

**MARINHA DO BRASIL
DIRETORIA DE ENSINO DA MARINHA
CENTRO DE INSTRUÇÃO ALMIRANTE ALEXANDRINO**

**CURSO DE APERFEIÇOAMENTO AVANÇADO EM
PROPULSÃO NAVAL**

TRABALHO DE CONCLUSÃO DE CURSO

**IDENTIFICAÇÃO DE FALHAS EM MÁQUINAS ROTATIVAS: *Machine Learning* e
Marinha do Brasil.**



1ºTen FELIPE ALVES DA SILVA

Rio de Janeiro
2023

1ºTen FELIPE ALVES DA SILVA

IDENTIFICAÇÃO DE FALHAS EM MÁQUINAS ROTATIVAS: *Machine Learning* e
Marinha do Brasil.

Monografia apresentada ao Centro de Instrução
Almirante Alexandrino como requisito parcial à
conclusão do Curso de Aperfeiçoamento Avançado
em Propulsão Naval.

Orientadores:

Dionísio Henrique Carvalho de Sá Só Martins, D.Sc

CC (EN) Júlio César Silva Pontes

Ulisses Admar Barbosa Vicente Monteiro, D.Sc

CIAA
Rio de Janeiro
2023

1ºTen FELIPE ALVES DA SILVA

IDENTIFICAÇÃO DE FALHAS EM MÁQUINAS ROTATIVAS: *Machine Learning*
e Marinha do Brasil.

Monografia apresentada ao Centro de Instrução Almirante Alexandrino como requisito parcial à conclusão do Curso de Aperfeiçoamento Avançado em Propulsão Naval.

Aprovada em _____

Banca Examinadora:

Dionísio Henrique Carvalho de Sá Só Martins, D.Sc – CECIERJ

Ulisses Admar Barbosa Vicente Monteiro, D.Sc – UFRJ

Júlio César Silva Pontes, M.Sc – DEN

Carlos Alfredo Orfão Martins, M.Sc – CIAA

CIAA
Rio de Janeiro
2023

AGRADECIMENTOS

Ao concluir meu trabalho de conclusão do Curso de Aperfeiçoamento Avançado em Propulsão Naval, não posso deixar de expressar minha profunda gratidão a todos aqueles que desempenharam um papel fundamental nessa jornada desafiadora e gratificante.

Primeiramente, quero agradecer à minha amada esposa. Seu apoio inabalável, paciência e encorajamento foram a âncora que me manteve firme em meio às tormentas pessoais e acadêmicas. Sem você, esse feito não teria sido possível, e sou eternamente grato por seu amor e apoio.

Aos meus pais e minha querida irmã, que sempre estiveram ao meu lado, fornecendo não apenas apoio emocional, mas também inspiração e encorajamento, agradeço do fundo do meu coração. Sua crença em mim foi um motor poderoso em minha jornada.

À coordenação do curso, aos orientadores, aos dedicados professores que compartilharam seu conhecimento e experiência, agradeço por sua orientação, paciência e apoio constante. Seus ensinamentos foram inestimáveis para o desenvolvimento do meu projeto.

Aos meus colegas de turma, quero expressar minha admiração e alegria por compartilhar essa jornada com vocês. Nossas discussões, colaborações e amizades tornaram essa experiência enriquecedora e memorável.

Este trabalho de conclusão de curso representa não apenas o meu esforço, mas um pouco de cada um que participou dessa jornada comigo.

À medida que encerro este capítulo e me preparo para novos desafios, levarei comigo não apenas o conhecimento adquirido, mas também a lembrança de como o apoio da família, dos amigos e dos mentores é inestimável.

IDENTIFICAÇÃO DE FALHAS EM MÁQUINAS ROTATIVAS: *Machine Learning* e Marinha do Brasil.

Resumo

O presente trabalho tem como objetivo prover um diagnóstico acurado para a identificação de falhas de desbalanceamento, desalinhamento horizontal e desalinhamento vertical em máquinas rotativas, a partir da utilização de técnicas de processamento de dados e *Machine Learning*, justificando a importância da implementação de procedimentos semelhantes na Marinha do Brasil.

Para isso foi utilizada uma base de dados com sinais de vibração adquiridos de uma bancada experimental, realizado o processamento dos dados adquiridos, a extração de característica desses dados, realizado o treinamento de dois algoritmos de inteligência artificial (*Adaptive Boosting tree* e Rede Neural Artificial) para efetuar o diagnóstico da máquina rotativa, execução de técnica de seleção de características com a finalidade de escolher os atributos com a maior capacidade de discriminação e, posteriormente, analisar e propor o algoritmo mais recomendado e os fatores que corroboram para a aplicação de técnicas de inteligência artificial na Marinha do Brasil.

Como resultado chegou-se a uma acurácia de 98,5% utilizando Rede Neural Artificial e 99,0% utilizando o *Adaptive Boosting tree* que, juntamente com a análise comparativa de outras métricas (especificidade, revocação, precisão, *F1-score* e velocidade de processamento) contribuem para a conclusão que, para o banco de dados utilizado, o algoritmo *Adaptive Boosting tree* é mais recomendado para a identificação de falhas de desbalanceamento, desalinhamento horizontal e desalinhamento vertical em máquinas rotativas.

Por fim, a partir da missão da Marinha do Brasil e do previsto no Plano Estratégico da Marinha, concluiu-se que a redução de custos de manutenção e o aumento da disponibilidade e da confiabilidade das máquinas rotativas são razões pelas quais se justifica a implementação de técnicas de aprendizagem de máquinas no sistema de manutenção da Marinha do Brasil.

Palavras-chave: *Machine Learning*, manutenção preditiva, falhas de desbalanceamento e desalinhamento, máquinas rotativas.

LISTA DE FIGURAS

Figura 1 - Resumo da metodologia utilizada.....	14
Figura 2 - <i>K</i> -fold com 5 partições.....	17
Figura 3 - Modelo de matriz confusão.....	18
Figura 4 - Bancada experimental ABVT do LEDAV/COPPE/UFRJ.....	20
Figura 5 - Identificação dos eixos de medição.....	21
Figura 6 - Estrutura de uma rede neural artificial.....	26
Figura 7 - Treinamento de um classificador AdaBoost.....	27
Figura 8 - Estrutura da árvore de decisão.....	28
Figura 9 - Sintonização de hiperparâmetros.....	28
Figura 10 - Matriz confusão para o classificador RNA com 42 características.....	30
Figura 11 - Matriz confusão para o classificador Adaboost tree com 42 características.....	30
Figura 12 - Matriz confusão do classificador RNA no teste com 24 características.....	34
Figura 13 - Matriz confusão do classificador Adaboost tree no teste com 30 características.....	34

LISTA DE TABELAS

Tabela 1 - Descrição do Objetivo Naval 7.	13
Tabela 2 - Métricas utilizadas para avaliação dos classificadores.	19
Tabela 3 - Descrição dos canais da COMFAULDA utilizados.	21
Tabela 4 - Cenários específicos utilizados da base de dados COMFAULDA. .	21
Tabela 5 - Cenários compilados utilizados da base de dados COMFAULDA. .	22
Tabela 6 - Aparições de características em estudos.	23
Tabela 7 - Definição das métricas estatísticas utilizadas.	24
Tabela 8 - Relevâncias das características selecionadas.	32

LISTA DE GRÁFICOS

Gráfico 1 - Resultado da seleção de características.	31
Gráfico 2 - Gráfico radar comparativo entre métricas dos classificadores RNA e Adaboost tree.	35

SUMÁRIO

1	INTRODUÇÃO.....	10
1.1	Objetivos gerais.....	11
1.2	Objetivos específicos.....	11
2	UTILIZAÇÃO NA MARINHA DO BRASIL.....	13
3	METODOLOGIA.....	14
4	REFERENCIAL TEÓRICO.....	15
4.1	Manutenção preditiva.....	15
4.2	Análise de vibração.....	15
4.3	<i>Machine Learning (ML)</i>	16
4.4	Validação cruzada.....	16
4.5	Extração de características.....	17
4.6	Matriz confusão.....	18
4.7	Métricas de desempenho.....	18
5	ESTRUTURAÇÃO DA BASE DE DADOS.....	20
5.1	COMFAULDA.....	20
6	EXTRAÇÃO DE CARACTERÍSTICAS.....	23
7	CLASSIFICADORES.....	25
7.1	Rede Neural Artificial (RNA).....	25
7.2	<i>Adaptive Boosting tree (AdaBoost tree)</i>	26
7.3	Sintonização de hiperparâmetros.....	28
8	DISCUSSÃO E RESULTADOS.....	29
9	CONCLUSÕES E RECOMENDAÇÕES.....	36

1 INTRODUÇÃO

As máquinas rotativas desempenham um papel fundamental na sociedade contemporânea, servindo como alicerces da produção mecanizada e autônoma, bem como fornecendo produtos de fácil acesso e rápido suprimento para atender às necessidades humanas. Uma interrupção inesperada de uma máquina rotativa devido a falha ou avaria não só pode perturbar a cadeia de fornecimento, mas também gerar enormes custos (Das, O.; Das; B.; Birant, 2023).

A redução na disponibilidade de equipamentos em embarcações navais, seja por falhas imprevistas ou manutenção programada, é igualmente indesejável. Para atender aos requisitos de disponibilidade e confiabilidade dessas embarcações, é necessário realizar investimentos significativos em manutenção, cujos valores tendem a crescer à medida que a complexidade das embarcações aumenta (Cullum; Binns; Lonsdale; Abbassi; Garaniya, 2018).

