pernambuco (UFPE), que possui 22 CPUs AMD Opteron 6000 e 384 GB de RAM compartilhada. No entanto, neste estudo, utilizou-se apenas um nó com 32 GB de RAM.

Para a detecção de falhas em cada cenário, foram utilizados três tipos de modelos de QML, aplicando dados clássicos em algoritmos de simulação quântica:

- a) Métodos de kernel quântico: *Quantum support vector machine* (QSVM) e agrupamento espectral quântico (AEQ);
- b) Classificador quântico variacional (CQV).

4.1 MÉTODOS DE KERNEL QUÂNTICO

Nesta análise, os métodos de kernel quânticos foram implementados com a mesma finalidade das funções de kernel clássicas: servir como medida de similaridade entre conjuntos de pontos. Esses métodos são empregados como truques de kernel em modelos clássicos, mas os cálculos são realizados por meio de circuitos quânticos.

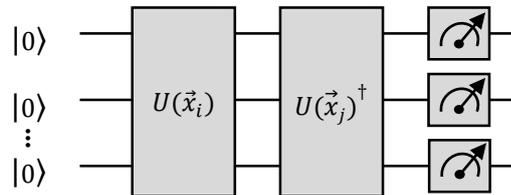
No contexto do QML, os métodos de kernel quânticos são aplicados por meio de mapas de atributos, ou *feature maps*. Os mapas de atributos servem como uma técnica para codificar os dados clássicos em estados quânticos, representados no espaço de Hilbert. O mapa de atributos é definido como um espaço de vetores complexos $|\phi(\vec{x})\rangle$, expresso pela Equação 48, onde $\vec{x} \in \mathbb{R}^d$. Nesta expressão, o operador $|\cdot\rangle\langle\cdot|$ representa a operação do produto tensorial, enquanto ϕ descreve o mapa de atributos (BARTKIEWICZ *et al.*, 2020).

$$\begin{aligned} \phi: \mathbb{R}^d &\rightarrow \mathcal{H} \\ \vec{x} &\rightarrow |\phi(\vec{x})\rangle\langle\phi(\vec{x})| \end{aligned} \quad (48)$$

Como visto nas seções 3.3.2 e 3.3.3 para os modelos de SVM e agrupamento espectral, os métodos de kernel clássicos são técnicas que superam a dificuldade de definir uma função $f(\vec{x})$ para mapear os dados $\vec{x} \in \mathbb{R}^d$ em um espaço de maior dimensionalidade, pela aplicação das funções de kernel $k(\vec{x}_i, \vec{x}_j) = \langle f(\vec{x}_i), f(\vec{x}_j) \rangle$ como uma medida de similaridade ou afinidade entre cada par de pontos do conjuntos de dados. No QML, as funções de kernel quânticos são estabelecidas através do produto interno de matrizes unitárias aplicadas em dois vetores \vec{x}_i e \vec{x}_j , conforme expresso na Equação 49. Estas funções são implementadas por meio de circuitos quânticos, como descrito na Figura 30 (GENTINETTA *et al.*, 2024).

$$k(\vec{x}_i, \vec{x}_j) = |\langle \phi(\vec{x}_j) | \phi(\vec{x}_i) \rangle|^2 = |\langle 0^{\otimes n} | U(\vec{x}_j)^\dagger U(\vec{x}_i) | 0^{\otimes n} \rangle|^2; |\phi(\vec{x})\rangle = U(\vec{x}) | 0 \rangle^{\otimes n} \quad (49)$$

Figura 30 - Circuito quântico para determinação da função de kernel.



Fonte: O autor (2024).

O Quadro 7 apresenta os mapas de atributos aplicados computacionalmente, utilizando a linguagem Python e as bibliotecas de computação quântica PennyLane e Qiskit, para o treinamento dos modelos QSVM (classificação supervisionada) e AEQ (agrupamento não supervisionado).

No algoritmo QSVM, as classes categorizadas como normal e falha foram representadas pelos rótulos -1 e 1, respectivamente. Essa convenção é comumente utilizada em modelos clássicos, pois simplifica o cálculo da distância dos pontos de dados ao hiperplano que maximiza a separação entre as classes. Para os modelos AEQ, uma estratégia adotada para reduzir o tempo de processamento foi utilizar apenas o conjunto de treinamento, aplicando a mesma técnica de divisão de dados utilizada nos modelos de classificação, com uma proporção de 70-30.

Quadro 7 - Mapas de atributos implementados pelas bibliotecas PennyLane e Qiskit no Python.

Biblioteca	Nome	Expressão	Descrição
PennyLane	Amplitude	$\phi: x \rightarrow x\rangle\langle x = \sum_{i,j=1}^N x_i x_j^* i\rangle\langle j $	Codifica 2^n atributos no vetor de amplitudes de n qubits, considerando $\ x\ ^2 = \sum_i x_i ^2 = 1$ (SCHULD, 2021).
	Angular	$\phi: x \rightarrow \phi(x)\rangle\langle\phi(x) $	Codifica N atributos nos ângulos de rotação das matrizes de Pauli de n qubits, com $N \leq n$ (SCHULD, 2021).
Qiskit	Pauli FeatureMap (X e Y)	$U_{\phi(x)} = \exp(i \sum_{S \in [n]} \Phi(x) \prod_{i \in S} P_i)$ $\Phi(x) = \begin{cases} x_i \text{ se } S = \{i\} \\ \prod_{j \in S} (\pi - x_j) \text{ se } S > 1 \end{cases}$	Codifica os dados por uma série de produtos de matrizes de Pauli, onde $P_i \in \{I, X, Y, Z\}$ (HAVLÍČEK <i>et al.</i> , 2019).
	Z FeatureMap	PauliFeatureMap com $P_0 = Z$ e $k = 1$	Primeira ordem do PauliFeatureMap.
	ZZ FeatureMap	PauliFeatureMap com $P_0 = Z$, $P_{0,1} = ZZ$ e $k = 2$	Segunda ordem do PauliFeatureMap.

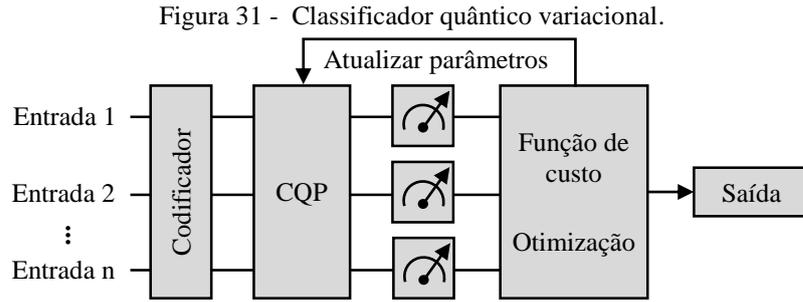
Fonte: O autor (2024).

4.2 CLASSIFICADOR QUÂNTICO VARIACIONAL

O classificador quântico variacional (CQV) pertence à classe de algoritmos de classificação supervisionada do QML, constituído pela definição de circuitos quânticos parametrizados (CQPs). CQPs são compostos por portas quânticas paramétricas cujos parâmetros são ajustados iterativamente para minimizar uma função de custo que mede a discrepância entre as classificações previstas e a classe real dos dados (BLANCE; SPANNOVSKY, 2021).

O CQV representa uma interseção entre a computação quântica e a computação clássica. Nessa abordagem híbrida, as previsões são geradas pelo CQP, enquanto a otimização dos parâmetros é realizada em um ambiente computacional clássico (LUBASCH *et al.*, 2020).

O modelo CQV adotado neste estudo, conforme ilustrado na Figura 31, é constituído por quatro partes: preparação do estado quântico, circuito quântico parametrizado, medição do estado quântico e otimização dos parâmetros.



Fonte: O autor (2024).

O treinamento do modelo de classificação pelo CQV segue o seguinte procedimento:

- 1) Preparação do estado quântico: Transformação da informação clássica para o estado quântico por meio de mapas de atributos;
- 2) Circuito quântico parametrizado: É definido como o circuito quântico $U(\theta) = U_1(\theta_1) \cdots U_2(\theta_2)U_1(\theta_1)$, formado por portas quânticas unitárias que atuam sobre o estado quântico inicial $|\psi_i\rangle$, otimizando seus parâmetros θ . Essas portas modificam o estado quântico de forma a proporcionar uma previsão precisa das classes alvo, conforme expresso pela Equação 50.

$$|\psi_{i+1}\rangle = U(\theta)|\psi_i\rangle \quad (50)$$

- 3) Medição do estado quântico: Retorna as informações dos estados quânticos para uma única possibilidade, colapsando-os em uma única possibilidade com uma probabilidade associada a cada resultado possível;
- 4) Otimização dos parâmetros: Avalia o desempenho do CQP por meio de uma função de custo apropriada e otimiza os parâmetros em uma rotina de pós-processamento clássica, conforme a Equação 51.

$$\theta = \operatorname{argmin}_{\theta} l(\theta) \quad (51)$$

Durante o processo de treinamento, os algoritmos quânticos variacionais podem apresentar o fenômeno conhecido como *barren plateaus*. Essencialmente, *barren plateaus* é uma região plana na região de busca dos parâmetros do CQP, onde o gradiente da função de custo em relação aos parâmetros é nulo (CEREZO *et al.*, 2021). De acordo com Li e Deng

(2022), a presença de *barren plateaus* depende da profundidade do circuito quântico e das propriedades da função de custo utilizada.

Os modelos CQVs foram desenvolvidas utilizando a biblioteca Qiskit, combinando mapas de atributos, CQPs e algoritmos de otimização. Conforme a Figura 32, foram desenvolvidas 18 combinações para a detecção de falhas para cada caso de estudo. No Apêndice B, são apresentadas as identificações dos modelos CQV utilizados.

Figura 32 - Pseudocódigo para combinação das estruturas do modelo CQV.

Pseudocódigo de combinações do CQV

- 1: Lista de mapa de atributos = [PauliFeatureMap-X, ZFeatureMap, ZZFeatureMap]
 - 2: Lista de CQPs = [RealAmplitudes, EfficientSU2, TwoLocal]
 - 3: Lista de otimizadores = [COBYLA, SPSA]
 - 4: Lista de combinações = []
 - 5: **Para cada** mapa de atributo em Lista de mapa de atributos **faça**
 - 6: **Para cada** CQP em Lista de CQPs **faça**
 - 7: **Para cada** otimizador em Lista de otimizadores **faça**
 - 8: **Adicione** [mapa de atributos, CQP, otimizador] em Lista de combinações
 - 9: **Fim**
-

Fonte: O autor (2024).

Para estimar o comportamento médio, os modelos foram treinados separadamente por 10 vezes, usando a entropia cruzada como função de custo, detalhada no Quadro 4.

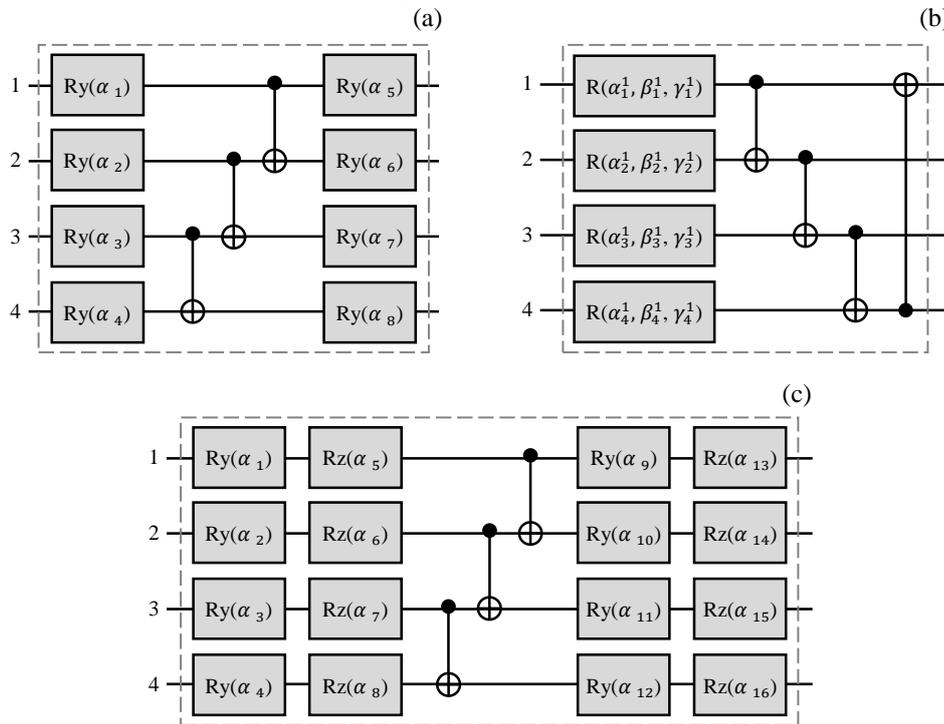
Os algoritmos COBYLA (*Constrained Optimization By Linear Approximation*) e SPSA (*Simultaneous Perturbation Stochastic Approximation*) foram empregados, com um limite de 100 iterações cada. Ambos os algoritmos de otimização não requerem o cálculo direto de gradientes, demonstrando uma taxa de convergência mais rápida do que os métodos baseados em diferenças finitas (PELLOW-JARMAN *et al.*, 2021; XU, R. *et al.*, 2024).

Na codificação dos dados clássicos para o estado quântico, os mapas de atributos utilizados foram PauliFeatureMap-X, ZFeatureMap e ZZFeatureMap. Enquanto para o CQP, foram utilizados três modelos de circuitos quânticos: RealAmplitudes, EfficientSU2 e TwoLocal, conforme apresentado na Figura 33. Esses circuitos são estruturas predefinidas na biblioteca Qiskit que são descritas por:

- a) RealAmplitudes: Consiste em camadas alternadas de rotações R_y e de emaranhamento CNOT-X , aplicadas com a configuração padrão de três repetições do circuito;

- b) EfficientSU2: Consiste em camadas alternadas de rotações R_y e R_z , seguida por portas de emaranhamento CNOT-X, utilizando três repetições padrão do circuito;
- c) TwoLocal: Circuito paramétrico com configurações personalizáveis pelo usuário. Desse modo, a configuração utilizada corresponde a primeira camada do circuito StronglyEntanglingLayers disponível na biblioteca PennyLane, conforme definido por SCHULD *et al.* (2020).

Figura 33 - Circuitos quânticos parametrizados utilizados no modelo CQV. (a) RealAmplitudes; (b) TwoLocal; (c) EfficientSU2.



Fonte: O autor (2024).