Assim, torna-se essencial a determinação de uma estratégia e elaborar um plano eficiente de manutenção. Nesse sentido, prognóstico e diagnóstico desempenham um papel fundamental nas estratégias contemporâneas de manutenção, uma vez que são amplamente reconhecidos que a identificação de falhas em curso e a estimativa do tempo remanescente antes da ocorrência de uma falha funcional são elementos essenciais.

No contexto das atividades de manutenção, os mantenedores dependem de muitas informações das máquinas, incluindo seus registros de falhas anteriores, métodos e diretrizes de reparo, bem como de novas pesquisas na área e, adicionalmente, o acesso a essas informações pode ser obtido por meios variados, o que pode afetar a qualidade das decisões de manutenção (Pascual, 2015).

Em geral, a tomada de decisões de manutenção é guiada por meio de processos não automatizados, o que introduz alguma incerteza na análise. Para minimizar esse problema novas metodologias de manutenção devem buscar a automação para garantir consistência e aumentar a eficiência do fluxo de trabalho (Cullum; Binns; Lonsdale; Abbassi; Garaniya, 2017). Nesse sentido, o uso de *Machine Learning* pode representar uma solução viável.

Através da coleta de dados históricos das variáveis operacionais de máquinas rotativas, é possível implementar uma estratégia de manutenção

preditiva, com enfoque na antecipação da ocorrência de falhas por meio do monitoramento de variáveis como temperatura, tensão, vibração e pressão. Um algoritmo é configurado para analisar esses dados e identificar irregularidades que podem indicar um comportamento atípico, que, se não corrigido pode levar a uma falha e a indisponibilidade do equipamento. Essa análise de dados baseada em algoritmos e modelos de aprendizado representa a abordagem de *Machine Learning* (Roque; Krebs; Figueiro; Jazdi, 2022).

Por fim, ao considerar a relevância do tema na atualidade, em termos financeiros, científicos, acadêmicos e por avaliar que para a Marinha do Brasil o assunto abordado tem muito a contribuir, visto a quantidade existente de máquinas rotativas tais como turbinas, bombas, ventiladores e compressores em seus meios, justifica-se a escolha do tema em questão.

1.1 Objetivos gerais

O objetivo geral deste trabalho de conclusão de curso concentra-se em, a partir de técnicas de processamento digital de sinais e inteligência artificial, prover um diagnóstico acurado para identificação de falhas de desbalanceamento, desalinhamento horizontal e desalinhamento vertical em máquinas rotativas e justificar a importância da implementação de procedimentos semelhantes no sistema de manutenção preditiva dos meios da Marinha do Brasil.

1.2 Objetivos específicos

Os objetivos específicos deste trabalho estão relacionados a seguir:

- utilização de base de dados de referência de sinais de vibrações a partir de uma bancada experimental de máquina rotativa, considerando diversos cenários de funcionamento;
- aplicação de algoritmos de inteligência artificial para efetuar um diagnóstico eficaz do estado de uma máquina rotativa em operação;
- execução de técnicas de processamento digital de sinais para realização de extração de características de sinais de vibração;
- execução de técnica de seleção de características com a finalidade de escolher os atributos com maior capacidade de discriminação;

- análise e proposição do algoritmo mais recomendado, dentre os estudados, para a identificação das condições de funcionamento das máquinas rotativas; e

- análise dos fatores que corroboram para a aplicação de técnicas de inteligência artificial na manutenção preditiva da Marinha do Brasil.

2 UTILIZAÇÃO NA MARINHA DO BRASIL

A Missão da Marinha do Brasil (MB) foi atualizada em 2016 e teve como condicionantes o artigo 142 da Constituição Federal (CF) e a Lei Complementar nº 97/99, sendo estabelecido o seguinte enunciado:

Preparar e empregar o Poder Naval, a fim de contribuir para a Defesa da Pátria; para a garantia dos poderes constitucionais e, por iniciativa de qualquer destes, da lei e da ordem; para o cumprimento das atribuições subsidiárias previstas em Lei; e para o apoio à Política Externa.

Nessa missão, a palavra “preparar” explicita a relevância da manutenção dos meios navais para o cumprimento do propósito da Marinha do Brasil. Nesse sentido a MB em seu Plano Estratégico estabelece o previsto na Tabela 1.

Tabela 1 - Descrição do Objetivo Naval 7.

Objetivo Naval 7 (OBNAV 7) – Obter a capacidade operacional plena (OCOP)	
Estratégia Naval 7.1 – OCOP	
Visa contribuir para a melhora da disponibilidade dos meios navais, aeronavais e de fuzileiros navais existentes por meio de processos de modernização e revitalização exequíveis e com custo adequado à realidade orçamentária da MB e incrementar o poder combatente por meio de modernização dos sistemas de combate.	
Ações Estratégicas Navais –	Descrição:
OCOP-1: Manutenir/modernizar os meios de superfície, submarinos, aeronavais e de Fuzileiros Navais existentes na MB.	Manter, no mínimo, a disponibilidade de 65% dos navios, submarinos, aeronaves e material de combate de Fuzileiros Navais de forma a cumprirem a Missão da Marinha.

Fonte: (Brasil, 2020).

Ao observar o Objetivo Naval 7 é possível analisar novamente a importância dada para a manutenção dos meios navais e que, conforme a Estratégia Naval 7.1 existe a necessidade que os custos sejam adequados à realidade orçamentária da MB.

3 METODOLOGIA

A fase inicial deste estudo envolveu a condução de uma revisão bibliográfica abrangente sobre o tema proposto. Esta revisão priorizou a análise de artigos relacionados a manutenção preditiva baseada na análise de sinais de vibração, diagnóstico de falhas, extração de características e técnicas e classificação de padrões, com o objetivo de descrever o assunto a partir de estudos atualizados e reconhecidos.

A próxima etapa consistiu em, a partir de uma base de dados de sinais de vibração de máquinas rotativas, contendo 4 cenários distintos (operação normal, desbalanceamento, desalinhamento horizontal e desalinhamento vertical), realizar a extração de características desses sinais. Para isso foi necessário primeiramente, identificar, através da revisão bibliográfica, quais as características de sinais de vibração são mais utilizadas para esse tipo de análise.

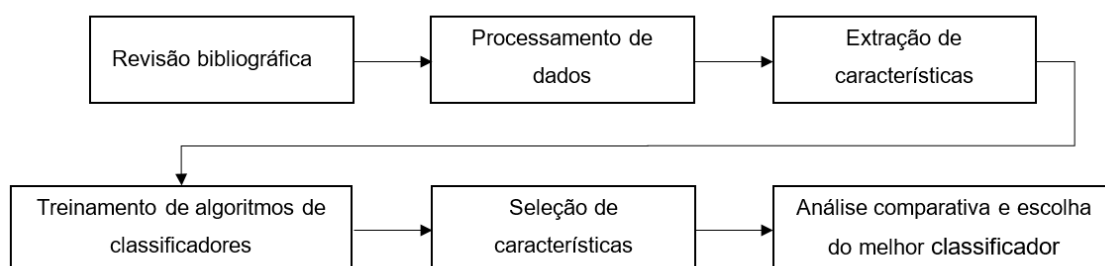
Posteriormente, dois algoritmos classificadores foram utilizados, a partir da aplicação *Classification Learner* do MATLAB, a fim de verificar seus desempenhos na identificação dos cenários existentes.

Com a finalidade de melhorar os resultados obtidos anteriormente, a quarta etapa consistiu na execução de técnica de seleção de características. Esse tipo de técnica propicia identificar quais as características são mais relevantes para a base de dados em análise.

A partir dos resultados obtidos foi possível fazer uma análise comparativa entre os dois algoritmos e concluir sobre a eficácia do método.

A Figura 1 resume as etapas metodologia utilizada.

Figura 1 - Resumo da metodologia utilizada.



Fonte: autoria própria.

4 REFERENCIAL TEÓRICO

Nesta seção, serão expostos os fundamentos teóricos que estabelecem a base para a descrição da pesquisa realizada neste trabalho. Serão abordados os princípios referentes à manutenção preditiva, à análise de vibrações em máquinas rotativas, ao processo de extração e seleção de características, à validação cruzada, bem como ao funcionamento dos algoritmos de classificação adotados nesta pesquisa.

4.1 Manutenção preditiva

A manutenção preditiva utiliza uma estimativa da condição atual e futura de um componente para realizar um planejamento de manutenção otimizado que previne falhas sem recorrer em manutenção excessiva. Esse planejamento proporciona prolongamento da vida útil da máquina, juntamente com redução dos custos de manutenção e tempo de inatividade. Para isso, é necessário que, para cada componente, se considere a condição presente e passada a fim de que se ofereça um plano de manutenção personalizado (Verbert; Schutter; Babuška, 2016).

Para estimar a condição passada e presente de componentes e máquinas, podem ser realizadas medições de diferentes tipos, a fim de terem seus dados adquiridos e processados, sendo as principais: medições de vibração, medições acústicas, análise de óleo, termografia, medições de deformação e medições de torque e momento fletor (Hossain; Abu-Siada; Muyeen, 2018).

A partir dos sinais de vibração é possível extrair diversas características que apresentam capacidade discriminativa, o que faz com que esta técnica seja uma das mais amplamente utilizadas para o diagnóstico de falhas. O objetivo de realizar esse tipo de medição é relacionar corretamente o sinal de vibração adquirido com a condição correspondente por meio de um algoritmo de classificação (Ahmed; Nandi, 2018).

4.2 Análise de vibração

Mesmo em boas condições, as máquinas geram vibrações. Muitas dessas vibrações estão diretamente ligadas a eventos periódicos na operação da

máquina, como eixos rotativos, engrenagens, campos elétricos rotativos, e assim por diante (Randall, 2011).

A análise de vibração tem sido usada com sucesso por muitos anos para monitorar equipamentos como um todo, medindo a vibração da máquina ou analisando seus componentes individualmente. Essas vibrações são utilizadas para acompanhar a tendência da máquina e de seus componentes, rastreando quaisquer mudanças que possam surgir. Essas alterações indicam possíveis problemas associados à máquina e exigem monitoramento adicional até que a manutenção possa ser realizada (Pascual, 2015).

Nesse sentido, técnicas *Machine Learning* (ML) têm sido muito bem-sucedidas em simplificar a construção de modelos de prognóstico para avaliação da condição de máquinas com base em dados históricos disponíveis provenientes de sistemas similares ou modelos físicos específicos (Berghout; Mouss; Bentrícia; Elbouchikhi; Benbouzid, 2021).

4.3 *Machine Learning* (ML)

A principal função do *Machine Learning* ou aprendizado de máquina é projetar e desenvolver algoritmos que permitam que sistemas, usando dados empíricos, experiência e treinamento, evoluam e se adaptem às mudanças em seu ambiente, ou seja, às mudanças no estado dos ativos monitorados. Uma parte importante na investigação do aprendizado de máquina é induzir automaticamente modelos, como regras e padrões, a partir dos dados de treinamento que estão sendo analisados (Pascual, 2015).

Algoritmos de ML analisam uma grande quantidade de dados por meio de métodos estatísticos, como o método de classificação, como o abordado no presente estudo, sendo possível identificar padrões no banco de dados e, em seguida, fazer previsões (Roque; Krebs; Figueiro; Jazdi, 2022).

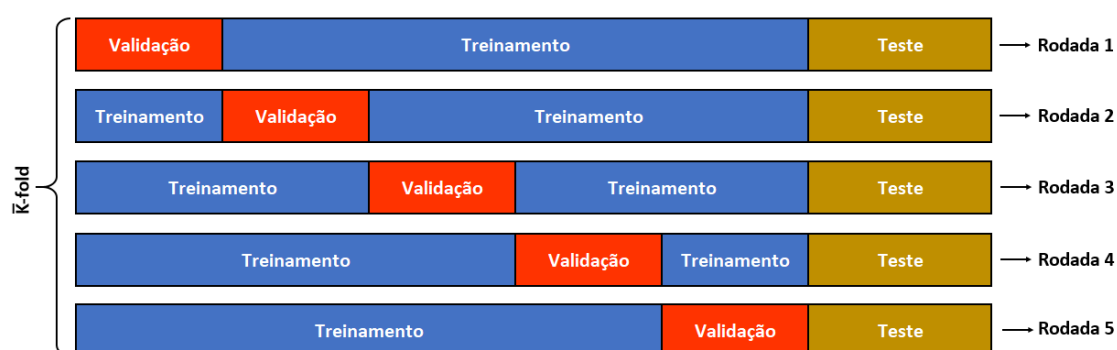
4.4 Validação cruzada

A validação cruzada é um método de amostragem iterativa, em que a cada iteração, o conjunto de dados original é dividido em um conjunto de treinamento e um conjunto de teste. A divisão pode ser feita de forma aleatória ou sistemática, conforme Figura 2.

Após a divisão dos dados, o conjunto de treinamento é usado para treinar um algoritmo, que é então aplicado ao conjunto de teste, resultando em uma estrutura de matrizes ou vetores com os valores de resposta previstos. Os resultados obtidos em todas as \bar{K} iterações são combinados em características de desempenho ou usados conforme necessário (Kucheryavskiy; Rodionova; Pomerantsev, 2023).

Os valores mais comuns de \bar{K} são 5 e 10, sendo os resultados sensíveis ao valor de \bar{K} . No entanto, o custo computacional aumenta quando \bar{K} é excessivamente grande. A precisão do modelo é calculada pela média da precisão alcançada em cada iteração (Chacon; Ramírez; Marquez, 2023). No presente trabalho foi escolhido o valor de $\bar{K} = 5$, devido à quantidade reduzida de dados.

Figura 2 - \bar{K} -fold com 5 partições.



Fonte: autoria própria.

Essa técnica permite uma avaliação mais robusta dos modelos, treinando e testando-os em múltiplos subconjuntos dos dados (Mahesh; V. Kumar; D. Kumar; Geman; Margala; Guduri, 2023).

4.5 Extração de características

A extração de características é amplamente reconhecida como a parte mais essencial no processo de aprendizado de máquina. Os dados brutos muitas vezes são extensos, contendo informações redundantes e irrelevantes. Na maioria das vezes, estão incompletos, com registros sujos, ausentes e ruidosos. A extração de características normalmente consiste em uma série de procedimentos de processamento de dados, incluindo padronização, normalização, aprimoramento de sinais, extração de características locais e

discretização de características. Um vetor de características, com menos dimensões que a matriz inicial pode então ser criado, sem que se perca informações importantes. Esse vetor posteriormente será utilizado como atributo de entrada nos algoritmos de aprendizagem de máquina (J. Li et al., 2023).

4.6 Matriz confusão

É uma tabela usada para avaliar o desempenho de um algoritmo de classificação em um conjunto de dados para o qual os valores verdadeiros são conhecidos. Resume o número de previsões corretas e incorretas feitas pelo algoritmo, para cada classe no problema (Bhattarai; Maraseni; Devkota; Apan, 2023), conforme demonstra a Figura 3.

Figura 3 - Modelo de matriz confusão.

		Condição Real	
		Positivo	Negativo
Classificação do Modelo	Positivo	Verdadeiro positivo (VP)	Falso Positivo (FP)
	Negativo	Falso Negativo (FN)	Verdadeiro Negativo (VN)

Fonte: Adaptado de (Bhattarai; Maraseni; Devkota; Apan, 2023)

4.7 Métricas de desempenho

Em geral, as métricas de desempenho podem ser descritas como ferramentas de avaliação do desempenho do classificador. Métricas diferentes avaliam características diferentes do algoritmo, como por exemplo sua capacidade de generalização. Tais métricas podem ser utilizadas também como avaliadores para a seleção de modelos ou da melhor solução entre as geradas durante um treinamento de classificação. A Tabela 2 mostra a descrição das métricas de desempenho utilizadas no presente trabalho.

Tabela 2 - Métricas utilizadas para avaliação dos classificadores.

Métrica	Fórmula	Foco da métrica
Acurácia	$\frac{VP + VN}{VP + FP + VN + FN}$	A acurácia mede a proporção de previsões corretas em relação ao número total de instâncias avaliadas.
Especificidade	$\frac{VN}{VN + FP}$	Métrica usada para medir a fração de padrões negativos que são classificados corretamente.
Revocação	$\frac{VP}{VP + VN}$	Usada para medir a fração de padrões positivos que são classificados corretamente.
Precisão	$\frac{VP}{VP + FP}$	Usado para medir os padrões positivos que são corretamente previstos a partir do total de padrões previstos em uma classe positiva.
<i>F1-score</i>	$\frac{2 \cdot p \cdot r}{p + r}$	Essa métrica representa a média harmônica entre os valores de recall e precisão.
Velocidade de processamento	$\frac{\text{Quantidade de instâncias}}{\text{Tempo de processamento}}$	A métrica avalia a velocidade com que o algoritmo consegue classificar uma instância do conjunto de teste.

Fonte: Adaptado de (HOSSIN & SULAIMAN, 2015).

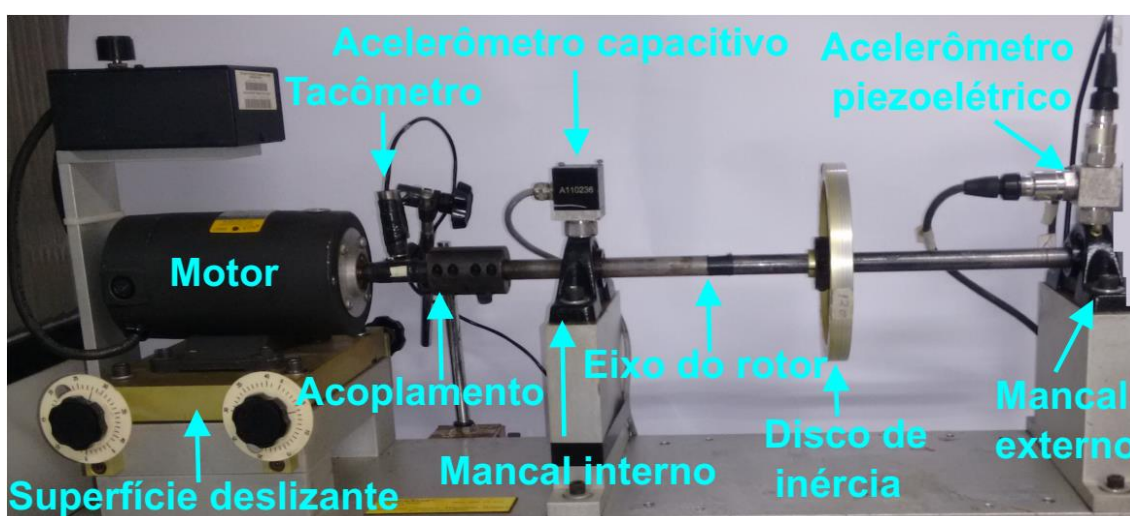
5 ESTRUTURAÇÃO DA BASE DE DADOS

Nesta seção, apresenta-se a base de dados empregada no presente estudo, a qual foi construída a partir de registros de defeitos induzidos em uma bancada experimental, apresentada na Figura 4. A utilização de uma bancada experimental física proporciona uma maior proximidade com as condições operacionais reais. Os defeitos analisados neste contexto consistem em falhas simples: desbalanceamento, desalinhamento horizontal e desalinhamento vertical (Martins, 2022).