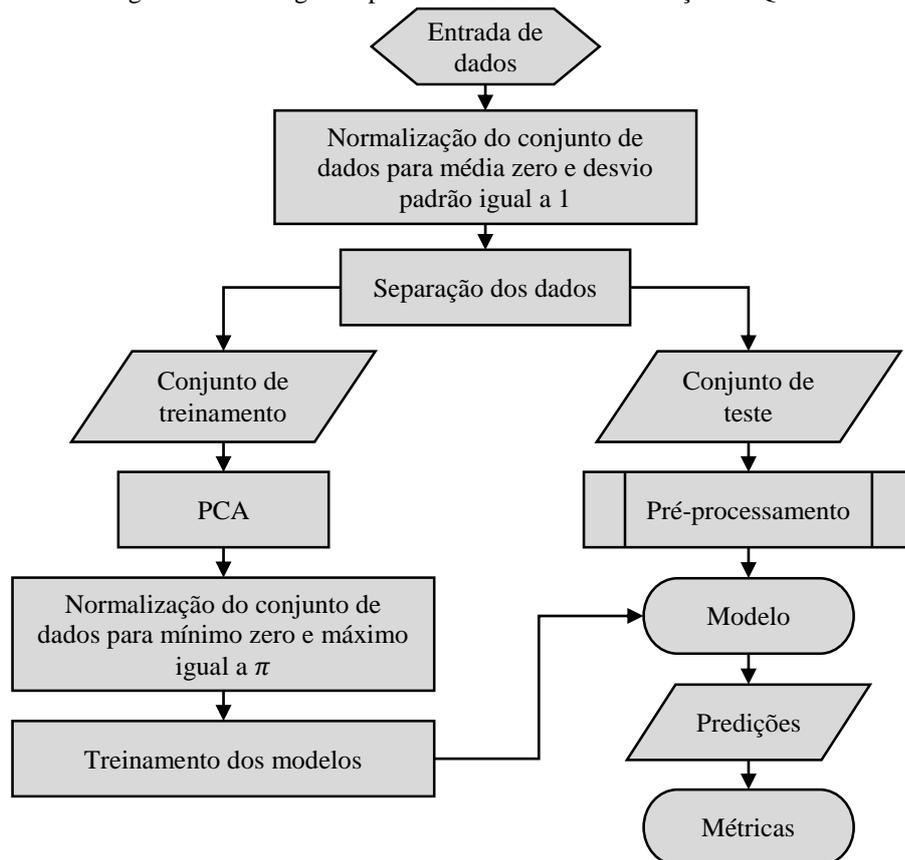
4.3 DESCRIÇÃO DA METODOLOGIA

Nesta seção, será apresentada a descrição da metodologia proposta para a detecção de falhas em rolamentos de TEs. A metodologia utiliza atributos extraídos dos dados de CMS de vibração coletados de um parque eólico em operação no Brasil. Inicialmente, serão descritos os procedimentos para os modelos quânticos de classificação, QSVM e CQV. Em seguida, será apresentado a abordagem para o modelo AEQ.

4.3.1 Modelos quânticos para classificação

Na metodologia de classificação utilizando modelos de QML, as etapas principais empregadas foram: 1) aquisição dos dados; 2) pré-processamento; 3) desenvolvimento dos modelos de classificação; 4) avaliação de desempenho por meio de métricas. A Figura 34 apresenta o fluxograma detalhado da abordagem dos modelos de classificação.

Figura 34 - Fluxograma para os modelos de classificação de QML.



Fonte: O autor (2024).

De acordo com a Figura 34, o primeiro passo da metodologia envolve acessar o conjunto de dados, que é composto pelos atributos dos domínios do tempo, frequência e mecânico extraídos dos dados CMS de vibração.

Para reduzir a dimensionalidade do conjunto de dados e, conseqüentemente, o tempo de processamento, foi adotada a estratégia de transformação dos dados para um espaço de menor dimensão por meio da PCA. Assim, o segundo passo na implementação da metodologia dos modelos de classificação consiste na normalização da base de dados em relação à média e ao desvio padrão, para evitar a influência da alta variabilidade entre os atributos da base de dados. Posteriormente, a base de dados foi separada em conjuntos de treinamento e teste, na proporção

de 70-30, respectivamente. A técnica de PCA foi então aplicada ao conjunto de dados destinado ao treinamento dos modelos.

A etapa anterior ao treinamento consiste na normalização da nova base de dados para mínimo zero e máximo igual a π . Essa prática é recomendada pelos seguintes motivos: 1) os valores de entrada representam ângulos de rotação nas portas quânticas; 2) assegura que todos os atributos contribuam de forma igualmente durante o treinamento dos modelos (BYRNE; COOK; EVANS, 2024).

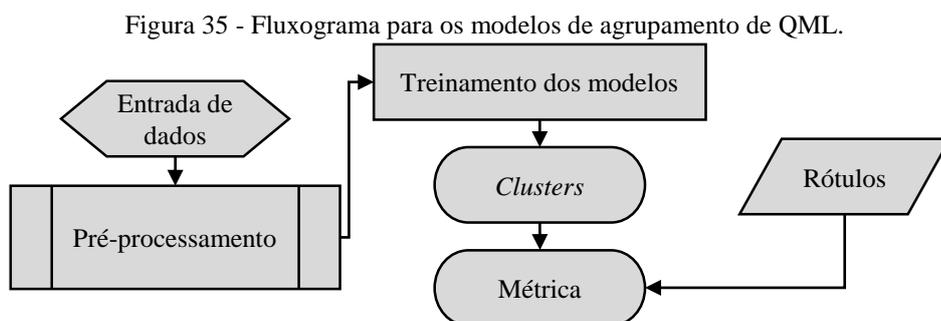
Na fase de avaliação de desempenho dos modelos de classificação, foram empregados os dados do conjunto de teste. Para os algoritmos de aprendizagem de máquina, os dados do conjunto de teste são utilizados para verificar como os modelos se comporta em dados não vistos durante a fase de treinamento. Os dados do conjunto de teste foram ajustados para as componentes principais e normalização definidas pelo conjunto de treinamento.

Neste estudo, os modelos foram avaliados por meio das métricas extraídas da matriz de confusão: acurácia balanceada, precisão, especificidade, *recall* e f1-score.

4.3.2 Modelos quânticos para agrupamento

A segunda abordagem para detecção de falhas em rolamentos de TEs envolveu a aplicação dos modelos AEQ. Esses modelos foram implementados para identificar, nos dados, comportamentos que se diferenciam dos padrões normais de operação.

De forma semelhante a metodologia dos modelos de classificação, as principais etapas utilizadas para os modelos de agrupamento incluíram: 1) aquisição dos dados; 2) pré-processamento; 3) desenvolvimento dos modelos de agrupamento; 4) avaliação de desempenho por meio da informação mútua normalizada (IMN). A Figura 35 ilustra o fluxograma da abordagem dos modelos de agrupamento.



Fonte: O autor (2024).

De acordo com a Figura 35, da mesma forma que nos modelos de classificação supervisionada, o primeiro passo da metodologia para os modelos de agrupamento quântico consiste em acessar o conjunto de dados composto pelos atributos extraídos do sinal de vibração das TEs. Para otimizar o tempo de processamento e avaliar o desempenho dos modelos, optou-se por utilizar apenas o conjunto de treinamento, seguindo a mesma abordagem utilizada no treinamento dos modelos de classificação.

No tocante ao pré-processamento do conjunto de dados, os procedimentos realizados foram: 1) normalização em relação à média e ao desvio padrão; 2) obtenção do conjunto de dados reduzido; 3) redução da dimensionalidade pela PCA; 4) normalização entre $[0, \pi]$.

Neste estudo, os modelos AEQ e QSVM foram treinados simultaneamente, uma vez que compartilham a mesma função de similaridade obtida pelos circuitos quânticos. Para avaliação dos modelos de agrupamento, as saídas foram comparadas com os rótulos estabelecidos na análise direta do sinal de vibração. Essa comparação foi conduzida usando a métrica IMN, que avalia como o algoritmo concorda com as classes reais, sem considerar os valores absolutos dos rótulos. Devido ao fato de que a identificação dos *clusters* como 0 e 1 não está diretamente ligada à classificação real dos dados, pois as saídas dos modelos de agrupamento indicam exclusivamente em qual *cluster* cada ponto do conjunto de dados foi atribuído, sem relação com a verdadeira classe.

5 RESULTADOS E DISCUSSÕES

Neste capítulo, serão apresentados os resultados dos modelos baseados no QML para detecção de falhas de rolamentos, utilizando dados reais de vibração de um parque eólico no Brasil. Inicialmente, serão detalhados os resultados do pré-processamento dos conjuntos de dados para cada caso de estudo. Em seguida, será avaliado o desempenho de modelos clássicos de classificação e agrupamento, visando compará-los posteriormente com o desempenho dos modelos de QML. Posteriormente, serão apresentados os resultados dos modelos de kernel quântico e do classificador quântico variacional, respectivamente. Por fim, os resultados das abordagens clássica e quântica serão comparados.

5.1 PRÉ-PROCESSAMENTO

O pré-processamento é uma etapa fundamental para o treinamento de modelos de ML, pois garante a qualidade dos dados que serão utilizados para treinar os modelos. Dentro dos primeiros passos da aplicação da metodologia deste estudo, a separação dos dados em conjuntos de treinamento e teste serve para garantir a generalização dos modelos e, por consequência, auxilia na avaliação do desempenho dos modelos para dados desconhecidos. A Tabela 3 apresenta a relação do número de amostras por classe para cada conjunto de dados, separados em 70% para treinamento e 30% para teste.

Tabela 3 - Distribuição das classes nos conjuntos de treinamento e teste.

Caso de estudo	Conjunto de treinamento		Conjunto de teste	
	Normal	Falha	Normal	Falha
MnBrg	457	348	204	141
GbxHss	530	389	240	154
GnNDE	854	172	355	85

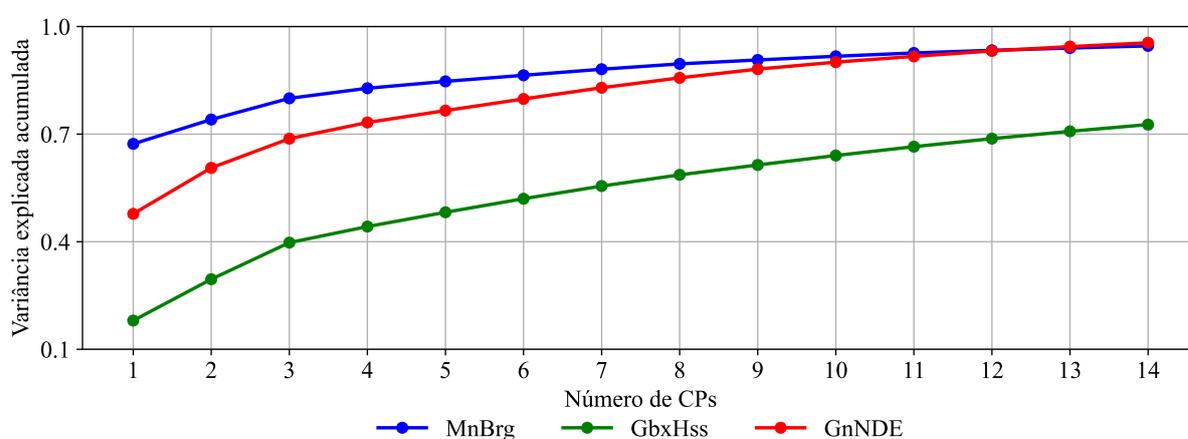
Fonte: O autor (2024).

Durante o tempo de operação de TEs, espera-se que o número de eventos de falhas seja menor do que o número de eventos em que as TEs funcionam sem problemas. Desse modo, modelos treinados com classes desbalanceadas podem ter dificuldade em generalizar e prever corretamente a ocorrência de falhas. Pela Tabela 3, os casos MnBrg e GbxHss foram treinados com mais 40% das amostras de falha, enquanto para o caso GnNDE foi de 17%

aproximadamente. Por esta razão, a métrica utilizada para medir a taxa de acertos foi a acurácia balanceada.

Neste estudo, a redução de dimensionalidade foi conduzida pela técnica de PCA. A Figura 36 apresenta o resultado da variância explicada acumulada pelo número de componentes principais, avaliada pelos conjuntos de treinamento de cada caso de estudo. A variância explicada é uma medida que indica quanto da variabilidade dos dados é explicada por cada componente principal (CP).

Figura 36 - Variância explicada acumulada pelo número de componentes principais.



Fonte: O autor (2024).

Pela Figura 36, nota-se que os casos MnBrg e GnNDE alcançam 70% da variância explicada após quatro CPs, enquanto para o caso GbxHss isso ocorre apenas após treze CPs. Além disso, observa-se que os casos MnBrg e GnNDE mostraram uma variância explicada acumulada semelhante após onze CPs. Para o número de CPs representados na Figura 36, o GbxHss exibiu uma menor variância explicada acumulada entre os casos de estudo, essa discrepância pode ser atribuída à diferença no número de atributos de cada caso de estudo, como evidenciado na Tabela 2. Notavelmente, para o caso GbxHss, mais componentes seriam necessárias para alcançar uma variância explicada mais elevada.

Na literatura, não há um limite teórico estabelecido para determinar o número ideal de CPs com base na variância explicada acumulada para uma representação adequada dos dados. Portanto, neste estudo, os modelos foram avaliados em uma faixa de 2 a 6 CPs, pelos seguintes motivos: 1) nos casos MnBrg e GnNDE, a variância explicada acumulada ultrapassa 50% a partir de 2 CPs; 2) para investigar como o número de CPs influencia o desempenho dos

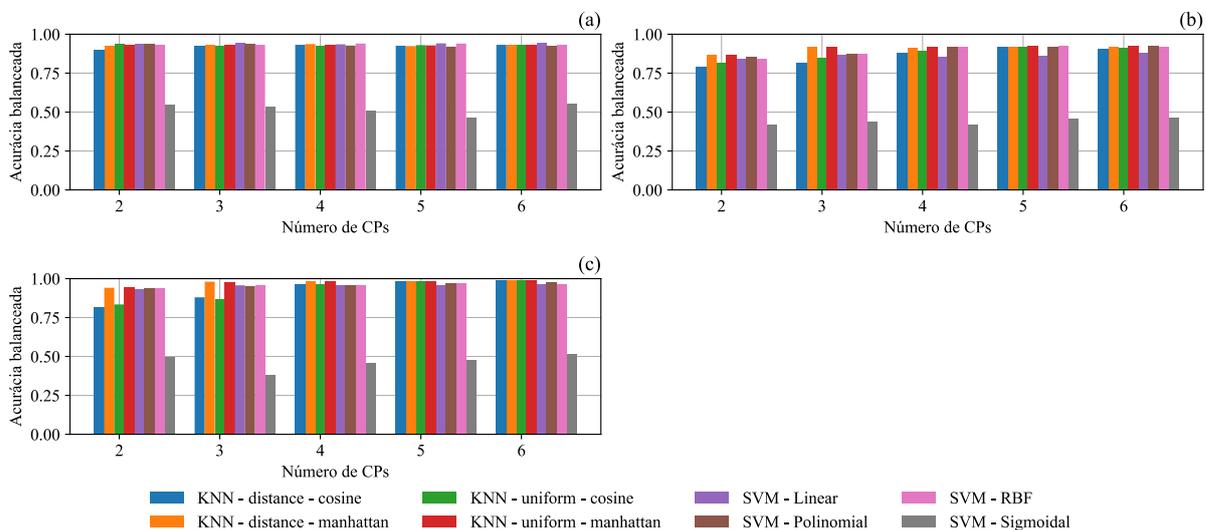
modelos; 3) para otimizar o tempo de processamento, utilizando um menor número de *qubits* nos circuitos quânticos.

5.2 MODELOS CLÁSSICOS PARA DETECÇÃO DE FALHAS

Nesta investigação, os modelos clássicos foram utilizados para comparação do desempenho dos modelos de QML, todos treinados sob condições idênticas. Para abordar o problema de classificação, foram utilizados os modelos KNN e SVM, amplamente reconhecidos e aplicados em problemas dessa natureza. No caso dos modelos clássicos de SVM, todas as funções de kernel listadas no Quadro 3 foram utilizadas. No que diz respeito aos modelos de KNN, foram implementadas quatro configurações distintas, variando entre os hiperparâmetros de pesos (*uniform* e *distance*) e as métricas de distância (*manhattan* e *cosine*).

A Figura 37 apresenta o resultado da acurácia balanceada dos modelos clássicos para o conjunto de teste. No Apêndice C, são apresentados todos os valores das métricas de performance para os modelos clássicos de classificação para o conjunto de teste.