5.1 COMFAULDA

No total foram utilizados 1019 sinais de vibração da base de dados denominada *Composed Fault Dataset* (COMFAULDA), que apresenta 4 cenários de operação diferentes da máquina rotativa: operação normal, desbalanceamento, desalinhamento vertical e desalinhamento horizontal.

Figura 4 - Bancada experimental ABVT do LEDAV/COPPE/UFRJ.



Fonte: Adaptado de (Martins, 2022).

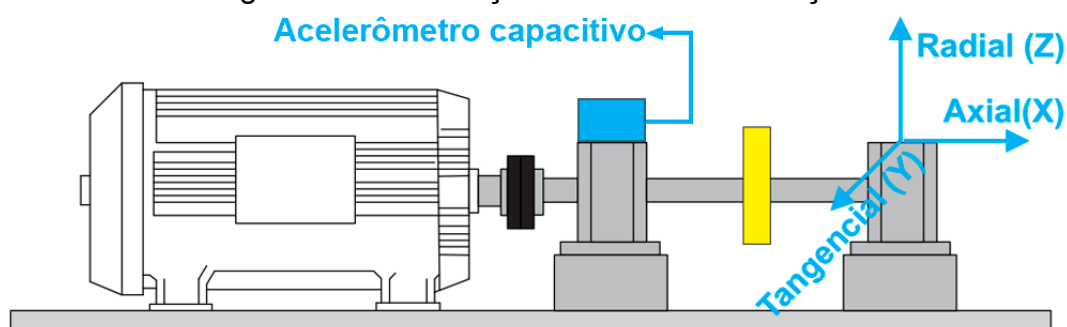
Os registros de vibração foram coletados ao longo de um período de 5 segundos, com uma taxa de amostragem $F_s = 50$ kHz. Conseqüentemente, cada entrada no conjunto de dados é composta por 250.000 amostras por canal. Cada aquisição inclui informações de três canais distintos (Martins, 2022), conforme detalhado na Tabela 3. Tais canais referem-se aos sinais do acelerômetro capacitivo nas direções axial, radial e tangencial, conforme Figura 5.

Tabela 3 - Descrição dos canais da COMFAULDA utilizados.

Canais	Descrição
Canal 3	Sinal do acelerômetro capacitivo da direção axial (X)
Canal 4	Sinal do acelerômetro capacitivo da direção radial (Z)
Canal 5	Sinal do acelerômetro capacitivo da direção tangencial (Y)

Fonte: Adaptado de (Martins, 2022).

Figura 5 - Identificação dos eixos de medição.



Fonte: Adaptado de (Martins, 2022).

As Tabela 4 e Tabela 5 mostram o número de sinais para cada cenário utilizado da base de dados.

Tabela 4 - Cenários específicos utilizados da base de dados COMFAULDA.

Cenários	Quantidade
Normal	245
Desbalanceamento (6 gramas)	48
Desbalanceamento (10 gramas)	48
Desbalanceamento (15 gramas)	48
Desbalanceamento (20 gramas)	48
Desbalanceamento (25 gramas)	48
Desbalanceamento (30 gramas)	48
Desbalanceamento (35 gramas)	48
Desalinhamento horizontal (0,5 mm)	48
Desalinhamento horizontal (1,0 mm)	49
Desalinhamento horizontal (1,5 mm)	49
Desalinhamento horizontal (2,0 mm)	49
Desalinhamento vertical (0,51 mm)	49
Desalinhamento vertical (1,27 mm)	48
Desalinhamento vertical (1,40 mm)	49
Desalinhamento vertical (1,78 mm)	48
Desalinhamento vertical (1,91 mm)	49

Fonte: Adaptado de (Martins, 2022).

Tabela 5 - Cenários compilados utilizados da base de dados COMFAULDA.

Cenários	Quantidade
Normal	245
Desbalanceamento	336
Desalinhamento horizontal	195
Desalinhamento vertical	243

Fonte: Adaptado de (Martins, 2022).

6 EXTRAÇÃO DE CARACTERÍSTICAS

Antes de iniciar o processo de extração de características, primeiramente realizou-se uma revisão bibliográfica a fim de verificar quais são as características mais usualmente utilizadas em cenários de identificação de falhas em máquinas. Para isso, foram selecionados 12 estudos, dos quais foram retiradas as 14 características com mais aparições, conforme apresentado na Tabela 6.

Tabela 6 - Aparições de características em estudos.

Métricas estatísticas	Aparições em artigos
Curtose	12
Assimetria	11
Fator de crista	10
Valor eficaz (RMS)	9
Fator de impulso	7
Média	6
Fator de forma	6
Desvio padrão	6
Valor de pico	5
Variância	4
Fator de margem	3
<i>Clearance factor</i>	2
Valor pico a pico	2
Mínimo	2

Fonte: autoria própria.

Após a análise de características relevantes para um cenário de identificação de falhas em máquinas, pode-se calcular os valores dessas características a partir do banco de dados. Os cálculos ocorrem de acordo com a Tabela 7.

A realização dos cálculos das características dos sinais é necessária para que os resultados sejam incluídos como entradas para os algoritmos classificadores, que serão temas da próxima seção deste trabalho.

Tabela 7 - Definição das métricas estatísticas utilizadas.

Métricas estatísticas	Definição
Máximo	$T_1 = \max(x(n))$
Mínimo	$T_2 = \min(x(n))$
Média	$T_3 = \sum_{n=1}^N \frac{x(n)}{N}$
Variância	$T_4 = \sum_{n=1}^N \frac{(x(n) - T_3)^2}{N - 1}$
Desvio padrão	$T_5 = \sqrt{T_4}$
Valor eficaz (RMS)	$T_6 = \sqrt{\sum_{n=1}^N \frac{(x(n))^2}{N}}$
Curtose	$T_7 = \frac{\sum_{n=1}^N (x(n) - T_3)^4}{(N - 1) \cdot T_5^4}$
Assimetria	$T_8 = \frac{\sum_{n=1}^N (x(n) - T_3)^3}{(N - 1) \cdot T_5^3}$
Fator de crista	$T_9 = \frac{\max x(n) }{T_6}$
Fator de impulso	$T_{10} = \frac{T_1}{\sum_{n=1}^N \frac{ x(n) }{N}}$
Fator de forma	$T_{11} = \frac{T_5}{\sum_{n=1}^N \frac{ x(n) }{N}}$
Fator de margem	$T_{12} = \frac{\max x(n) }{\left(\sum_{n=1}^N \frac{\sqrt{ x(n) }}{N}\right)^2}$
<i>Clearance factor</i>	$T_{13} = \frac{\max x(n) }{\frac{1}{N} \cdot \sum_{n=1}^N x(n) }$
Valor pico a pico	$T_{14} = T_1 - T_2$

Fonte: Adaptado de (Martins, 2022).

7 CLASSIFICADORES

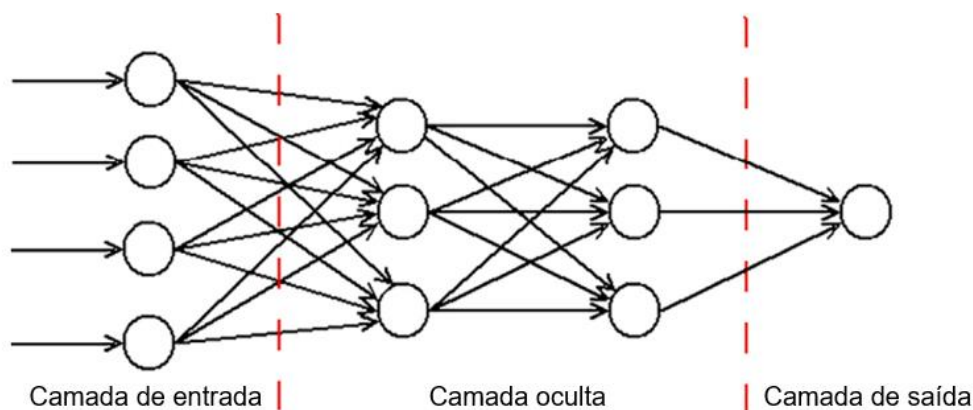
Algoritmos de aprendizado de máquina geralmente consomem e processam dados para aprender os padrões relacionados a indivíduos, processos de negócios, transações, eventos, entre outras situações. Os algoritmos de ML são principalmente divididos em quatro categorias: aprendizado supervisionado, aprendizado não supervisionado, aprendizado semi-supervisionado e aprendizado por reforço (Sarker, 2021).

Uma vez que o principal objetivo do estudo é identificar falhas em máquinas rotativas, utilizou-se técnicas de classificação supervisionada. O modelo de classificação é treinado com base em variáveis de entrada e aprende com os dados de treinamento para classificar as instâncias. O conjunto de dados foi dividido em um conjunto de treinamento (80%) e um conjunto de teste (20%). Na etapa de treinamento foi usado o procedimento de validação cruzada com *K-fold* de 5 partições, utilizando quatro partições para etapa de aprendizado e uma para a validação. O melhor modelo foi selecionado e aplicado ao conjunto de teste. Em seguida, dois modelos classificadores foram considerados para realizar o treinamento com os dados. Cada modelo de classificação é descrito nos itens 7.1 e 7.2.

7.1 Rede Neural Artificial (RNA)

Os modelos de redes neurais foram desenvolvidos para imitar a estrutura do cérebro humano. A ideia é tratar o cérebro como composto por elementos altamente interconectados (neurônios) que trabalham juntos para resolver problemas específicos. Os modelos de redes neurais podem ser descritos por um gráfico organizado de acordo com diferentes níveis: a camada de entrada, a camada oculta e a camada de saída (Giudici; Gramegna; Raffinetti, 2023), como exibido na Figura 6.

Figura 6 - Estrutura de uma rede neural artificial.



Fonte: Adaptado de (Giudici; Gramegna; Raffinetti, 2023).

Cada camada consiste em um número de neurônios. Os neurônios recebem entradas, multiplicam-nas pelos pesos de cada entrada e combinam os resultados da multiplicação. Em seguida, as multiplicações combinadas dos sinais e pesos são enviadas a uma função de transferência para gerar a saída do neurônio (Ahmed; Nandi, 2018).

7.2 Adaptive Boosting tree (AdaBoost tree)

Adaptive Boosting (AdaBoost) é um processo de aprendizado em conjunto que emprega uma abordagem iterativa para melhorar classificadores, aprendendo com seus erros. Cria um classificador poderoso combinando classificadores de baixo desempenho para obter um bom classificador de alta precisão. Nesse sentido, o *AdaBoost* é chamado de classificador adaptativo, pois melhora significativamente a eficiência do classificador. O *AdaBoost* é melhor utilizado para aprimorar o desempenho de árvores de decisão, em problemas de classificação binária, no entanto, é sensível a dados ruidosos e valores discrepantes (Sarker, 2021).

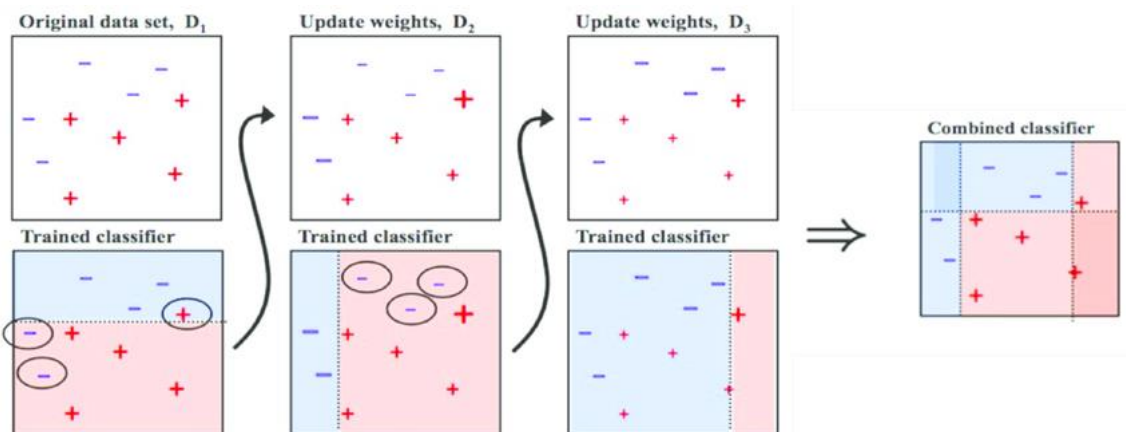
É um método geral que pode ser aplicado a vários classificadores, como árvores de decisão, para melhorar confiabilidade, desempenho e resistência ao sobreajuste. Cada árvore de decisão individual é chamada de aprendiz fraco no sentido de que é apenas um entre muitos classificadores. Então inicia-se o aumento adaptativo: cada aprendiz fraco é treinado iterativamente para melhorar em relação ao anterior.

O primeiro aprendiz fraco é treinado como uma árvore de decisão normal a partir dos dados de treinamento. No entanto, os resultados do primeiro

aprendiz fraco são então usados para ponderar a importância dos dados de treinamento para o próximo aprendiz fraco, pontos que foram classificados corretamente recebem pequenos pesos, enquanto pontos classificados incorretamente recebem pesos grandes.

Dessa forma, o próximo aprendiz fraco é treinado focando em pontos que não foram classificados corretamente pelo aprendiz fraco anterior. Esse processo continua, de modo que cada aprendiz fraco se concentra em corrigir os erros do último, melhorando a cada etapa (Marsh, 2016), conforme verificasse na Figura 7.

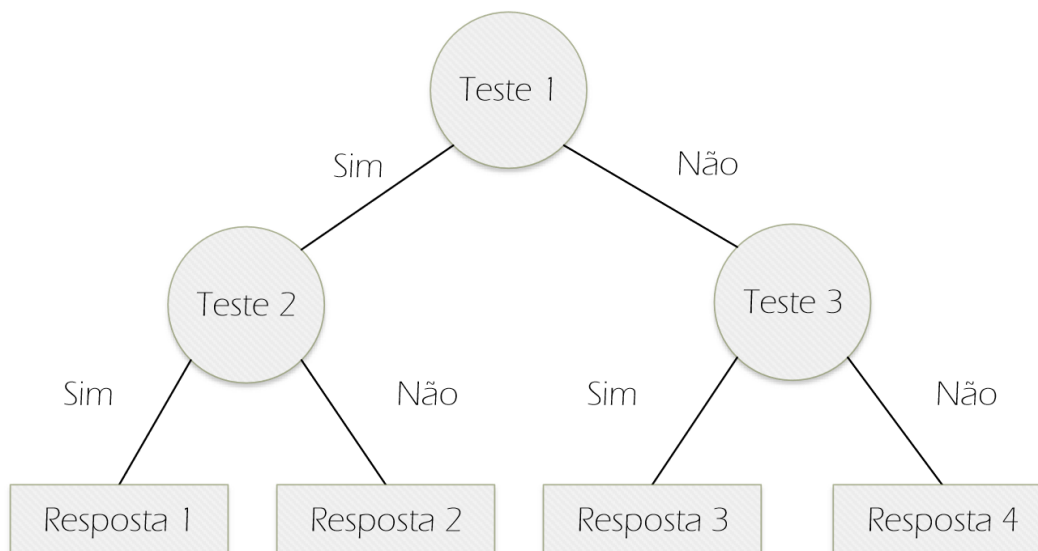
Figura 7 - Treinamento de um classificador AdaBoost.



Fonte: adaptado de (Marsh, 2016).

O classificador de árvore de decisão divide o conjunto de dados em subconjuntos menores em forma de uma estrutura de árvore para separar os dados com base em diferentes características, tendo o potencial para lidar com um grande volume de informações. Através dos métodos de árvore de decisão, nós e ramos são criados em forma de árvore, conforme Figura 8. Pode ser usado para resolver problemas de regressão e classificação. Cada nó na árvore atua como um caso de teste para alguma característica. Cada aresta do nó está associada às possíveis respostas do caso de teste. É um processo recursivo que se repete para cada subárvore no nó raiz. As árvores de decisão calculam o ganho de informação de cada árvore para dividir um nó (Gholamzadeh; Abtahi; Safdari, 2023).

Figura 8 - Estrutura da árvore de decisão.



Fonte: (Rolim; Cordeiro; Ferreira, 2014).

7.3 Sintonização de hiperparâmetros

Para a sintonização dos algoritmos foi utilizado o método de otimização Bayesiana que é uma forma de matematicamente encontrar um máximo ou um mínimo global de uma função.

A otimização Bayesiana se inicia com uma ideia sobre como a função que estamos tentando otimizar se comporta. À medida que os dados são observados, esse modelo é gradualmente ajustado. Ao longo do tempo verifica-se se esses ajustes foram satisfatórios, de forma a tentar minimizar a perda esperada. Após cada avaliação, o modelo é atualizado para torná-lo melhor para as futuras escolhas.

A principal ideia da otimização Bayesiana é criar um modelo que pode ser ajustado e consultado para auxiliar nas decisões de otimização (Shahriari et al., 2016). No presente trabalho os hiperparâmetros foram sintonizados acordo Figura 9.

Figura 9 - Sintonização de hiperparâmetros.

```

▼ Optimizer: Bayesian optimization
Optimizer: Bayesian optimization
Acquisition function: Expected improvement per second plus
Iterations: 10
Training time limit: false
  
```

Fonte: autoria própria.

8 DISCUSSÃO E RESULTADOS

Para realizar a análise dos resultados obtidos utilizou-se como principal parâmetro de medida de desempenho a acurácia, que pode ser definida como a porcentagem das classificações que obtiveram o resultado correto. Essa medida é frequentemente usada para avaliar problemas de classificação de aprendizado de máquina quando não há desbalanceamento de classes ou quando este desbalanceamento é pequeno, que é o caso deste trabalho (Manning; Raghavan; Schütze, 2009), podendo ser calculada através da fórmula apresentada na Tabela 2.