Figura 37 - Resultado da acurácia balanceada dos modelos clássicos para o conjunto de teste. (a) MnBrg; (b) GbxHss; (c) GnNDE.



Fonte: O autor (2024).

De acordo com a Figura 37, verifica-se que, em todos os casos avaliados, os modelos apresentaram uma acurácia balanceada superior a 0,75, exceto o modelo com a configuração SVM-sigmoidal. Além disso, observa-se que entre os casos de estudo, a melhor performance foi para o caso GnNDE. Por outro lado, os modelos analisados apresentaram menor performance para o caso GbxHss. Conforme destacado no Apêndice C, o melhor resultado de

acurácia balanceada para o caso MnBrg foi obtido com o modelo SVM-linear. Já para os casos GbxHss e GnNDE, o melhor desempenho foi para o modelo KNN.

Ainda na Figura 37, percebe-se que a acurácia balanceada dos modelos SVM foi pouco afetada pelo aumento no número de atributos utilizados. Este fato é evidenciado na Tabela 4, que mostra valores de desvio padrão próximos de zero para a acurácia balanceada dos modelos clássicos de todos os casos de estudo.

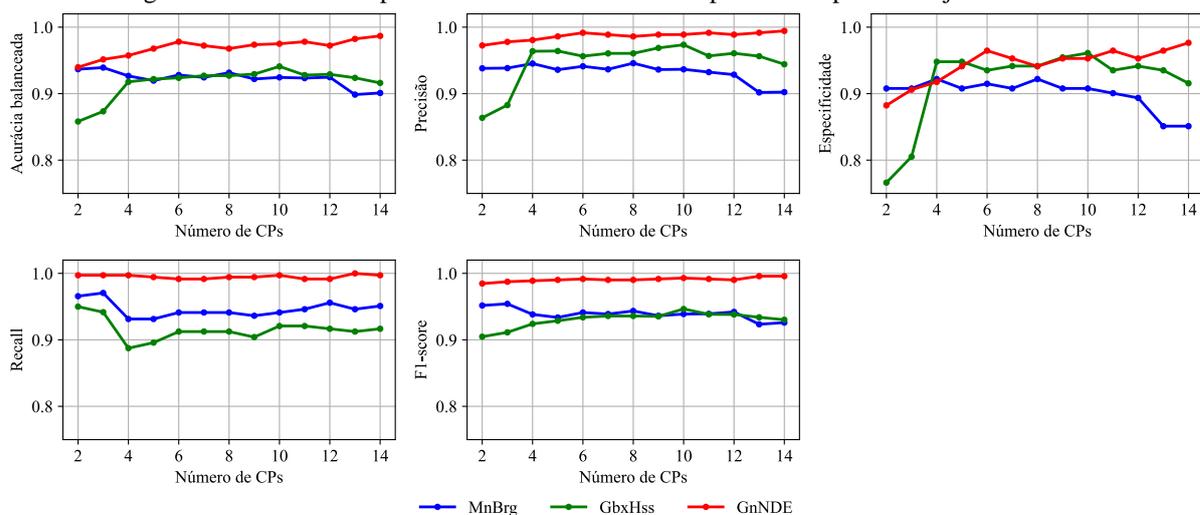
Tabela 4 - Desvio padrão da acurácia balanceada dos modelos clássicos para o conjunto de teste.

Caso de estudo	KNN				SVM			
	UM	UC	DM	DC	L	S	P	RBF
MnBrg	0,003	0,005	0,006	0,014	0,004	0,036	0,008	0,005
GbxHss	0,026	0,045	0,022	0,055	0,012	0,021	0,031	0,037
GnNDE	0,018	0,071	0,020	0,075	0,012	0,052	0,015	0,012

Fonte: O autor (2024).

Entre os modelos de SVM, o SVM-polinomial demonstrou o melhor desempenho médio de acurácia balanceada nos casos MnBrg e GbxHss, alcançando 0,959 e 0,899, respectivamente. Para o caso GnNDE, o modelo SVM-RBF obteve o maior desempenho médio, com 0,935. A Figura 38 apresenta como o desempenho do modelo SVM-polinomial varia conforme o aumento das CPs nos casos de estudo.

Figura 38 - Métricas de performance do modelo SVM-polinomial para o conjunto de teste.

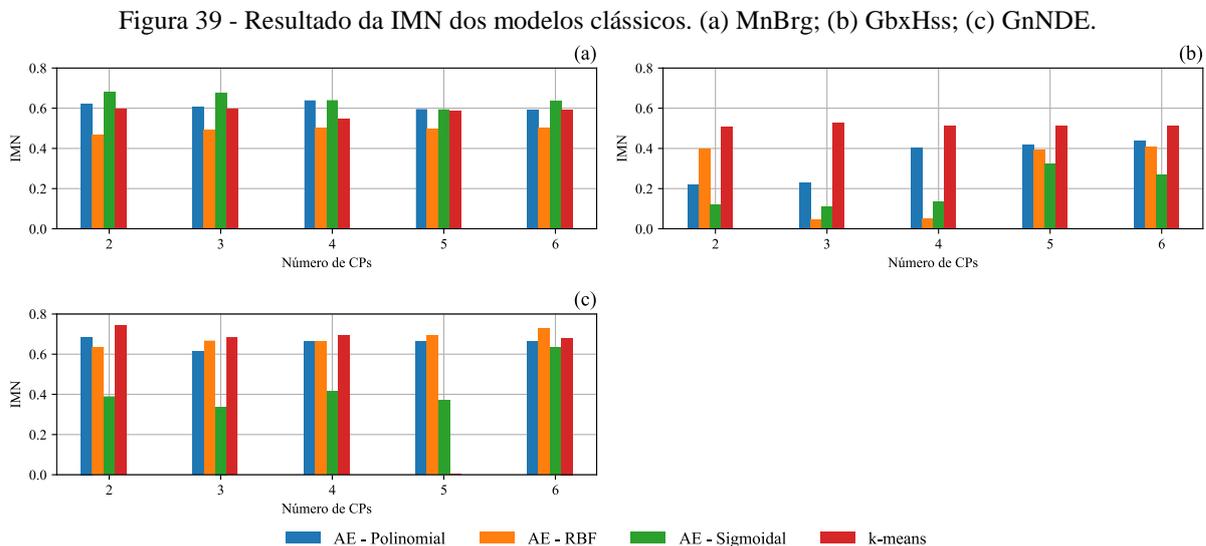


Fonte: O autor (2024).

De acordo com a Figura 38, os melhores resultados da acurácia balanceada foram observados no caso GnNDE. Em contrapartida, os casos MnBrg e GbxHss mostraram desempenho semelhante. Adicionalmente, na Figura 38, nota-se que para os casos MnBrg e GbxHss, a acurácia balanceada mantém-se praticamente constante a partir de quatro CPs. Em termos da capacidade de prever corretamente a condição de falha dos rolamentos, para o modelo SVM-polinomial, observa-se que os casos GbxHss e GnNDE apresentaram desempenho semelhante a partir de três CPs.

Para abordar o problema de agrupamento, foram empregados os modelos *k-means* e agrupamento espectral (AE) para comparar o desempenho nos dados AEQ. No modelo AE, foram aplicadas as funções polinomial, RBF e sigmoidal para avaliar a afinidade entre os pares de pontos do conjunto de dados. Nesta abordagem, optou-se por utilizar o conjunto de dados de treinamento para caracterização dos *clusters*, visando a redução do tempo de processamento, ao mesmo tempo assegurando uma representação suficiente em relação ao número de amostras.

A Figura 39 apresenta o resultado da performance dos modelos clássicos de agrupamento pela métrica IMN. No Apêndice D, são apresentados todos os valores da métrica IMN.



Fonte: O autor (2024).

Pelo resultado apresentado na Figura 39, observa-se que os modelos obtiveram valores para a métrica IMN inferiores a 0,8. Além disso, nota-se que os modelos enfrentaram maior dificuldade em agrupar o conjunto de dados em subgrupos que representem o comportamento normal e de falha do rolamento no caso GbxHss.

No geral, o modelo *k-means* alcançou o melhor desempenho para os casos GbxHss e GnNDE, enquanto para o caso MnBrg, o modelo AE-sigmoidal apresentou o melhor resultado. Por outra perspectiva, o modelo AE-RBF obteve performance inferior nos casos MnBrg e GbxHss, enquanto para o caso GnNDE, foi o modelo *k-means* para o conjunto de dados com cinco CPs.

5.3 MÉTODOS DE KERNEL QUÂNTICO

Nesta análise, foram implementados os modelos de QML que se baseiam em uma medida de similaridade, utilizando funções de kernel avaliadas através de circuitos quânticos. Dessa forma, primeiramente serão apresentados os resultados dos algoritmos QSVM para classificação. Em seguida, serão exibidos os resultados dos modelos AEQ para o agrupamento dos conjuntos de dados em *clusters*. Nos modelos quânticos, cada CP do conjunto de dados reduzido corresponde a um *qubit* no circuito quântico.

Nesta abordagem, foram consideradas seis configurações para o mapa de atributos utilizado no modelo kernel quântico. A métrica principal para a avaliação do desempenho do problema de classificação foi a acurácia balanceada, entretanto, também foram calculadas as métricas de precisão, especificidade, *recall* e f1-score, cujos resultados estão apresentados para caso de estudo no Apêndice E. Para a tarefa de agrupamento a métrica considerada foi a IMN.

Neste estudo, os mapas de atributos disponíveis nas bibliotecas PennyLane e Qiskit. Para os modelos do Qiskit, foram consideradas a configuração padrão do circuito quântico com duas repetições.

5.3.1 *Quantum support vector machine (QSVM)*

A Tabela 5 apresenta o resultado da acurácia balanceada do modelo de QSVM para o conjunto de teste dos casos de estudo, considerando diferentes dimensões dos dados. Em resumo, percebe-se que os modelos quânticos obtiveram uma acurácia balanceada superior a 0,80, exceto os modelos configurados com PauliFeatureMap-Y. Além disso, nota-se uma tendência no aumento da acurácia balanceada com o aumento da dimensionalidade dos dados.

Comparando o desempenho dos modelos entre os diferentes casos de estudo, observa-se que apenas no caso MnBrg todos os modelos atingiram uma acurácia superior a 0,9 na classificação correta dos dados.

Tabela 5 - Resultado da acurácia balanceada dos modelos de QSVM para o conjunto de teste.

Caso de estudo	Nº de CPs	Mapa de atributos					
		Angular	Amplitude	X	Y	Z	ZZ
MnBrg	2	0,929	0,926	0,939	0,500	0,923	0,908
	3	0,929	0,930	0,946	0,500	0,931	0,932
	4	0,933	0,936	0,942	0,500	0,934	0,902
	5	0,932	0,941	0,942	0,500	0,938	0,912
	6	0,938	0,938	0,933	0,500	0,926	0,910
	2	0,841	0,829	0,841	0,500	0,853	0,825
GbxHss	3	0,862	0,820	0,867	0,500	0,880	0,851
	4	0,896	0,868	0,922	0,500	0,903	0,898
	5	0,902	0,900	0,919	0,500	0,905	0,912
	6	0,913	0,887	0,925	0,500	0,900	0,888
	2	0,932	0,878	0,938	0,500	0,849	0,913
GnNDE	3	0,957	0,830	0,957	0,500	0,860	0,764
	4	0,957	0,894	0,963	0,500	0,896	0,849
	5	0,957	0,940	0,969	0,500	0,902	0,909
	6	0,962	0,962	0,963	0,500	0,937	0,878

Fonte: O autor (2024).

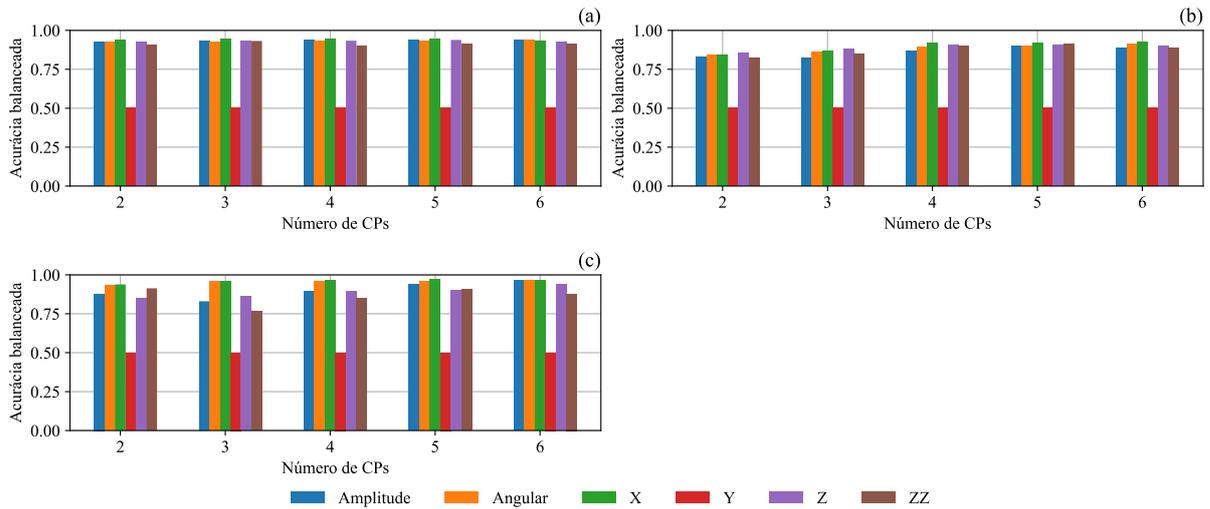
De acordo com a Tabela 5, para os casos GbxHss e GnNDE, verifica-se que o mapa de atributos Angular apresentou um desempenho superior ao mapa de Amplitude. Ademais, percebe-se que, de maneira geral, os modelos utilizando ZFeatureMap superaram aqueles com ZZFeatureMap.

Os resultados apresentados na Tabela 5 indicam que, para todos os casos de estudo, a maior performance para classificação do conjunto de teste foi para o modelo com o PauliFeatureMap-X. No entanto, ao avaliar a capacidade de identificar corretamente as amostras de falha, conforme mostrado no Apêndice E, observa-se que, para o caso MnBrg, os modelos com melhor desempenho foram o Angular e o Amplitude.

Analisando o desempenho do QSVM para classificação da condição do rolamento de turbinas eólicas, destaca-se que as melhores performances foram de 0,969, 0,946 e 0,925 para os casos GnNDE, MnBrg e GbxHss, respectivamente. Em contrapartida, em relação à métrica de especificidade, os melhores resultados foram de 0,965, 0,961 e 0,941 para os casos MnBrg, GbxHss e GnNDE, respectivamente.

A Figura 40 ilustra graficamente o resultado da acurácia balanceada dos modelos de QSVM para o conjunto de teste. A Figura 40 mostra o tempo para o treinamento dos modelos de QSVM.

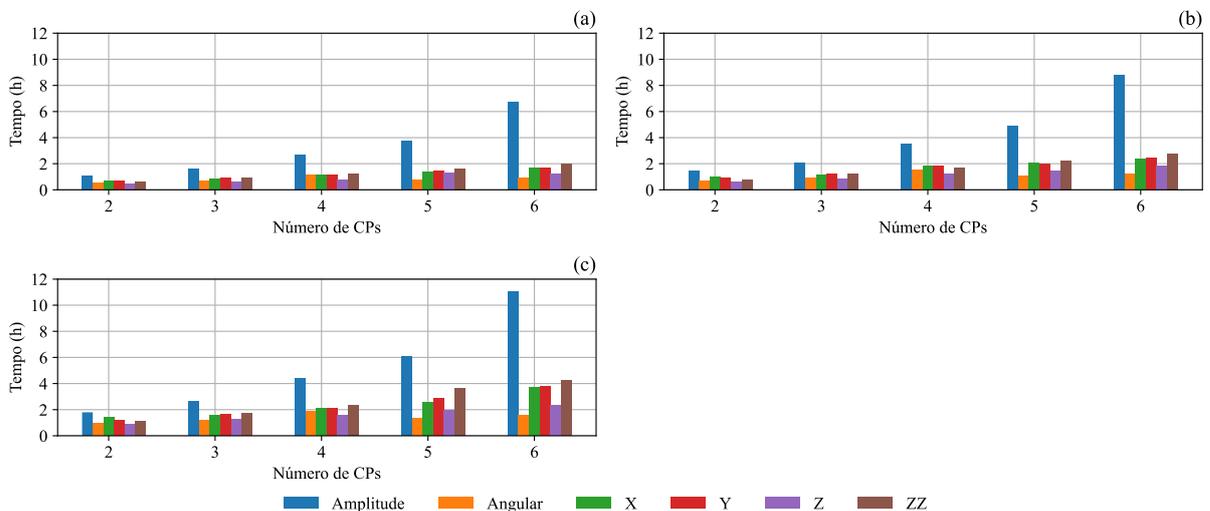
Figura 40 - Resultado da acurácia balanceada dos modelos QSVM para o conjunto de teste. (a) MnBrg; (b) GbxHss; (c) GnNDE.



Fonte: O autor (2024).

A partir da Figura 40, é possível identificar visualmente a diferença do comportamento dos modelos para cada caso de estudo. Para o caso MnBrg, os resultados mostram pouca diferença na performance entre os mapas de atributos Angular, Amplitude e PauliFeatureMap-X, mesmo com o aumento do número de atributos no conjunto de dados. Além disso, percebe-se que existe uma tendência do aumento da acurácia para os casos GbxHss e GnNDE.

Figura 41 - Tempo de treinamento dos modelos de QSVM (a) MnBrg; (b) GbxHss; (c) GnNDE.



Fonte: O autor (2024).

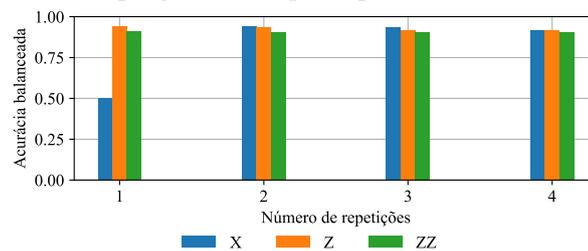
Adicionalmente, nota-se na Figura 40 que, em todos os casos de estudo, o modelo PauliFeatureMap-Y teve o desempenho inferior, classificando todas as amostras como normais. Esse comportamento pode indicar que a configuração utilizada possui uma limitação significativa para diferenciar as amostras entre normais e com falhas.

Em relação ao tempo de execução para o treinamento dos modelos, a Figura 41 mostra que a quantidade de amostras do conjunto de dados influencia significativamente no aumento do tempo de processamento. Isso ocorre porque um conjunto de dados maior exige mais operações para a construção da matriz de kernel.

Além disso, observa-se na Figura 41 que a aplicação do mapa de atributos de Amplitude resultou em um aumento exponencial no tempo de execução com o aumento do número de CPs (*qubits*). Isso se deve ao fato de que, para o mapa de atributos de Amplitude, são necessários $n \geq \log_2(NM)$ *qubits*, onde N é o número de amostras e M é o número de atributos. Para atender a esse critério, foi utilizado o recurso `pad_with = 0` da biblioteca PennyLane, que ajusta automaticamente os atributos para a dimensão 2^n . Isso resultou em um aumento no tamanho do conjunto de dados e, conseqüentemente, no tempo de processamento. Apesar disso, a aplicação dessa abordagem não comprometeu a identificação correta da condição dos rolamentos.

Para avaliar a influência do número de repetições nos mapas de atributos PauliFeatureMap-X, ZFeatureMap e ZZFeatureMap da biblioteca Qiskit, foi realizado um experimento com o caso MnBrg utilizando quatro *qubits*, conforme mostra a Figura 42.

Figura 42 - Avaliação do número de repetições com 4 *qubits* para o modelo QSVM aplicado para o caso MnBrg.



Fonte: O autor (2024).

Pode ser visto na Figura 42, que a configuração PauliFeatureMap-X teve um melhor desempenho ao aumentar o número de repetições de 1 para 2. Entre duas e três repetições, a acurácia balanceada para o conjunto de teste foi melhor na seguinte ordem: PauliFeatureMap-X, ZFeatureMap e ZZFeatureMap. Além disso, observa-se que o uso da porta de controle

CNOT não resultou em uma melhora na acurácia balanceada para o número de repetições considerados.

5.3.2 Agrupamento espectral quântico (AEQ)

Para os modelos de agrupamento AEQ, foram utilizadas as mesmas configurações dos mapas de atributos que foram aplicadas aos modelos de classificação QSVM. No entanto, para o treinamento, foram utilizados os mesmos dados do conjunto de treinamento dos modelos de classificação. Essa abordagem visou reduzir a quantidade de amostras e foi utilizada para garantir uma maior representatividade entre as classes. A Tabela 6 mostra o resultado da métrica IMN para os modelos de AEQ considerados.

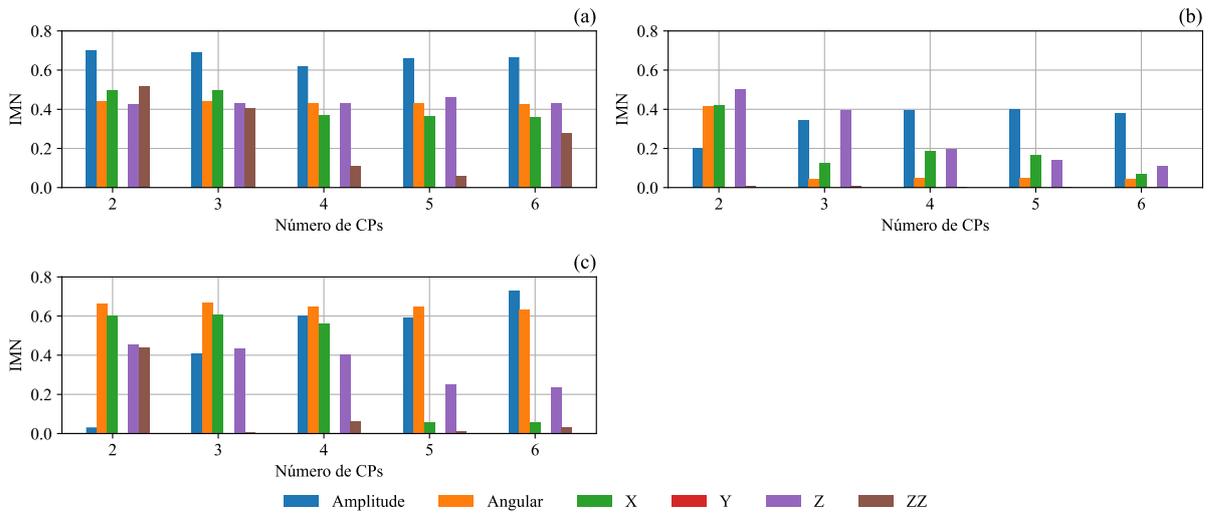
Tabela 6 - Resultados da IMN para os modelos de AEQ.

Caso de estudo	Nº de CPs	Mapa de atributos					
		Angular	Amplitude	X	Y	Z	ZZ
MnBrg	2	0,442	0,699	0,496	0,000	0,426	0,519
	3	0,442	0,691	0,496	0,003	0,431	0,403
	4	0,431	0,620	0,369	0,000	0,431	0,110
	5	0,431	0,662	0,366	0,000	0,459	0,059
	6	0,426	0,666	0,359	0,003	0,431	0,277
GbxHss	2	0,415	0,201	0,419	0,000	0,503	0,009
	3	0,044	0,345	0,123	0,000	0,395	0,005
	4	0,048	0,393	0,186	0,000	0,199	0,000
	5	0,048	0,399	0,167	0,000	0,138	0,001
	6	0,042	0,381	0,070	0,003	0,108	0,000
GnNDE	2	0,662	0,032	0,603	0,000	0,453	0,440
	3	0,667	0,407	0,609	0,000	0,433	0,005
	4	0,651	0,604	0,562	0,002	0,404	0,062
	5	0,651	0,594	0,057	0,000	0,252	0,013
	6	0,635	0,731	0,057	0,000	0,236	0,033

Fonte: O autor (2024).

A Figura 43 apresenta graficamente o resultado da métrica IMN para os modelos de AEQ, mostrando o impacto do aumento no número de atributos no conjunto de dados.

Figura 43 - Resultado da IMN dos modelos de AEQ. (a) MnBrg; (b) GbxHss; (c) GnNDE.



Fonte: O autor (2024).

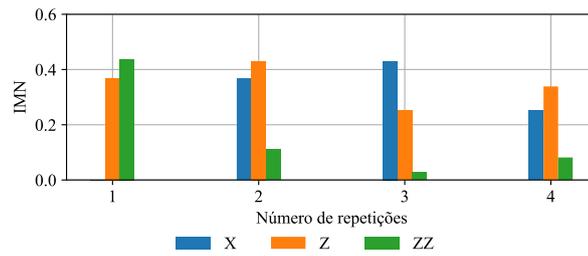
O principal objetivo da aplicação da abordagem de agrupamento foi avaliar a capacidade dos modelos em classificar corretamente os pontos de dados em subgrupos que representem comportamentos normais e de falha dos rolamentos. Conforme mostrado na Figura 29, observa-se um nível de separabilidade entre as classes, com a condição normal apresentando maior compacidade em comparação à condição de falha. Dessa maneira, os modelos de AEQ foram treinados sob a mesma condição de inicialização dos modelos clássicos.

Os resultados mostram que todos os modelos de AEQ apresentaram uma performance inferior a 0,8 em todos os casos de estudo. Para os casos MnBrg e GnNDE, o modelo Amplitude foi o que obteve o melhor desempenho, enquanto para o caso GbxHss, foi o modelo ZFeatureMap.

Como mostrado na Figura 43, o modelo Amplitude demonstrou uma tendência de melhoria na performance com o aumento do número de atributos para os casos GbxHss e GnNDE. Em contraste, o modelo Angular mostrou um comportamento aproximadamente constante em relação ao aumento do número de atributos nos casos MnBrg e GnNDE. Quanto ao desempenho dos modelos ZFeatureMap e ZZFeatureMap, observa-se que o ZFeatureMap conseguiu agrupar de forma mais eficaz os subconjuntos para as condições normais e de falha. Assim como para a tarefa de classificação com QSVM, o modelo PauliFeatureMap-Y para o modelo de AEQ apresentou performance inferior.

A Figura 44 mostra os resultados da IMN para o caso MnBrg, considerando a variação no número de repetições nas configurações dos modelos PauliFeatureMap-X, ZFeatureMap e ZZFeatureMap.

Figura 44 - Avaliação do número de repetições com 4 *qubits* para o modelo AEQ aplicado para o caso MnBrg.

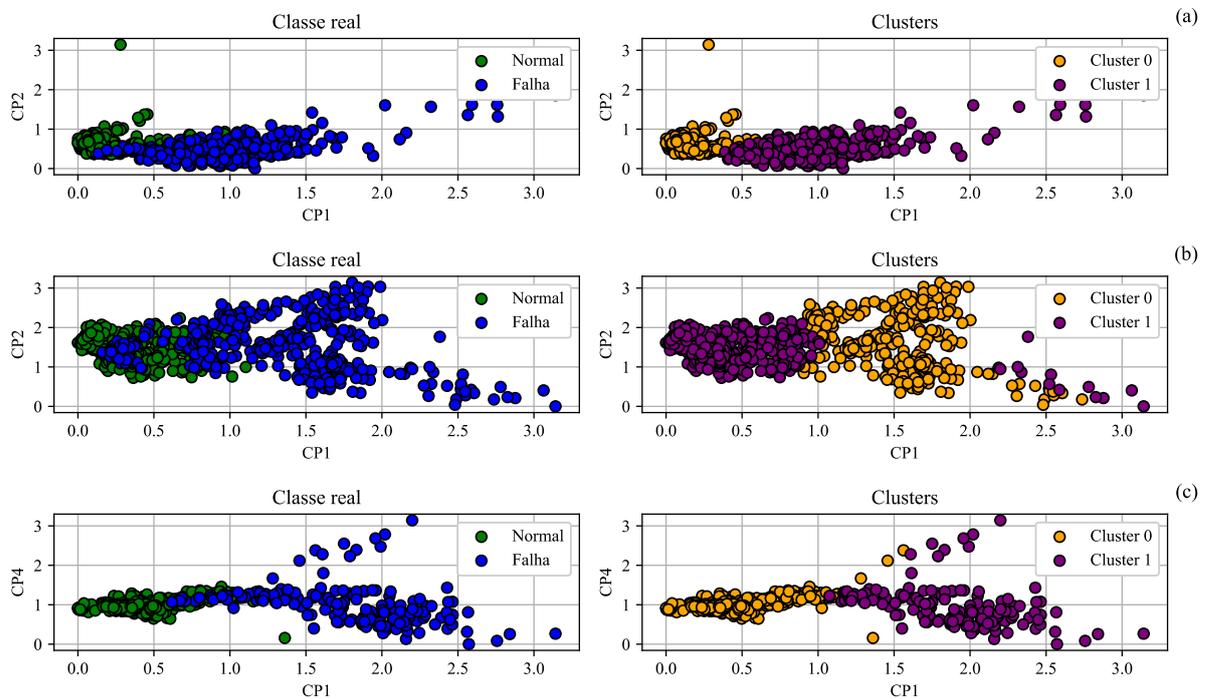


Fonte: O autor (2024).

A variação no número de repetições do modelo PauliFeatureMap-X resultou em um aumento da capacidade de separação correta dos clusters entre 1 e 3 repetições. No entanto, na mesma faixa de repetições, o modelo ZZFeatureMap apresentou uma queda de performance.

Outra maneira adotada para visualizar o comportamento dos *clusters* em relação à classe real dos dados, foi obtida por meio dos gráficos de dispersão bidimensionais ilustrados na Figura 45.

Figura 45 - Comparação da classe real pela posição dos *clusters* do modelo de AEQ de melhor performance por caso de estudo. (a) MnBrg; (b) GbxHss; (c) GnNDE.



Fonte: O autor (2024).

Para o problema de agrupamento, a identificação dos *clusters*, 0 e 1, não pode ser relacionada com a real classificação dos dados, pois indicam exclusivamente em qual *cluster* cada ponto de dados foi atribuído, sem relação com a verdadeira classe. Assim, a métrica

apropriada para avaliar a acurácia é a IMN, que compara a distribuição dos rótulos com a distribuição dos clusters.

Pela Figura 45, é perceptível que os modelos conseguiram separar as áreas densas de cada classe. No entanto, nas regiões de interseção, pode ser afirmado que os modelos tiveram dificuldade em agrupar os dados, diminuindo a métrica de IMN. Como visto na Tabela 6, entre os casos de estudo, o modelo Amplitude aplicado para o caso GnNDE teve maior performance.

5.4 CLASSIFICADOR QUÂNTICO VARIACIONAL (CQV)

Nesta seção, serão apresentados o resultado do modelo CQV aplicado aos três casos de estudo. Esse modelo pertence à classe de algoritmos de aprendizado cujos parâmetros são ajustados continuamente por meio de uma rotina de otimização, avaliada através de uma função de custo.