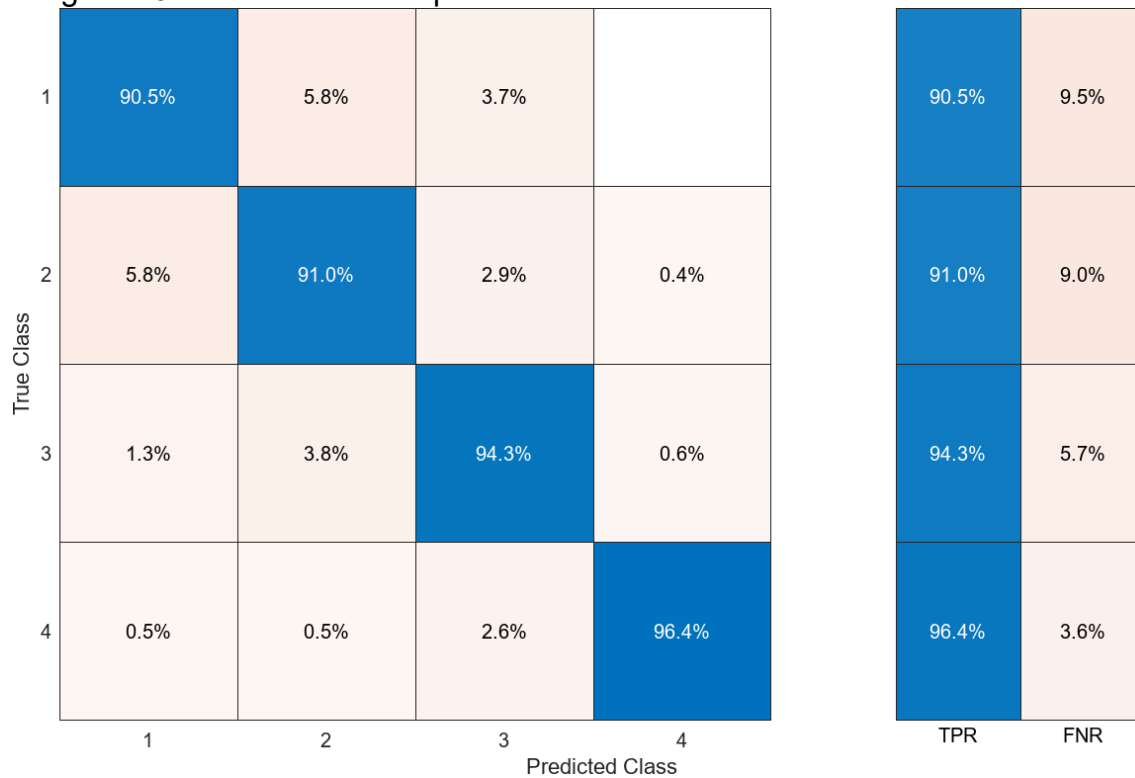
Inicialmente todas as 42 características foram utilizadas (14 características extraídas em três eixos) para realizar o treinamento dos modelos classificadores RNA e *Adaboost tree*. As Figura 5, Figura 10 e Figura 11 representam os resultados obtidos na validação dos modelos através de matrizes de confusão, que foram divididas em 4 classes. A classe 1 representa a condição normal de funcionamento, a classe 2 o desbalanceamento, a terceira classe o desalinhamento horizontal e a quarta o desalinhamento vertical.

Para esse cenário a performance em termos de acurácia foi de 92,8% com RNA e 91,7% com *Adaboost tree*.

Tendo em vista que o resultado obtido ainda não foi satisfatório optou-se pela utilização da técnica de seleção de características, com a finalidade de aumentar a acurácia dos modelos e avaliar as características mais relevantes para classificação destes cenários.

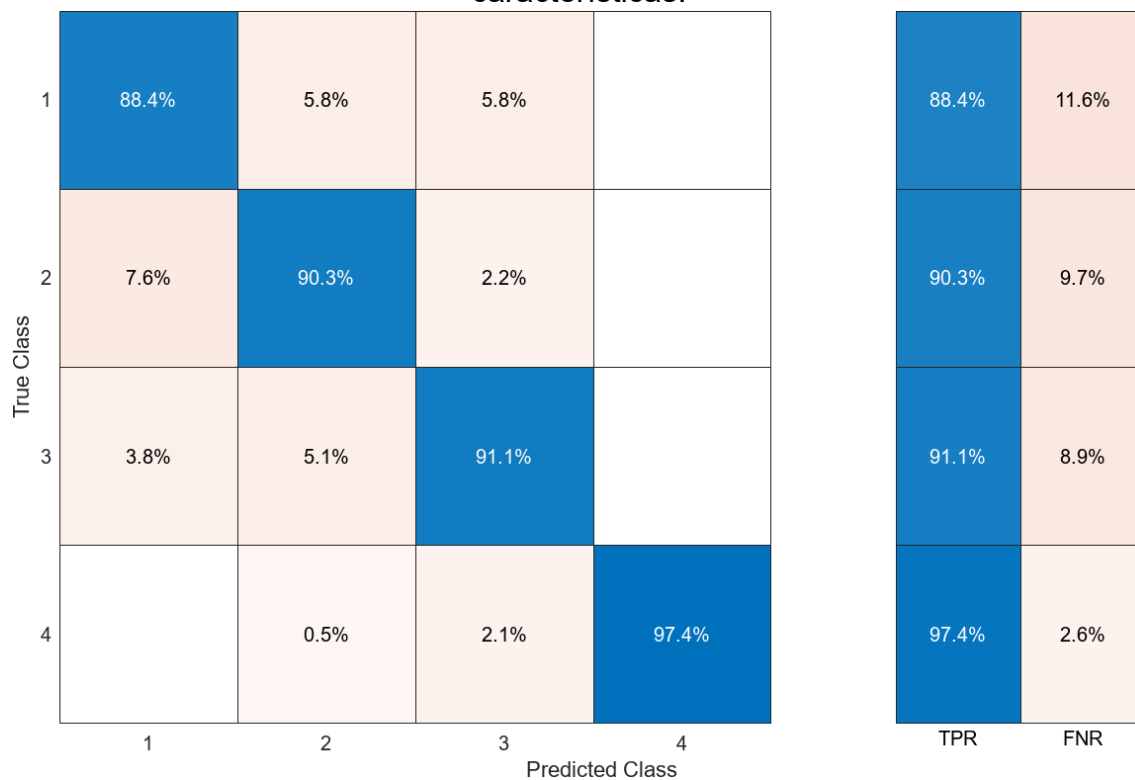
Para a execução da seleção de características foi utilizado o algoritmo RELIEFF, que é considerado um dos mais bem-sucedidos devido à sua simplicidade e eficácia. Este algoritmo tem como ideia principal estimar iterativamente os pesos dos atributos em relação a padrões de vizinhos mais próximos (Sun, 2007).

Figura 10 - Matriz confusão para o classificador RNA com 42 características.



Fonte: autoria própria.

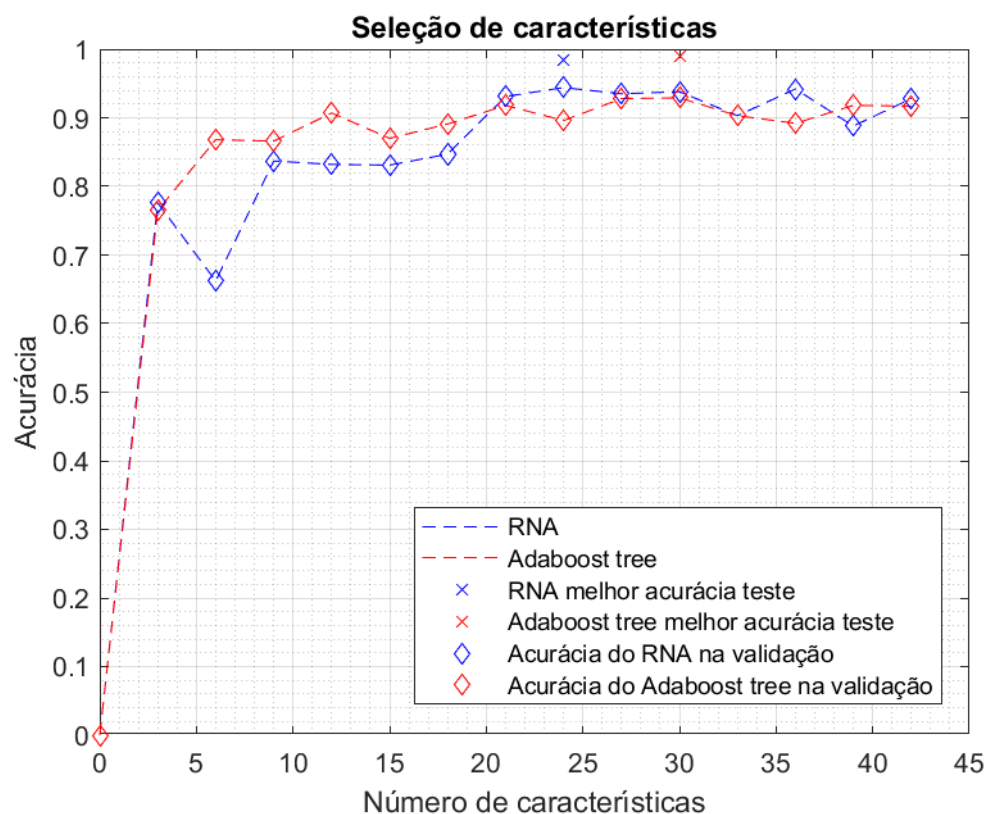
Figura 11 - Matriz confusão para o classificador *Adaboost tree* com 42 características.



Fonte: autoria própria.

Através da aplicação “*Classification Learner*” do MATLAB configurou-se o algoritmo RELIEFF, a fim de verificar qual seria a quantidade de características a ser utilizada que retornaria a melhor acurácia no conjunto de validação. O Gráfico 1 mostra os resultados obtidos com a seleção de características, no qual observa-se que o acréscimo da quantidade de características utilizadas não representa diretamente melhora no desempenho do modelo, já que a maior acurácia obtida com o classificador RNA se deu com 24 características (94,4% na validação e 98,5% no teste) e com o classificador *Adaboost tree* ocorreu com 30 características (92,9% na validação e 99,0% no teste). Em ambos os casos foi realizado um incremento de 3 características a cada iteração devido à quantidade de simulações a serem executadas.

Gráfico 1 - Resultado da seleção de características.



Fonte: autoria própria.

A Tabela 8 apresenta as características mais significantes de acordo com o algoritmo RELIEFF.

Tabela 8 - Relevâncias das características selecionadas.