Nesta abordagem, foram avaliadas 18 variações da configuração do modelo CQV, resultantes da combinação entre o mapa de atributos, o CQP e o algoritmo de otimização, conforme apresentado no Apêndice B. A função de custo utilizada foi a entropia cruzada. No modelo de CQV utilizado, a saída do classificador é determinada por uma função de paridade aplicada a sequência de bits medidos.

Inicialmente, as configurações do modelo CQV foram testadas para 4 *qubits* (4 CPs) pelos seguintes motivos: 1) com 4 CPs nos casos MnBrg e GnNDE, obtém-se uma variância explicada acumulada que ultrapassa 70%; 2) os modelos clássicos e o de QSVM apresentaram pouca variação na acurácia balanceada; 3) com 4 *qubits* e 2 classes, o circuito quântico pode representar um espaço 16 estados, que são mapeadas para uma das 2 classes. Cada configuração foi treinada individualmente 10 vezes para calcular a média das métricas obtidas. Posteriormente, as configurações com a maior média de acurácia balanceada em cada caso de estudo foram avaliadas em uma faixa de 2 a 6 CPs.

A Tabela 7 apresenta o resultado da média e do desvio padrão da acurácia balanceada para as 18 configurações aplicadas aos casos de estudo. A Figura 46 mostra o gráfico de boxplot da acurácia balanceada para o conjunto de teste.

No caso MnBrg, a média da acurácia balanceada variou de 0,656 na combinação 17 a 0,911 na combinação 7. As combinações que utilizaram o mapa de atributos ZFeatureMap e o CQP RealAmplitudes destacaram-se por apresentar as médias mais altas, enquanto aquelas com o mapa de atributos ZZFeatureMap e o CQP TwoLocal mostraram os valores mais baixos. Em

termos de mapas de atributos utilizados, os melhores resultados individuais foram obtidos com o ZFeatureMap, seguido pelo PauliFeatureMap-X.

Tabela 7 - Resultado da acurácia balanceada dos modelos de VQC para o conjunto de teste com 4 qubits.

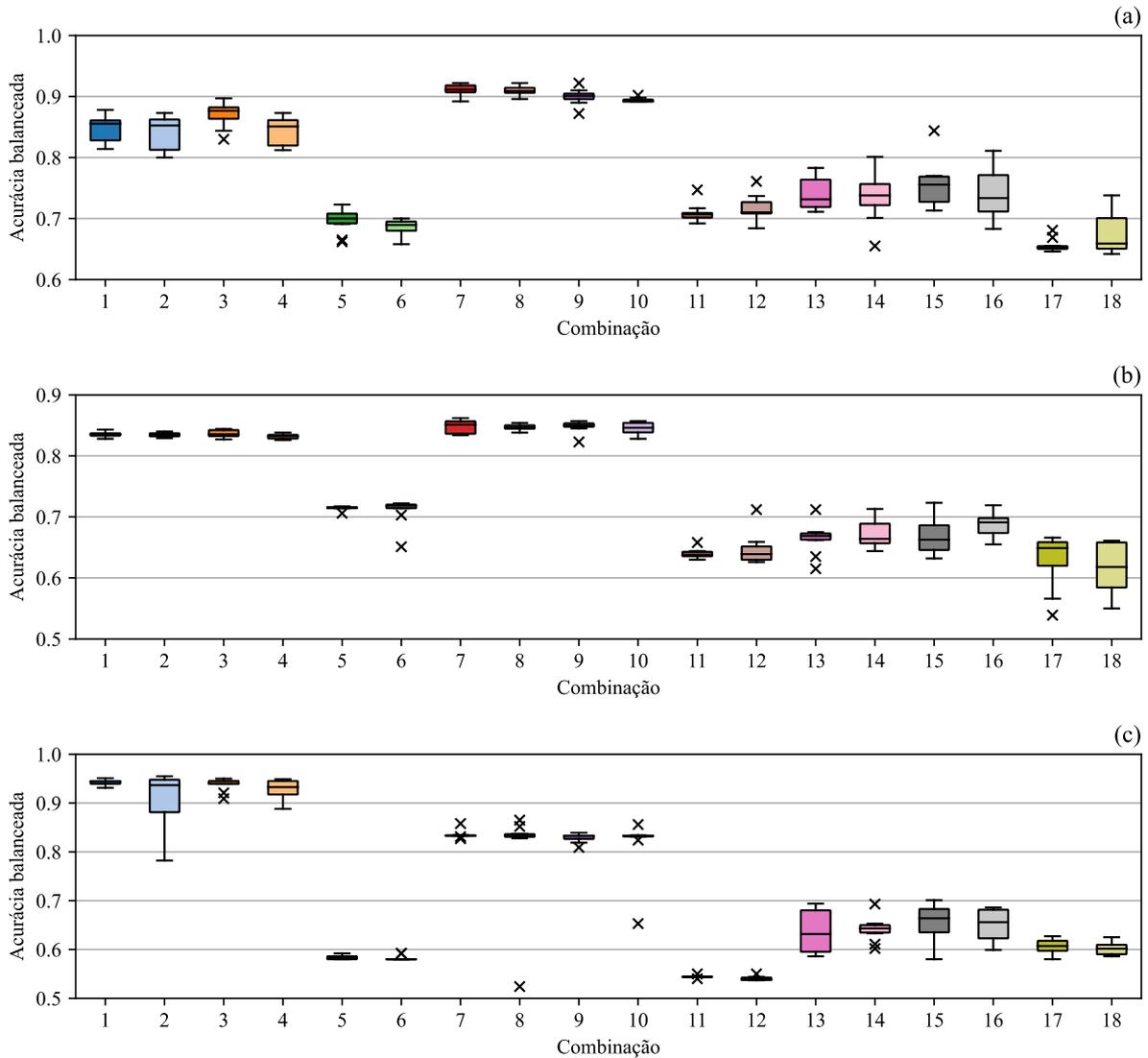
Combinação	Caso de estudo					
	MnBrg		GbxHss		GnNDE	
	Média	Desvio	Média	Desvio	Média	Desvio
1	0,848	0,021	0,836	0,005	0,943	0,005
2	0,840	0,028	0,835	0,004	0,910	0,055
3	0,870	0,020	0,837	0,006	0,938	0,013
4	0,843	0,023	0,832	0,004	0,928	0,021
5	0,697	0,020	0,715	0,003	0,584	0,004
6	0,684	0,015	0,710	0,021	0,582	0,005
7	0,911	0,010	0,848	0,011	0,835	0,008
8	0,909	0,008	0,847	0,005	0,807	0,100
9	0,900	0,013	0,848	0,010	0,829	0,009
10	0,894	0,004	0,845	0,010	0,816	0,058
11	0,709	0,015	0,640	0,008	0,544	0,002
12	0,716	0,022	0,646	0,026	0,540	0,004
13	0,741	0,027	0,665	0,026	0,637	0,047
14	0,736	0,039	0,671	0,023	0,642	0,025
15	0,756	0,038	0,669	0,031	0,653	0,040
16	0,741	0,041	0,687	0,021	0,651	0,033
17	0,656	0,011	0,630	0,044	0,607	0,014
18	0,674	0,033	0,617	0,042	0,602	0,014
Média	0,785	0,022	0,743	0,017	0,725	0,025
Desvio	0,090	0,011	0,093	0,013	0,147	0,026

Fonte: O autor (2024).

Em relação ao caso GbxHss, os maiores valores médios da acurácia balanceada foram para a configuração com ZFeatureMap e os CQPs RealAmplitudes e EfficientSU2. No entanto, a configuração mais eficiente foi a combinação 9 (ZFeatureMap, EfficientSU2 e COBYLA), que obteve uma média de 0,8484 para a acurácia balanceada. Isso demonstra que, para esta análise, a inclusão da porta R_z resultou em um desempenho superior em comparação à

combinação 7 (ZFeatureMap, RealAmplitudes e COBYLA), que alcançou uma média de 0,8477. O valor mais baixo de acurácia balanceada foi registrado para a combinação 18, com uma média de 0,617.

Figura 46 - Boxplot da acurácia balanceada do conjunto de teste. (a) MnBrg; (b) GbxHss; (c) GnNDE.



Fonte: O autor (2024).

No cenário do GnNDE, observa-se uma variação na performance dos modelos quando comparados aos casos MnBrg e GbxHss. Neste caso, os maiores valores médios da acurácia balanceada foram para a configuração com PauliFeatureMap-X e os CQPs RealAmplitudes e EfficientSU2, variando de 0,910 na combinação 2 a 0,943 na combinação 1. A combinação 12 apresentou o menor valor de acurácia balanceada, com uma média de 0,540.

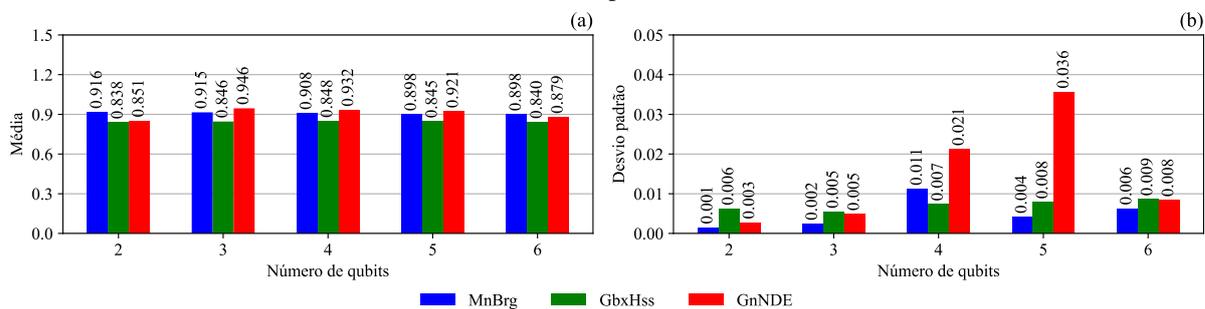
No geral, ao considerar 4 *qubits*, os valores mais altos de acurácia balanceada média foram de 0,943, 0,911 e 0,848 para os casos GnNDE, MnBrg e GbxHss, respectivamente. Os

melhores resultados foram obtidos com os CQPs RealAmplitudes e EfficientSU2. Esse desempenho superior pode estar relacionado tanto à configuração específica do circuito quanto ao fato de terem sido aplicados com 3 repetições. No entanto, seria necessário avaliar o impacto do número de repetições para confirmar essa hipótese. Conforme apresentado na Tabela 7, as médias de desempenho de todos os modelos foram 0,785 para o caso MnBrg, 0,743 para o caso GbxHss e 0,725 para o caso GnNDE. O comportamento geral da aplicação do CQV para o

De acordo com a Figura 46, para todos os casos de estudo, as configurações mais simples para o mapa de atributos (PauliFeatureMap-X e ZFeatureMap) resultaram em uma melhor performance. Além disso, observa-se que os modelos configurados com o CQP TwoLocal apresentaram uma distribuição da média da acurácia balanceada inferior em comparação com os demais modelos.

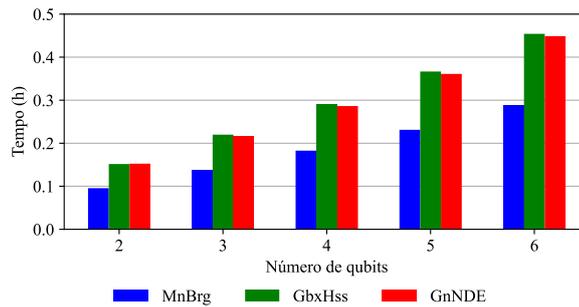
Conforme mencionado, as configurações com as maiores médias de acurácia balanceada no conjunto de teste foram selecionadas para investigar a influência do número de *qubits* no modelo de CQV. As configurações selecionadas foram: MnBrg com ZFeatureMap, RealAmplitudes e COBYLA; GbxHss com ZFeatureMap, EfficientSU2 e COBYLA; e GnNDE com PauliFeatureMap-X, RealAmplitudes e COBYLA. Da mesma forma, as configurações individuais foram treinadas 10 vezes para obter a média das métricas de performance. A Figura 47 exibe os resultados da acurácia balanceada. Outras métricas para as combinações do modelo CQV em cada caso de estudo estão detalhadas no Apêndice F. A Figura 48 mostra o tempo da execução do treinamento do modelo CQV pelo aumento do número de *qubits*.

Figura 47 - Resultado da métrica de acurácia balanceada do modelo CQV para o conjunto de teste. (a) Média; (b) Desvio padrão.



Fonte: O autor (2024).

Figura 48 - Tempo médio de treinamento pelo aumento do número de *qubits* no modelo CQV.



Fonte: O autor (2024).

Para esta análise, observa-se uma diferença pouco significativa no aumento do número de *qubits*, que não resultou em um aumento substancial na média da acurácia balanceada. No entanto, nota-se que o modelo do caso GnNDE apresentou um crescimento no desvio padrão, indicando maior variabilidade na acurácia balanceada com a inicialização de parâmetros diferentes. Além disso, o aumento do número de *qubits* resultou em um maior tempo para o treinamento dos modelos, já que são aplicadas mais operações e mais parâmetros precisam ser atualizados na fase de otimização.

5.5 COMPARAÇÃO DA PERFORMANCE ENTRE OS MODELOS CÁSSICOS E QUÂNTICOS

Nesta seção, serão comparados os resultados dos modelos clássicos e quânticos nas tarefas de classificação e agrupamento. A análise inclui a quantidade de CPs necessárias para alcançar o melhor desempenho, bem como as métricas de avaliação e o tempo de treinamento.

Para o problema de classificação, observa-se que, entre os modelos de melhor desempenho do QSVM e do CQV, a acurácia balanceada foi superior a 0,9 nos casos MnBrg e GnNDE. No caso GbxHss, percebe-se que o modelo do CQV teve maior dificuldade em identificar os estados do rolamento.

Para ambos os modelos, ao utilizar um computador clássico para simular os fenômenos quânticos, nota-se um aumento no tempo de execução à medida que a dimensionalidade dos dados aumenta. No entanto, os modelos do CQV apresentaram um tempo menor tanto no treinamento quanto no teste, devido a diferença característica do funcionamento de cada algoritmo.

A Tabela 8 compara o desempenho dos modelos clássicos e quânticos avaliados neste estudo, considerando a acurácia balanceada e o tempo de treinamento.

Tabela 8 - Comparação dos modelos de melhor performance para a classificação.

Estudo de caso	Abordagem	Modelo	Nº de CPs	Acurácia balanceada	Tempo de treinamento
MnBrg	Clássica	SVM-Linear	3	0,946	0,040 s
	Quântica	QSVM-X	3	0,946	0,860 h
GbxHss	Clássica	SVM-RBF	5	0,924	0,177 s
	Quântica	QSVM-X	6	0,925	2,396 h
GnNDE	Clássica	KNN	6	0,987	0,052 s
	Quântica	QSVM-X	5	0,969	2,536 h

Fonte: O autor (2024).

Primeiramente, observa-se que nos casos MnBrg e GbxHss, o modelo de QSVM com PauliFeatureMap-X apresentou um desempenho igual ou superior aos modelos de SVM Linear e RBF, respectivamente. No entanto, para o caso GbxHss mais informações foram necessárias para atingir uma acurácia balanceada superior.

Para o caso GnNDE, embora o modelo QSVM tenha apresentado uma acurácia balanceada superior a 0,9, este modelo não conseguiu superar o desempenho do modelo KNN. Contudo, considerado o mesmo número de CPs o modelo de QSVM para o caso GnNDE obteve desempenho igual ao modelo de SVM-RBF.

Para o problema de agrupamento, apenas o modelo AEQ aplicado ao caso MnBrg conseguiu superar a métrica de IMN do modelo clássico, conforme apresentado na Tabela 9.