Relevância	RNA			Adaboost tree		
	Característica	Direção	RELIEFF	Característica	Direção	RELIEFF
1	Média	Y	0,1428	Média	Y	0,1428
2	Média	Z	0,0475	Média	Z	0,0475
3	Média	X	0,0419	Média	X	0,0419
4	Fator de crista	Z	0,0147	Fator de crista	Z	0,0147
5	Fator de impulso	Z	0,0135	Fator de impulso	Z	0,0135
6	Curtose	Z	0,0134	Curtose	Z	0,0134
7	Assimetria	Z	0,0128	Assimetria	Z	0,0128
8	Clearance	Z	0,0114	Clearance	Z	0,0114
9	Assimetria	Y	0,0091	Assimetria	Y	0,0091
10	Clearance	X	0,0086	Clearance	X	0,0086
11	Fator de forma	Z	0,0066	Fator de forma	Z	0,0066
12	Mínimo	X	0,0061	Mínimo	X	0,0061
13	Curtose	Y	0,0052	Curtose	Y	0,0052
14	Fator de impulso	Y	0,0051	Fator de impulso	Y	0,0051
15	Fator de crista	Y	0,0048	Fator de crista	Y	0,0048
16	Fator de impulso	X	0,0047	Fator de impulso	X	0,0047
17	Pico a pico	X	0,0036	Pico a pico	X	0,0036
18	Curtose	X	0,0036	Curtose	X	0,0036
19	Fator de crista	X	0,0031	Fator de crista	X	0,0031
20	Fator de forma	Y	0,0014	Fator de forma	Y	0,0014
21	Fator de forma	X	0,0005	Fator de forma	X	0,0005
22	Pico a pico	Z	0,0002	Pico a pico	Z	0,0002
23	Mínimo	Z	0,0002	Mínimo	Z	0,0002
24	Desvio padrão	Z	$4,36 \cdot 10^{-5}$	Desvio padrão	Z	$4,36 \cdot 10^{-5}$
25	-	-	-	RMS	Z	$4,07 \cdot 10^{-5}$
26	-	-	-	Valor de pico	X	-0,0001
27	-	-	-	Desvio padrão	X	-0,0006
28	-	-	-	RMS	X	-0,0006
29	-	-	-	Assimetria	X	-0,0013
30	-	-	-	Valor de pico	Z	-0,0019

Fonte: autoria própria.

A Tabela 8 mostra a importância que características como média, curtose e assimetria possuem na identificação das falhas em máquinas rotativas e que a ausência do fator de margem entre as características selecionadas indica que para a situação analisada essa característica não é muito significativa.

Outra análise possível é a relevância das direções em que foram realizadas as medições. Apesar da característica mais relevante ter sido medida na direção Y, entre as 10 características mais relevantes 6 foram medidas na direção Z, o que demonstra que variações nessa direção tendem a ser mais significativas ao se identificar falhas.

Ao identificar quais são os melhores resultados em termos de acurácia se torna interessante analisar com mais detalhes esses casos a partir de suas matrizes confusão, dispostas nas Figura 12 e Figura 13.

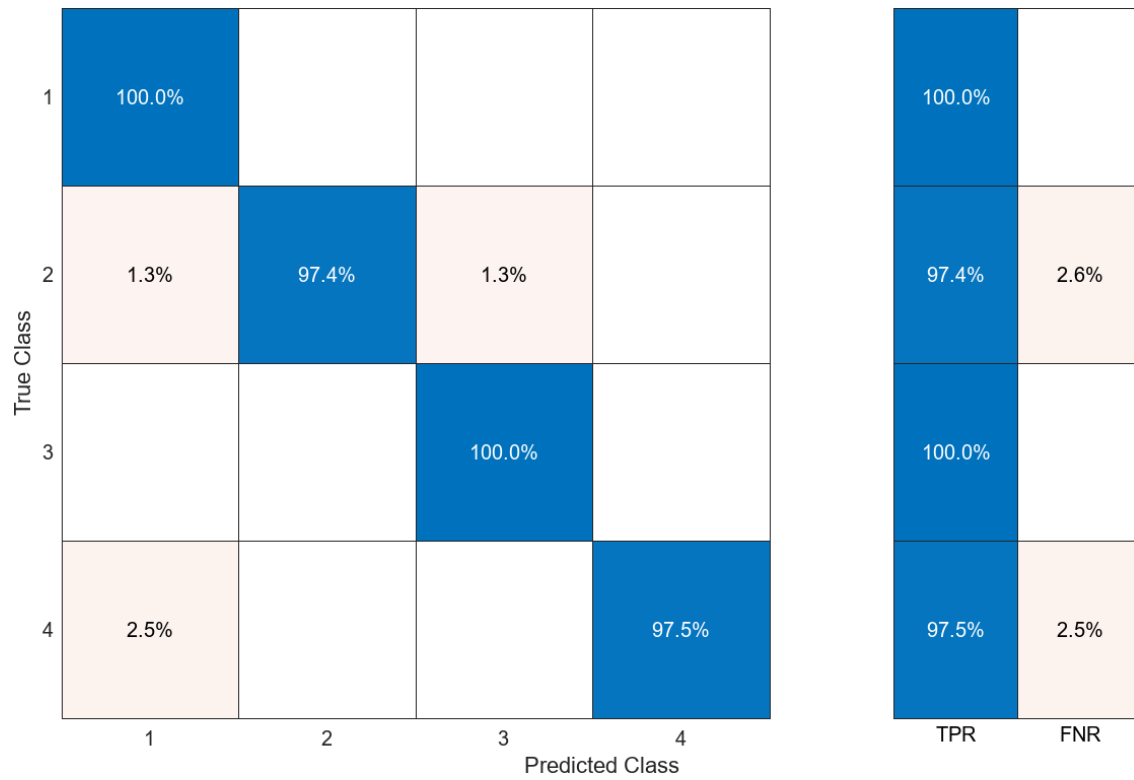
A matriz da Figura 12 (Matriz confusão do classificador RNA com 24 características) mostra que o modelo identificou corretamente todos os casos das classes 1 (condição normal) e 3 (desalinhamento horizontal) e apresentou pequena taxa de erro para as classes 2 (desbalanceamento) e 4 (desalinhamento vertical), taxas essas de 2,6% e 2,5%, respectivamente.

Para o caso do classificador *Adaboost tree* com 30 características (Figura 13), verifica-se a permanência da taxa de 100% de acerto para as classes 1 e 3, melhora na identificação da classe 2, com 98,7% das identificações corretas e, para classe 4, taxa de erro de 2,5%, mesma obtida pelo classificador RNA.

Além da acurácia, outras medidas de desempenho podem ser usadas ao se analisar uma classificação, sendo utilizadas para avaliar a capacidade de generalização do classificador treinado, para determinar o melhor classificador a ser utilizado e para distinguir e selecionar a melhor solução entre as geradas durante um treinamento de classificação (HOSSIN & SULAIMAN, 2015).

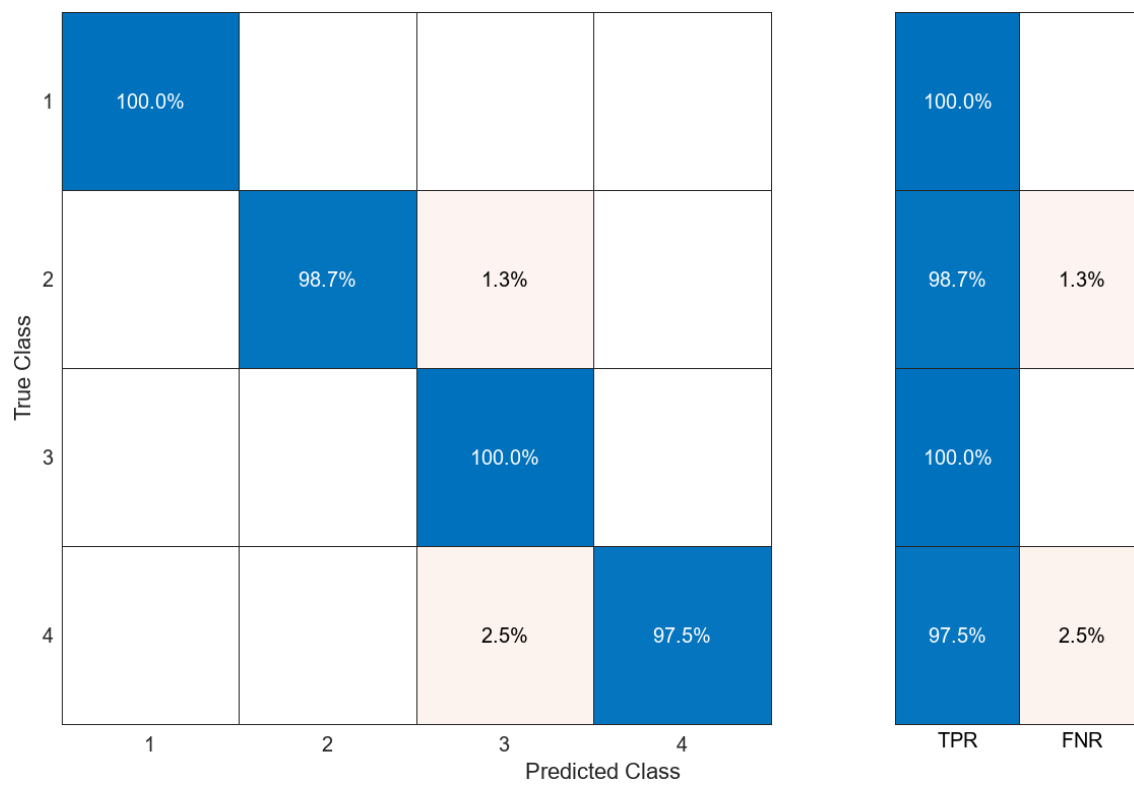
Com a finalidade de aprofundar a análise comparativa entre os classificadores RNA e *Adaboost tree* foram utilizadas sete métricas, descritas na Tabela 2.