Tabela 9 - Comparação dos modelos de melhor performance para o agrupamento.

Estudo de caso	Abordagem	Modelo	Nº de CPs	IMN	Tempo de treinamento
MnBrg	Clássica	AE - Sigmoidal	2	0,684	0,372 s
	Quântica	AEQ-Amplitude	2	0,699	1,106 h
GbxHss	Clássica	<i>k-means</i>	3	0,526	0,093 s
	Quântica	AEQ-Z	2	0,503	0,632 h
GnNDE	Clássica	<i>k-means</i>	2	0,741	0,075 s
	Quântica	AEQ-Amplitude	6	0,731	11,034 h

Fonte: O autor (2024).

Conforme a Tabela 9, nos casos de GbxHss e GnNDE, os modelos quânticos mostraram pouca diferença em relação aos modelos clássicos correspondentes. Entretanto, quando

comparamos apenas o desempenho entre os modelos AE e AEQ, os modelos quânticos de melhor desempenho superaram os clássicos em todos os casos de estudo.

Para ambos os problemas, classificação e agrupamento, observa-se que os modelos quânticos demandaram um tempo de treinamento maior em comparação com os modelos clássicos analisados. Como mencionado anteriormente, essa diferença se deve ao fato de que as propriedades quânticas foram simuladas em um ambiente clássico. Para obter uma avaliação mais precisa do desempenho dos modelos quânticos, seria necessário analisar seu comportamento em hardwares quânticos reais.

Além disso, ao comparar os resultados obtidos, observa-se que os modelos encontraram maior dificuldade na detecção da falha no caso GbxHss, inclusive para os modelos de QML. Esse desafio pode estar associado ao ambiente e ao ponto de medição do sinal de vibração, uma vez que o rolamento do eixo de alta velocidade pode estar sujeito a perturbações adicionais em comparação aos casos MnBrg e GnNDE. Associado a isso, a detecção de falhas pode ter sido impactada pela diversidade das turbinas eólicas utilizadas, que operam sob condições distintas e podem apresentar comportamentos de falhas variados.

6 CONCLUSÕES

Neste estudo foram aplicadas técnicas de *Quantum Machine Learning* (QML) para detecção de falha em rolamentos de turbinas eólicas utilizando dados de vibração de um parque eólico localizado no Brasil. A abordagem consistiu em empregar atributos extraídos nos domínios do tempo, frequência e mecânico para detecção de falhas no rolamento principal (MnBrg), rolamento do eixo de alta velocidade da caixa de engrenagens (GbxHss) e rolamento da extremidade sem acionamento do gerador (GnNDE), considerando a técnica de PCA para a redução da dimensionalidade.

Neste trabalho, foram explorados dois métodos de QML: os métodos de kernel quântico e o classificador quântico variacional (CQV). Os métodos de kernel quântico atuam como medidas de similaridade que podem ser integradas em modelos clássicos de aprendizado de máquina, como *Support Vector Machine* (SVM) e agrupamento espectral. Já o classificador quântico variacional utiliza circuitos quânticos parametrizáveis para realizar a classificação de dados. Para os modelos de kernel quântico, foram consideradas 6 configurações, enquanto para o modelo CQV, foram testadas 18 combinações, utilizando estruturas de circuitos pré-configuradas fornecidas por bibliotecas de simulação quântica. Os modelos foram avaliados com a dimensionalidade dos dados aumentando de 2 para até 6 componentes principais.

Para o problema de classificação, as principais conclusões com base nos resultados são:

- 1) Os modelos clássicos utilizados para comparar a performance dos modelos de QML apresentaram pouca variação na acurácia balanceada com o aumento do número de atributos utilizados;
- 2) Para os modelos de QSVM, observou-se que os casos GbxHss e GnNDE exibiram uma tendência de aumento na acurácia balanceada com o incremento no número de *qubits*. A melhor performance foi alcançada para todos os casos de estudo utilizando o mapa de atributos PauliFeatureMap-X com duas repetições. No entanto, o aumento do número de repetições dos mapas de atributos baseados no PauliFeatureMap para o caso MnBrg resultou em uma leve diminuição na performance de classificação dos dados;
- 3) Nos modelos baseados no CQV, o melhor desempenho com 4 *qubits* para os casos MnBrg e GbxHss foi observado com o mapa de atributos ZFeatureMap, enquanto o caso GnNDE teve o melhor desempenho com o PauliFeatureMap-X. A configuração TwoLocal do CQP mostrou um desempenho inferior em comparação com as configurações RealAmplitudes e EfficientSU2. Além disso, o aumento no número de

qubits não resultou em um ganho significativo na acurácia média dos modelos com melhor desempenho;

- 4) Ao comparar o desempenho dos modelos quânticos, o QSVM demonstrou uma superioridade sobre o CQV com base na acurácia balanceada do conjunto de teste. Por outro lado, os modelos QSVM mostraram um tempo de execução mais elevado durante o treinamento e teste em comparação com os modelos CQV;
- 5) A comparação entre os modelos clássicos e quânticos revelou que, para os casos MnBrg e GbxHss, o modelo QSVM superou todos os modelos clássicos considerados neste estudo. No entanto, para o caso GnNDE, o QSVM não alcançou um desempenho superior aos modelos de KNN, embora tenha superado os modelos de SVM.

De acordo com os resultados analisados para o problema de agrupamento, as principais conclusões são:

- 1) Para os modelos clássicos de agrupamento, os melhores resultados da métrica IMN foram alcançados com 2 e 3 componentes principais. Além disso, embora ainda inferiores ao melhor resultado geral, foi observado um aumento na IMN para o caso GbxHss quando o número de componentes principais foi aumentado de 3 para 6 nos modelos de AE;
- 2) Nos modelos de AEQ, os melhores resultados foram obtidos para os casos MnBrg e GnNDE usando o mapa de atributos Amplitude, enquanto para o caso GbxHss, o ZFeatureMap apresentou o melhor desempenho;
- 3) Na comparação do desempenho dos modelos de agrupamento, observou-se que a versão quântica apresentou o melhor desempenho em relação aos modelos AE.

No geral, tanto os modelos clássicos quanto os quânticos enfrentaram maior dificuldade para detectar as condições do rolamento no caso GbxHss em comparação aos casos MnBrg e GnNDE. Além disso, os resultados mostraram que, neste estudo, os mapas de atributos do PauliFeatureMap que utilizaram exclusivamente portas de rotação (PauliFeatureMap-X e ZFeatureMap) tiveram um desempenho melhor do que o ZZFeatureMap, que inclui portas CNOT.

Os resultados deste estudo sugerem que a aplicação de técnicas de QML para a detecção de falhas em turbinas eólicas possui um potencial para auxiliar as equipes de operação e

manutenção de parques eólicos. Contudo, este estudo não considerou a performance dos modelos em *hardwares* quânticos, o que impactaria na aplicação real dos modelos.

6.1 TRABALHOS FUTUROS

Para aprofundar a aplicação do QML nos mesmos conjuntos de dados, é sugerido investigar diferentes arquiteturas de circuitos quânticos que possam resultar em uma acurácia aprimorada, além de avaliar outras estruturas de modelos e técnicas de combinação entre os modelos de QML ou entre modelos os clássicos e quânticos.

Para futuras investigações do QML, no contexto do setor de energia eólica, problemas envolvendo modelos quânticos de otimização, regressão ou prognóstico, poderão ampliar a aplicação do QML e, conseqüentemente, auxiliar no aumento do tempo de operação das máquinas, reduzir custos e promover uma manutenção mais eficiente das turbinas eólicas.

REFERÊNCIAS

- ABIODUN, O. I. *et al.* State-of-the-art in artificial neural network applications: A survey. **Heliyon**, 2018. v. 4, n. 11.
- ALI, J. Ben *et al.* Online automatic diagnosis of wind turbine bearings progressive degradations under real experimental conditions based on unsupervised machine learning. **Applied Acoustics**, 2018. v. 132, p. 167–181.
- ALPAYDIN, E. **Introduction to machine learning**. [S.l.]: MIT press, 2020.
- APICELLA, A. *et al.* A survey on modern trainable activation functions. **Neural Networks**, 2021. v. 138, p. 14–32.
- AYESHA, S.; HANIF, M. K.; TALIB, R. Overview and comparative study of dimensionality reduction techniques for high dimensional data. **Information Fusion**, 2020. v. 59, p. 44–58.
- AZEVEDO, H. D. M.; ARAÚJO, A. M.; BOUCHONNEAU, N. A review of wind turbine bearing condition monitoring: State of the art and challenges. **Renewable and Sustainable Energy Reviews**, 2016. v. 56, p. 368–379.
- BARTKIEWICZ, K. *et al.* Experimental kernel-based quantum machine learning in finite feature space. **Scientific Reports**, 2020. v. 10, n. 1, p. 12356.
- BENBOUZID, M. *et al.* Intelligent condition monitoring of wind power systems: State of the art review. **Energies**, 2021. v. 14, n. 18, p. 5967.
- BERETTA, M. *et al.* Improved ensemble learning for wind turbine main bearing fault diagnosis. **Applied Sciences**, 2021. v. 11, n. 16, p. 7523.
- BIAMONTE, J. *et al.* Quantum machine learning. **Nature**, 2017. v. 549, n. 7671, p. 195–202.
- BLANCE, A.; SPANNOVSKY, M. Quantum machine learning for particle physics using a variational quantum classifier. **Journal of High Energy Physics**, 2021. v. 2021, n. 2, p. 1–20.
- BONACCORSO, G. **Machine learning algorithms**. [S.l.]: Packt Publishing Ltd, 2017.
- BREZAK, D. *et al.* A comparison of feed-forward and recurrent neural networks in time series forecasting. [S.l.]: IEEE, 2012. p. 1–6.
- BURGES, C. J. C. A tutorial on support vector machines for pattern recognition. **Data mining and knowledge discovery**, 1998. v. 2, n. 2, p. 121–167.
- BYRNE, D.; COOK, M. G.; EVANS, E. N. A Quick Introduction to Quantum Machine Learning for Non-Practitioners. **arXiv preprint arXiv:2402.14694**, 2024.
- CEREZO, M. *et al.* Variational quantum algorithms. **Nature Reviews Physics**, 2021. v. 3, n. 9, p. 625–644.
- CERVANTES, J. *et al.* A comprehensive survey on support vector machine classification:

- Applications, challenges and trends. **Neurocomputing**, 2020. v. 408, p. 189–215.
- CHAUHAN, V. K.; DAHIYA, K.; SHARMA, A. Problem formulations and solvers in linear SVM: a review. **Artificial Intelligence Review**, 2019. v. 52, n. 2, p. 803–855.
- CORREA-JULLIAN, C. *et al.* Exploring Quantum Machine Learning and Feature Reduction Techniques for Wind Turbine Pitch Fault Detection. **Energies**, 2022. v. 15, n. 8, p. 2792.
- CORTES, C.; VAPNIK, V. Support-vector networks. **Machine learning**, 1995. v. 20, p. 273–297.
- COSTA, Á. M. *et al.* New tendencies in wind energy operation and maintenance. **Applied Sciences**, 2021. v. 11, n. 4.
- CUI, B.; WENG, Y.; ZHANG, N. A feature extraction and machine learning framework for bearing fault diagnosis. **Renewable Energy**, 2022. v. 191, p. 987–997.
- DEMOLLI, H. *et al.* Wind power forecasting based on daily wind speed data using machine learning algorithms. **Energy Conversion and Management**, 2019. v. 198, p. 111823.
- DHIMAN, H. S. *et al.* Wind turbine gearbox anomaly detection based on adaptive threshold and twin support vector machines. **IEEE Transactions on Energy Conversion**, 2021. v. 36, n. 4.
- DIBAJ, A.; GAO, Z.; NEJAD, A. R. Fault detection of offshore wind turbine drivetrains in different environmental conditions through optimal selection of vibration measurements. **Renewable Energy**, 2023. v. 203, p. 161–176.
- DNV. **Energy transition outlook 2023: A global and regional forecast to 2050**. Høvik, Noruega: [s.n.], 2023.
- DONG, Wei *et al.* Intelligent fault diagnosis of wind turbine gearboxes based on refined generalized multi-scale state joint entropy and robust spectral feature selection. **Nonlinear Dynamics**, 2022. v. 107, n. 3, p. 2485–2517.
- EL-THALJI, I.; JANTUNEN, E. A summary of fault modelling and predictive health monitoring of rolling element bearings. **Mechanical systems and signal processing**, 2015. v. 60, p. 252–272.
- ELASHA, F. *et al.* Prognosis of a wind turbine gearbox bearing using supervised machine learning. **Sensors**, 2019. v. 19, n. 14, p. 3092.
- ENCALADA-DÁVILA, Á. *et al.* Wind turbine main bearing fault prognosis based solely on scada data. **Sensors**, 2021. v. 21, n. 6, p. 2228.
- EPE. Consumo Mensal de Energia Elétrica por Classe (regiões e subsistemas). 2024. Disponível em: <<https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/consumo-de-energia-eletrica>>. Acesso em: 22 jun. 2024.

- GAO, Z.; LIU, X. An overview on fault diagnosis, prognosis and resilient control for wind turbine systems. **Processes**, 2021. v. 9, n. 2.
- GARG, S.; RAMAKRISHNAN, G. Advances in quantum deep learning: An overview. **arXiv preprint arXiv:2005.04316**, 2020.
- GENTINETTA, G. *et al.* The complexity of quantum support vector machines. **Quantum**, 2024. v. 8, p. 1225.
- GHAZALI, M. H. M.; RAHIMAN, W. Vibration analysis for machine monitoring and diagnosis: a systematic review. **Shock and Vibration**, 2021. v. 2021, p. 1–25.
- GREENACRE, M. *et al.* Principal component analysis. **Nature Reviews Methods Primers**, 2022. v. 2, n. 1, p. 100.
- GUPTA, P.; PRADHAN, M. K. Fault detection analysis in rolling element bearing: A review. **Materials Today: Proceedings**, 2017. v. 4, n. 2, p. 2085–2094.
- GWEC. **Global wind report 2023**. Bélgica: Global wind energy council, 2023.
- HAHN, B.; DURSTEWITZ, M.; ROHRIG, K. Reliability of wind turbines. [S.l.]: Springer, 2007. p. 329–332.
- HAVLÍČEK, V. *et al.* Supervised learning with quantum-enhanced feature spaces. **Nature**, 2019. v. 567, n. 7747, p. 209–212.
- HEYDARI, A. *et al.* A hybrid intelligent model for the condition monitoring and diagnostics of wind turbines gearbox. **IEEE Access**, 2021. v. 9.
- HIRVENSALO, M. **Quantum computing**. [S.l.]: Springer Science & Business Media, 2013.
- HORODECKI, R. *et al.* Quantum entanglement. **Reviews of modern physics**, 2009. v. 81, n. 2, p. 865.
- HOSSIN, M.; SULAIMAN, M. N. A review on evaluation metrics for data classification evaluations. **International journal of data mining & knowledge management process**, 2015. v. 5, n. 2, p. 1.
- HUANG, X.; WU, L.; YE, Y. A review on dimensionality reduction techniques. **International Journal of Pattern Recognition and Artificial Intelligence**, 2019. v. 33, n. 10, p. 1950017.
- IKOTUN, A. M. *et al.* K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. **Information Sciences**, 2023. v. 622, p. 178–210.
- IRENA. **Renewable Generation Costs in 2022**. Abu Dhabi: [s.n.], 2023.
- IUNUSOVA, E. *et al.* Early fault diagnosis in rolling element bearings: comparative analysis of a knowledge-based and a data-driven approach. **Journal of Intelligent Manufacturing**,

2024. v. 35, p. 2327–2347.

JANIESCH, C.; ZSCHECH, P.; HEINRICH, K. Machine learning and deep learning.

Electronic Markets, 2021. v. 31, n. 3, p. 685–695.

JAVED, K.; GOURIVEAU, R.; ZERHOUNI, N. State of the art and taxonomy of prognostics approaches, trends of prognostics applications and open issues towards maturity at different technology readiness levels. **Mechanical Systems and Signal Processing**, 2017. v. 94.

JIA, W. *et al.* Feature dimensionality reduction: a review. **Complex & Intelligent Systems**, 2022. v. 8, n. 3, p. 2663–2693.

JIANG, Z. *et al.* Structural reliability analysis of wind turbines: A review. **Energies**, 2017. v. 10, n. 12, p. 2099.

JIMÉNEZ, A. A.; MUÑOZ, C. Q. G.; MÁRQUEZ, F. P. G. Machine learning for wind turbine blades maintenance management. **Energies**, 2017. v. 11, n. 1, p. 13.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. **Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences**, 2016. v. 374, n. 2065, p. 20150202.

KARAMAN, Ö. A. Prediction of wind power with machine learning models. **Applied Sciences**, 2023. v. 13, n. 20, p. 11455.

KHOSRAVI, A.; MACHADO, L.; NUNES, R. O. Time-series prediction of wind speed using machine learning algorithms: A case study Osorio wind farm, Brazil. **Applied Energy**, 2018. v. 224, p. 550–566.

KNOPS, Z. F. *et al.* Normalized mutual information based registration using k-means clustering and shading correction. **Medical image analysis**, 2006. v. 10, n. 3, p. 432–439.

KURANI, A. *et al.* A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. **Annals of Data Science**, 2023. v. 10, n. 1, p. 183–208.

LADOSZ, P. *et al.* Exploration in deep reinforcement learning: A survey. **Information Fusion**, 2022. v. 85, p. 1–22.

LEI, Y. *et al.* Applications of machine learning to machine fault diagnosis: A review and roadmap. **Mechanical Systems and Signal Processing**, 2020. v. 138, p. 106587.

LEITE, G. N. P. *et al.* Alternative fault detection and diagnostic using information theory quantifiers based on vibration time-waveforms from condition monitoring systems: Application to operational wind turbines. **Renewable Energy**, 2021. v. 164.

_____ *et al.* A robust fleet-based anomaly detection framework applied to wind turbine vibration data. **Engineering Applications of Artificial Intelligence**, 2023. v. 126.

- LI, W.; DENG, D.-L. Recent advances for quantum classifiers. **Science China Physics, Mechanics & Astronomy**, 2022. v. 65, n. 2, p. 220301.
- LILLICRAP, T. P. *et al.* Backpropagation and the brain. **Nature Reviews Neuroscience**, 2020. v. 21, n. 6, p. 335–346.
- LIN, H.; YE, Y. Reviews of bearing vibration measurement using fast Fourier transform and enhanced fast Fourier transform algorithms. **Advances in Mechanical Engineering**, 2019. v. 11.
- LINS, I. D. *et al.* Quantum machine learning for drowsiness detection with EEG signals. **Process Safety and Environmental Protection**, 2024. v. 186.
- LUBASCH, M. *et al.* Variational quantum algorithms for nonlinear problems. **Physical Review A**, 2020. v. 101, n. 1, p. 10301.
- LUO, Y. *et al.* A study on the modeling method of cage slip and its effects on the vibration response of rolling-element bearing. **Energies**, 2022. v. 15, n. 7.
- LUXBURG, U. VON. A tutorial on spectral clustering. **Statistics and computing**, 2007. v. 17, p. 395–416.
- MAIOR, C. B. S. *et al.* Prognostics and Health Management of Rotating Machinery via Quantum Machine Learning. **IEEE Access**, 2023. v. 11, p. 25132–25151.
- MASSERAN, N. Evaluating wind power density models and their statistical properties. **Energy**, 2015. v. 84, p. 533–541.
- MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. **The bulletin of mathematical biophysics**, 1943. v. 5, p. 115–133.
- MCMAHON, D. **Quantum computing explained**. Estados Unidos: John Wiley & Sons, 2007.
- MENG, D. *et al.* Fault analysis of wind power rolling bearing based on EMD feature extraction. **CMES-Computer Modeling in Engineering & Sciences**, 2022. v. 130, n. 1, p. 543–558.
- MUNGUBA, C. F. De L. *et al.* Ensemble learning framework for fleet-based anomaly detection using wind turbine drivetrain components vibration data. **Engineering Applications of Artificial Intelligence**, 2024. v. 133.
- NALEPA, J.; KAWULOK, M. Selecting training sets for support vector machines: a review. **Artificial Intelligence Review**, 2019. v. 52, n. 2, p. 857–900.
- NEWMAN, M. E. J.; CANTWELL, G. T.; YOUNG, J. G. Improved mutual information measure for clustering, classification, and community detection. **Physical Review E**, 2020. v. 101, n. 4.

- NG, A.; JORDAN, M.; WEISS, Y. On spectral clustering: Analysis and an algorithm. **Advances in neural information processing systems**, 2001. v. 14.
- NGUYEN, P. T. *et al.* Evaluation of a machine learning classifier for metamodels. **Software and Systems Modeling**, 2021. v. 20, n. 6, p. 1797–1821.
- NIELSEN, M. A.; CHUANG, I. L. **Quantum computation and quantum information**. 10. ed. [S.l.]: Cambridge University Press, 2010.
- OLABI, A. G. *et al.* A review on failure modes of wind turbine components. **Energies**, 2021. v. 14, n. 17, p. 5241.
- PANDIT, R.; ASTOLFI, D.; DURAZO, I. A review of predictive techniques used to support decision making for maintenance operations of wind turbines. **Energies**, 2023. v. 16, n. 4, p. 1654.
- PELLOW-JARMAN, A. *et al.* A comparison of various classical optimizers for a variational quantum linear solver. **Quantum Information Processing**, 2021. v. 20, n. 6, p. 202.
- PENG, H. *et al.* A review of research on wind turbine bearings' failure analysis and fault diagnosis. **Lubricants**, 2022. v. 11, n. 1, p. 14.
- RANDALL, R. B.; ANTONI, J. Rolling element bearing diagnostics—A tutorial. **Mechanical systems and signal processing**, 2011. v. 25, n. 2, p. 485–520.
- REBALA, G.; RAVI, A.; CHURIWALA, S. **An introduction to machine learning**. [S.l.]: Springer, 2019.
- RENEWABLES.NIJA. Renewables.nija. 2024. Disponível em: <<https://www.renewables.ninja/>>. Acesso em: 22 jun. 2024.
- RESCH, S.; KARPUZCU, U. R. Benchmarking quantum computers and the impact of quantum noise. **ACM Computing Surveys (CSUR)**, 2021. v. 54, n. 7, p. 1–35.
- REZAMAND, M. *et al.* Critical wind turbine components prognostics: A comprehensive review. **IEEE Transactions on Instrumentation and Measurement**, 2020. v. 69, n. 12, p. 9306–9328.
- RIEFFEL, E. G.; POLAK, W. H. **Quantum computing: A gentle introduction**. Inglaterra: MIT Press, 2011.
- RINGNÉR, M. What is principal component analysis? **Nature biotechnology**, 2008. v. 26, n. 3, p. 303–304.
- ROSENBLATT, F. The perceptron: a probabilistic model for information storage and organization in the brain. **Psychological review**, 1958. v. 65, n. 6, p. 386.
- SAIDI, L.; BENBOUZID, M. Prognostics and health management of renewable energy systems: state of the art review, challenges, and trends. **Electronics**, 2021. v. 10, n. 22.

- SCHULD, M. *et al.* Circuit-centric quantum classifiers. **Physical Review A**, 2020. v. 101, n. 3.
- _____. Supervised quantum machine learning models are kernel methods. **arXiv preprint arXiv:2101.11020**, 2021.
- _____; PETRUCCIONE, F. **Supervised learning with quantum computers**. Berlin: Springer, 2018.
- _____; SINAYSKIY, I.; PETRUCCIONE, F. An introduction to quantum machine learning. **Contemporary Physics**, 2015. v. 56, n. 2, p. 172–185.
- SHER, F.; CURNICK, O.; AZIZAN, M. T. Sustainable conversion of renewable energy sources. **Sustainability**, 2021. v. 13, n. 5.
- SIEGEL, D. *et al.* A comparative study on vibration-based condition monitoring algorithms for wind turbine drive trains. **Wind energy**, 2014. v. 17, n. 5, p. 695–714.
- SILVA, G. S. M.; DROGUETT, E. L. Quantum Machine Learning for Health State Diagnosis and Prognostics. [S.l.]: IEEE, 2022. p. 1–17.
- SINGRU, P. *et al.* Bearing failure prediction using Wigner-Ville distribution, modified Poincare mapping and fast Fourier transform. **Journal of Vibroengineering**, 2018. v. 20, n. 1, p. 127–137.
- SMITH, W. A.; RANDALL, R. B. Rolling element bearing diagnostics using the Case Western Reserve University data: A benchmark study. **Mechanical systems and signal processing**, 2015. v. 64.
- STRÖMBERGSSON, D.; MARKLUND, P.; BERGLUND, K. Increasing wind turbine drivetrain bearing vibration monitoring detectability using an artificial neural network implementation. **Applied Sciences**, 2021. v. 11, n. 8, p. 3588.
- TIAN, M. *et al.* Bearing fault diagnosis of wind turbines based on dynamic multi-adversarial adaptive network. **Journal of Mechanical Science and Technology**, 2023. v. 37, n. 4.
- TSALLIS, C. Entropy. **Encyclopedia**, 2022. v. 2, n. 1, p. 264–300.
- TURBOSQUID. Wind turbine model. 2024. Disponível em: <<https://www.turbosquid.com/FullPreview/755190>>. Acesso em: 26 jun. 2024.
- VEERS, P. *et al.* **Grand Challenges Revisited: Wind Energy Research Needs for a Global Energy Transition**. [S.l.]: National Renewable Energy Laboratory (NREL), Golden, CO (United States), 2023.
- VIEIRA, J. L. De M. *et al.* Remaining Useful Life Estimation Framework for the Main Bearing of Wind Turbines Operating in Real Time. **Energies**, 2024. v. 17.
- WANG, Q. *et al.* A comprehensive survey of loss functions in machine learning. **Annals of**

Data Science, 2020. p. 1–26.

WIERZCHOŃ, S. *et al.* Spectral clustering. **Modern Algorithms of Cluster Analysis**, 2018. p. 181–259.

WILLIAMS, C. P. **Explorations in quantum computing**. 2. ed. Inglaterra: Springer Science & Business Media, 2011.

XIANG, L. *et al.* Fault detection of wind turbine based on SCADA data analysis using CNN and LSTM with attention mechanism. **Measurement**, 2021. v. 175, p. 109094.

XU, R. *et al.* Benchmarking hybrid digitized-counterdiabatic quantum optimization. **Physical Review Research**, 2024. v. 6, n. 1, p. 13147.

XU, Z.; LI, C.; YANG, Y. Fault diagnosis of rolling bearings using an improved multi-scale convolutional neural network with feature attention mechanism. **ISA transactions**, 2021. v. 110.

YANG, S. *et al.* A 2DCNN-RF model for offshore wind turbine high-speed bearing-fault diagnosis under noisy environment. **Energies**, 2022. v. 15, n. 9, p. 3340.

YANG, W.; JIANG, J. Wind turbine condition monitoring and reliability analysis by SCADA information. [S.l.]: IEEE, 2011. p. 1872–1875.

YANG, Y. *et al.* Nonlinear vibration signatures for localized fault of rolling element bearing in rotor-bearing-casing system. **International Journal of Mechanical Sciences**, 2020. v. 173.

YANOFSKY, N. S.; MANNUCCI, M. A. **Quantum computing for computer scientists**. 1. ed. Estados Unidos: Cambridge University Press, 2008.

ZHANG, Y.; NI, Q. Recent advances in quantum machine learning. **Quantum Engineering**, 2020. v. 2.

ZHOU, Z.-H. **Machine learning**. [S.l.]: Springer Nature, 2021.

APÊNDICE A – DESCRIÇÃO DETALHADA DO CONJUNTO DE DADOS POR CASO DE FALHA

Tabela 10 - Descrição do conjunto de dados para o caso MnBrg.

TE	Total de amostras	Amostras normais	Amostras de falha	Percentual de amostras de falha (%)
WT01	97	2	95	98
WT02	91	91	0	0
WT05	90	90	0	0
WT06	90	90	0	0
WT08	114	113	1	1
WT12	103	98	5	5
WT13	102	102	0	0
WT16	236	10	226	96
WT18	113	42	71	63
WT21	114	23	91	80

Fonte: O autor (2024).

Tabela 11 - Descrição do conjunto de dados para o caso GbxHss.

TE	Total de amostras	Amostras normais	Amostras de falha	Percentual de amostras de falha (%)
WT03	88	88	0	0
WT04	89	0	89	100
WT05	90	90	0	0
WT07	90	90	0	0
WT09	66	66	0	0
WT10	91	90	1	1
WT11	91	91	0	0
WT15	90	90	0	0
WT17	89	44	45	51
WT19	91	7	84	92
WT22	91	24	67	74
WT24	88	6	82	93
WT25	88	2	86	98
WT26	86	0	86	100
WT28	85	82	3	4

Fonte: O autor (2024).

Tabela 12 - Descrição do conjunto de dados para o caso GnNDE.

TE	Total de amostras	Amostras normais	Amostras de falha	Percentual de amostras de falha (%)
WT01	90	90	0	0
WT04	90	90	0	0
WT05	97	97	0	0
WT06	96	96	0	0
WT09	141	141	0	0
WT12	92	0	92	100
WT14	91	91	0	0
WT16	91	91	0	0
WT17	107	107	0	0
WT18	91	91	0	0
WT19	91	89	2	2
WT20	91	0	91	100
WT22	91	91	0	0
WT23	116	96	20	17
WT27	91	39	52	57

Fonte: O autor (2024).

APÊNDICE B – IDENTIFICAÇÃO DAS COMBINAÇÕES UTILIZADAS PARA O MODELO CQV

Quadro 8 - Identificação das combinações entre mapa de atributos, CQPs e otimizadores para os modelos CQV.

Combinação	Mapa de atributos	CQPs	Otimizador
1	PauliFeatureMap-X	RealAmplitudes	COBYLA
2	PauliFeatureMap-X	RealAmplitudes	SPSA
3	PauliFeatureMap-X	EfficientSU2	COBYLA
4	PauliFeatureMap-X	EfficientSU2	SPSA
5	PauliFeatureMap-X	TwoLocal	COBYLA
6	PauliFeatureMap-X	TwoLocal	SPSA
7	ZFeatureMap	RealAmplitudes	COBYLA
8	ZFeatureMap	RealAmplitudes	SPSA
9	ZFeatureMap	EfficientSU2	COBYLA
10	ZFeatureMap	EfficientSU2	SPSA
11	ZFeatureMap	TwoLocal	COBYLA
12	ZFeatureMap	TwoLocal	SPSA
13	ZZFeatureMap	RealAmplitudes	COBYLA
14	ZZFeatureMap	RealAmplitudes	SPSA
15	ZZFeatureMap	EfficientSU2	COBYLA
16	ZZFeatureMap	EfficientSU2	SPSA
17	ZZFeatureMap	TwoLocal	COBYLA
18	ZZFeatureMap	TwoLocal	SPSA

Fonte: O autor (2024).

APÊNDICE C – MÉTRICAS DE PERFORMANCE DOS MODELOS CLÁSSICOS DE CLASSIFICAÇÃO PARA O CONJUNTO DE TESTE

Tabela 13 - Métricas dos modelos clássicos para o conjunto de teste do caso MnBrg.

Métrica	Nº de CPs	KNN				SVM			
		UM	UC	DM	DC	L	S	P	RBF
Acurácia balanceada	2	0,934	0,940	0,923	0,897	0,939	0,547	0,937	0,929
	3	0,934	0,927	0,935	0,926	0,946	0,533	0,939	0,932
	4	0,935	0,928	0,937	0,932	0,935	0,508	0,927	0,940
	5	0,928	0,930	0,924	0,922	0,941	0,467	0,920	0,940
	6	0,933	0,933	0,929	0,932	0,944	0,557	0,928	0,935
Precisão	2	0,955	0,943	0,941	0,905	0,955	0,629	0,938	0,964
	3	0,955	0,937	0,959	0,941	0,965	0,619	0,938	0,969
	4	0,959	0,950	0,960	0,946	0,959	0,598	0,945	0,969
	5	0,950	0,950	0,945	0,936	0,964	0,565	0,941	0,969
	6	0,950	0,950	0,950	0,946	0,969	0,622	0,941	0,959
Especificidade	2	0,936	0,915	0,915	0,858	0,936	0,461	0,908	0,950
	3	0,936	0,908	0,943	0,915	0,950	0,454	0,908	0,957
	4	0,943	0,929	0,943	0,922	0,943	0,418	0,922	0,957
	5	0,929	0,929	0,922	0,908	0,950	0,355	0,915	0,957
	6	0,929	0,929	0,929	0,922	0,957	0,177	0,915	0,943
Recall	2	0,931	0,966	0,931	0,936	0,941	0,632	0,966	0,907
	3	0,931	0,946	0,926	0,936	0,946	0,613	0,971	0,907
	4	0,926	0,926	0,931	0,941	0,926	0,598	0,931	0,922
	5	0,926	0,931	0,926	0,936	0,931	0,578	0,931	0,922
	6	0,936	0,936	0,936	0,941	0,931	0,936	0,941	0,926
F1-score	2	0,943	0,954	0,936	0,920	0,948	0,631	0,952	0,934
	3	0,943	0,941	0,943	0,939	0,955	0,616	0,954	0,937
	4	0,943	0,938	0,945	0,943	0,943	0,598	0,938	0,945
	5	0,938	0,941	0,936	0,936	0,948	0,571	0,936	0,945
	6	0,943	0,943	0,943	0,943	0,950	0,748	0,941	0,943

Fonte: O autor (2024).

Tabela 14 - Métricas dos modelos clássicos para o conjunto de teste do caso GbxHss.

Métrica	Nº de CPs	KNN				SVM			
		UM	UC	DM	DC	L	S	P	RBF
Acurácia balanceada	2	0,865	0,817	0,870	0,794	0,845	0,421	0,858	0,843
	3	0,921	0,848	0,919	0,819	0,870	0,438	0,873	0,872
	4	0,922	0,895	0,915	0,880	0,858	0,421	0,918	0,921
	5	0,924	0,922	0,921	0,918	0,864	0,461	0,922	0,924
	6	0,923	0,911	0,919	0,909	0,878	0,465	0,922	0,921
Precisão	2	0,881	0,869	0,893	0,851	0,863	0,548	0,864	0,850
	3	0,948	0,866	0,948	0,853	0,888	0,565	0,883	0,880
	4	0,956	0,939	0,948	0,918	0,904	0,553	0,964	0,968
	5	0,964	0,956	0,960	0,952	0,919	0,583	0,964	0,973
	6	0,952	0,948	0,948	0,947	0,932	0,590	0,956	0,960
Especificidade	2	0,805	0,805	0,831	0,779	0,773	0,292	0,766	0,740
	3	0,922	0,779	0,922	0,766	0,818	0,260	0,805	0,799
	4	0,935	0,909	0,922	0,877	0,857	0,234	0,948	0,955
	5	0,948	0,935	0,942	0,929	0,883	0,247	0,948	0,961
	6	0,929	0,922	0,922	0,922	0,903	0,084	0,935	0,942
Recall	2	0,925	0,829	0,908	0,808	0,917	0,550	0,950	0,946
	3	0,921	0,917	0,917	0,871	0,921	0,617	0,942	0,946
	4	0,908	0,892	0,908	0,883	0,858	0,608	0,888	0,888
	5	0,900	0,908	0,900	0,908	0,846	0,675	0,896	0,888
	6	0,917	0,904	0,917	0,896	0,854	0,846	0,908	0,900
F1-score	2	0,902	0,849	0,901	0,829	0,889	0,549	0,905	0,895
	3	0,934	0,891	0,932	0,862	0,904	0,590	0,911	0,912
	4	0,932	0,915	0,928	0,900	0,880	0,579	0,924	0,926
	5	0,931	0,932	0,929	0,930	0,881	0,625	0,929	0,928
	6	0,934	0,925	0,932	0,921	0,891	0,695	0,932	0,929

Fonte: O autor (2024).

Tabela 15 - Métricas dos modelos clássicos para o conjunto de teste do caso GnNDE.

Métrica	Nº de CPs	KNN				SVM			
		UM	UC	DM	DC	L	S	P	RBF
Acurácia balanceada	2	0,943	0,834	0,941	0,817	0,932	0,494	0,940	0,938
	3	0,974	0,869	0,980	0,880	0,957	0,380	0,952	0,957
	4	0,980	0,962	0,985	0,965	0,957	0,459	0,957	0,957
	5	0,985	0,985	0,985	0,985	0,957	0,478	0,968	0,969
	6	0,987	0,987	0,987	0,987	0,962	0,517	0,978	0,962
Precisão	2	0,975	0,930	0,975	0,928	0,970	0,805	0,972	0,973
	3	0,989	0,947	0,992	0,953	0,981	0,761	0,981	0,978
	4	0,992	0,983	0,994	0,986	0,981	0,792	0,981	0,981
	5	0,994	0,994	0,994	0,994	0,981	0,800	0,986	0,986
	6	0,994	0,994	0,994	0,994	0,983	0,813	0,983	0,992
Especificidade	2	0,894	0,694	0,894	0,694	0,871	0,000	0,882	0,882
	3	0,953	0,776	0,965	0,800	0,918	0,000	0,906	0,918
	4	0,965	0,929	0,976	0,941	0,918	0,082	0,918	0,918
	5	0,976	0,976	0,976	0,976	0,918	0,012	0,941	0,941
	6	0,976	0,976	0,976	0,976	0,929	0,059	0,965	0,929
Recall	2	0,992	0,975	0,989	0,941	0,994	0,989	0,997	0,994
	3	0,994	0,961	0,994	0,961	0,997	0,761	0,997	0,997
	4	0,994	0,994	0,994	0,989	0,997	0,837	0,997	0,997
	5	0,994	0,994	0,994	0,994	0,997	0,944	0,994	0,997
	6	0,997	0,997	0,997	0,997	0,994	0,977	0,992	0,994
F1-score	2	0,983	0,952	0,982	0,934	0,982	0,887	0,985	0,983
	3	0,992	0,954	0,993	0,957	0,989	0,761	0,987	0,989
	4	0,993	0,989	0,994	0,987	0,989	0,814	0,989	0,989
	5	0,994	0,994	0,994	0,994	0,989	0,866	0,990	0,992
	6	0,996	0,996	0,996	0,996	0,989	0,887	0,992	0,989

Fonte: O autor (2024).

**APÊNDICE D – MÉTRICA DE PERFORMANCE DOS MODELOS CLÁSSICOS DE
AGRUPAMENTO**

Tabela 16 - Resultado da IMN para os modelos clássicos de agrupamento.

Caso de estudo	Nº de CPs	<i>k-means</i>	AE		
			P	RBF	S
MnBrg	2	0,597	0,623	0,470	0,684
	3	0,597	0,609	0,495	0,676
	4	0,548	0,639	0,503	0,640
	5	0,587	0,595	0,500	0,594
	6	0,592	0,595	0,503	0,638
GbxHss	2	0,509	0,219	0,399	0,120
	3	0,526	0,230	0,048	0,112
	4	0,511	0,404	0,051	0,135
	5	0,511	0,418	0,395	0,324
	6	0,511	0,439	0,411	0,272
GnNDE	2	0,741	0,686	0,635	0,390
	3	0,685	0,614	0,667	0,338
	4	0,694	0,665	0,663	0,415
	5	0,005	0,665	0,691	0,367
	6	0,680	0,665	0,726	0,634

Fonte: O autor (2024).

**APÊNDICE E – MÉTRICA DE PERFORMANCE DOS MODELOS QUÂNTICOS DE
QSVM PARA O CONJUNTO DE TESTE**

Tabela 17 - Métricas dos modelos de QSVM para o conjunto de teste do caso MnBrg.

Métrica	Nº de CPs	Mapa de atributos					
		Angular	Amplitude	X	Y	Z	ZZ
Precisão	2	0,954	0,963	0,955	0,591	0,949	0,957
	3	0,954	0,959	0,965	0,591	0,964	0,969
	4	0,974	0,974	0,969	0,591	0,964	0,934
	5	0,969	0,974	0,969	0,591	0,964	0,944
	6	0,974	0,974	0,950	0,591	0,941	0,935
Especificidade	2	0,936	0,950	0,936	0,000	0,929	0,943
	3	0,936	0,943	0,950	0,000	0,950	0,957
	4	0,965	0,965	0,957	0,000	0,950	0,908
	5	0,957	0,965	0,957	0,000	0,950	0,922
	6	0,965	0,965	0,929	0,000	0,915	0,908
Recall	2	0,922	0,902	0,941	1,000	0,917	0,873
	3	0,922	0,917	0,941	1,000	0,912	0,907
	4	0,902	0,907	0,926	1,000	0,917	0,897
	5	0,907	0,917	0,926	1,000	0,926	0,902
	6	0,912	0,912	0,936	1,000	0,936	0,912
F1-score	2	0,938	0,932	0,948	0,743	0,933	0,913
	3	0,938	0,937	0,953	0,743	0,937	0,937
	4	0,936	0,939	0,947	0,743	0,940	0,915
	5	0,937	0,944	0,947	0,743	0,945	0,922
	6	0,942	0,942	0,943	0,743	0,939	0,923

Fonte: O autor (2024).

Tabela 18 - Métricas dos modelos de QSVM para o conjunto de teste do caso GbxHss.

Métrica	Nº de CPs	Mapa de atributos					
		Angular	Amplitude	X	Y	Z	ZZ
Precisão	2	0,854	0,851	0,854	0,609	0,860	0,834
	3	0,872	0,820	0,876	0,609	0,890	0,859
	4	0,924	0,874	0,972	0,609	0,935	0,931
	5	0,942	0,950	0,968	0,609	0,936	0,944
	6	0,955	0,941	0,960	0,609	0,931	0,919
	Especificidade	2	0,753	0,753	0,753	0,000	0,760
3		0,786	0,669	0,792	0,000	0,818	0,760
4		0,883	0,786	0,961	0,000	0,903	0,896
5		0,916	0,929	0,955	0,000	0,903	0,916
6		0,935	0,916	0,942	0,000	0,896	0,877
Recall		2	0,929	0,904	0,929	1,000	0,946
	3	0,938	0,971	0,942	1,000	0,942	0,942
	4	0,908	0,950	0,883	1,000	0,904	0,900
	5	0,888	0,871	0,883	1,000	0,908	0,908
	6	0,892	0,858	0,908	1,000	0,904	0,900
	F1-score	2	0,890	0,877	0,890	0,757	0,901
3		0,904	0,889	0,908	0,757	0,915	0,899
4		0,916	0,910	0,926	0,757	0,919	0,915
5		0,914	0,909	0,924	0,757	0,922	0,926
6		0,922	0,898	0,934	0,757	0,918	0,909

Fonte: O autor (2024).

Tabela 19 - Métricas dos modelos de QSVM para o conjunto de teste do caso GnNDE.

Métrica	Nº de CPs	Mapa de atributos					
		Angular	Amplitude	X	Y	Z	ZZ
Precisão	2	0,970	0,948	0,972	0,807	0,934	0,962
	3	0,981	0,930	0,981	0,807	0,939	0,900
	4	0,981	0,952	0,983	0,807	0,954	0,935
	5	0,981	0,973	0,986	0,807	0,957	0,961
	6	0,983	0,983	0,983	0,807	0,972	0,948
Especificidade	2	0,871	0,776	0,882	0,000	0,706	0,835
	3	0,918	0,694	0,918	0,000	0,729	0,541
	4	0,918	0,788	0,929	0,000	0,800	0,718
	5	0,918	0,882	0,941	0,000	0,812	0,835
	6	0,929	0,929	0,929	0,000	0,882	0,776
Recall	2	0,994	0,980	0,994	1,000	0,992	0,992
	3	0,997	0,966	0,997	1,000	0,992	0,986
	4	0,997	1,000	0,997	1,000	0,992	0,980
	5	0,997	0,997	0,997	1,000	0,992	0,983
	6	0,994	0,994	0,997	1,000	0,992	0,980
F1-score	2	0,982	0,964	0,983	0,893	0,962	0,976
	3	0,989	0,948	0,989	0,893	0,964	0,941
	4	0,989	0,975	0,990	0,893	0,972	0,957
	5	0,989	0,985	0,992	0,893	0,974	0,972
	6	0,989	0,989	0,990	0,893	0,982	0,964

Fonte: O autor (2024).

**APÊNDICE F – MÉTRICA DE PERFORMANCE DOS MODELOS QUÂNTICOS
CQV PARA O CONJUNTO DE TESTE**

Tabela 20 - Métricas dos modelos de CQV para o conjunto de teste.

Métrica	Nº de CPs	Caso de estudo					
		MnBrg		GbxHss		GnNDE	
		Média	Desvio	Média	Desvio	Média	Desvio
Acurácia balanceada	2	0,916	0,001	0,838	0,006	0,851	0,003
	3	0,915	0,002	0,846	0,005	0,946	0,005
	4	0,908	0,011	0,848	0,007	0,932	0,021
	5	0,899	0,004	0,845	0,008	0,921	0,036
	6	0,898	0,006	0,840	0,009	0,879	0,008
Precisão	2	0,931	0,000	0,864	0,004	0,935	0,001
	3	0,932	0,002	0,866	0,002	0,982	0,002
	4	0,929	0,016	0,854	0,012	0,976	0,011
	5	0,918	0,006	0,864	0,008	0,970	0,015
	6	0,917	0,012	0,860	0,015	0,953	0,003
Especificidade	2	0,901	0,000	0,779	0,007	0,714	0,006
	3	0,902	0,003	0,781	0,008	0,925	0,006
	4	0,898	0,026	0,746	0,028	0,901	0,047
	5	0,881	0,010	0,775	0,021	0,876	0,064
	6	0,879	0,019	0,765	0,040	0,804	0,011
<i>Recall</i>	2	0,932	0,004	0,897	0,012	0,988	0,001
	3	0,928	0,006	0,910	0,018	0,966	0,011
	4	0,919	0,010	0,950	0,018	0,963	0,018
	5	0,916	0,005	0,916	0,030	0,965	0,009
	6	0,918	0,009	0,916	0,037	0,954	0,007
F1-score	2	0,932	0,002	0,880	0,006	0,961	0,001
	3	0,930	0,003	0,887	0,008	0,974	0,005
	4	0,924	0,007	0,899	0,004	0,970	0,008
	5	0,917	0,003	0,889	0,011	0,968	0,011
	6	0,917	0,004	0,886	0,011	0,954	0,004

Fonte: O autor (2024).