Figura 12 - Matriz confusão do classificador RNA no teste com 24 características.



Fonte: autoria própria.

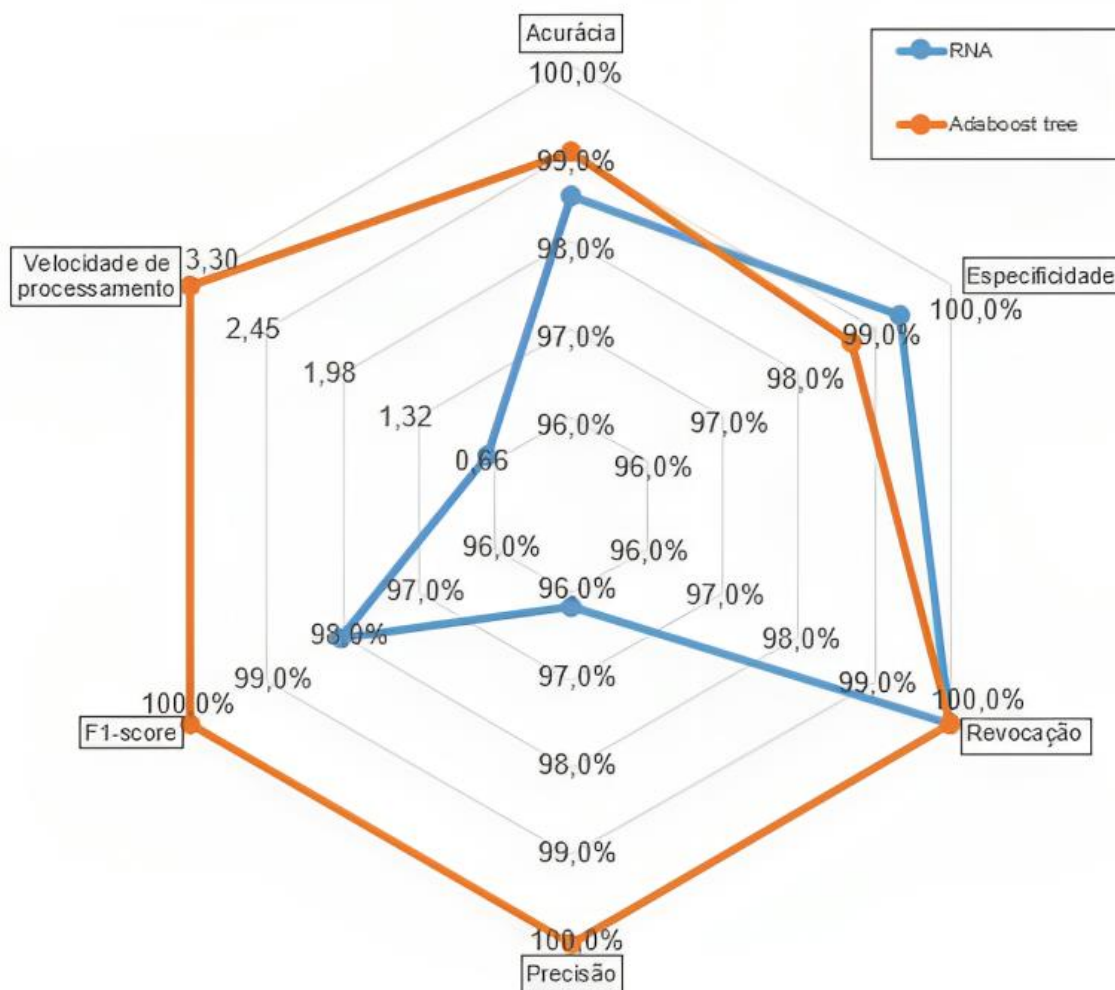
Figura 13 - Matriz confusão do classificador *Adaboost tree* no teste com 30 características.



Fonte: autoria própria.

O Gráfico 2 apresenta um gráfico comparativo entre o desempenho do classificador RNA (24 características) e o *Adaboost tree* (30 características) no conjunto de teste, considerando as métricas apresentadas na Tabela 2.

Gráfico 2 - Gráfico radar comparativo entre métricas dos classificadores RNA e *Adaboost tree*.



Fonte: autoria própria.

O gráfico comparativo mostra que em quase todas as métricas calculadas o classificador *Adaboost tree* possui melhor desempenho que o RNA, principalmente em precisão e velocidade de processamento, já em especificidade o classificador RNA obteve um desempenho melhor, porém a diferença não foi muito significativa, concluindo que, para a base de dados utilizada, o classificador *Adaboost tree* com 30 características possui melhor performance para a identificação de falhas em máquinas rotativas.

9 CONCLUSÕES E RECOMENDAÇÕES

A partir dos resultados alcançados neste trabalho verifica-se que a acurácia alcançada a partir da metodologia utilizada é satisfatória para a aplicação na identificação de falhas de desbalanceamento, desalinhamento horizontal e desalinhamento vertical de máquinas rotativas, possibilitando a utilização de procedimentos semelhantes na manutenção de equipamentos em indústrias e embarcações.

Ao ser feita a análise das vantagens de se utilizar tais técnicas constata-se que a detecção e o diagnóstico de falhas são úteis para otimizar e garantir a segurança na operação de máquinas, resultando em maior produtividade e eficiência do processo, com benefícios como redução dos custos operacionais, maior vida útil da máquina e maior tempo de funcionamento operacional (Muñiz; Díaz; Cuadrado; 2020),

Sendo assim, conclui-se que a utilização de aprendizado de máquina para a identificação de falhas em máquinas rotativas está alinhada com a missão, estratégias e objetivos da Marinha do Brasil, sendo, portanto, um tema válido para que seja realizado um estudo de viabilidade de execução deste projeto.

REFERÊNCIAS

AHMED, H.; NANDI, A. K. Compressive sampling and feature ranking framework for bearing fault classification with vibration signals. **IEEE access: practical innovations, open solutions**, v. 6, p. 44731–44746, 2018.

BERGHOUT, T. et al. A deep supervised learning approach for condition-based maintenance of naval propulsion systems. **Ocean engineering**, v. 221, n. 108525, p. 108525, 2021.

BHATTARAI U., MARASENI T., DEVKOTA L.P., APAN A. Application of machine learning to assess people's perception of household energy in the developing world: A case of Nepal. **Energy and AI**, v. 14, 9 Primavera 2023.

BRASIL. [Constituição (1988)]. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República. Disponível em: https://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm. Acesso em: 23 out. 2023.

BRASIL. **Plano Estratégico da Marinha (PEM 2040)**. Brasília, DF: 2020. Marinha do Brasil. Disponível em: https://www.marinha.mil.br/sites/all/modules/pub_pem_2040/book.html. Acesso em: 23 out. 2023.

CULLUM, J. et al. Risk-Based Maintenance Scheduling with application to naval vessels and ships. **Ocean engineering**, v. 148, p. 476–485, 2018.

DAS, O.; BAGCI DAS, D.; BIRANT, D. Machine learning for fault analysis in rotating machinery: A comprehensive review. **Heliyon**, v. 9, n. 6, p. e17584, 2023.

DE SÁ SÓ MARTINS, D. H. C. **Diagnóstico de Falhas Combinadas em Máquinas Rotativas**. [s.l.] CEFET/RJ e UFF, 2022.

DIKIS, K. K.; LAZAKIS I., **Dynamic risk and reliability assessment of ship machinery and equipment**, Proc. 26th Int. Ocean Polar Eng. Conf., Rhodes, Greece, vol. 2016, pp. 969–976.

GALAR PASCUAL, D. **Artificial intelligence tools: Decision support systems in condition monitoring and Diagnosis**. Londres, England: CRC Press, 2020.

GHOLAMZADEH, M.; ABTAHI, H.; SAFDARI, R. Comparison of different machine learning algorithms to classify patients suspected of having sepsis

infection in the intensive care unit. **Informatics in medicine unlocked**, v. 38, n. 101236, p. 101236, 2023.

GIUDICI, P.; GRAMEGNA, A.; RAFFINETTI, E. Machine learning classification model comparison. **Socio-economic planning sciences**, v. 87, n. 101560, p. 101560, 2023.

GONZÁLEZ-MUÑIZ, A.; DÍAZ, I.; CUADRADO, A. A. DCNN for condition monitoring and fault detection in rotating machines and its contribution to the understanding of machine nature. **Heliyon**, v. 6, n. 2, p. e03395, 2020.

HOSSAIN, M.; ABU-SIADA, A.; MUYEEN, S. Methods for advanced wind turbine condition monitoring and early diagnosis: A literature review. **Energies**, v. 11, n. 5, p. 1309, 2018.

KUCHERYAVSKIY, S.; RODIONOVA, O.; POMERANTSEV, A. Procrustes cross-validation of multivariate regression models. **Analytica chimica acta**, v. 1255, n. 341096, p. 341096, 2023.

LI, J.-B. et al. Feedback on a shared big dataset for intelligent TBM Part I: Feature extraction and machine learning methods. **Underground space**, v. 11, p. 1–25, 2023.

MAHESH et al. The stratified K-folds cross-validation and class-balancing methods with high-performance ensemble classifiers for breast cancer classification. **Healthcare Analytics**, v. 4, n. 100247, p. 100247, 2023.

MANNING, C. D.; RAGHAVAN, P.; SCHUTZE, H. **Introduction to Information Retrieval**. Cambridge, England: Cambridge University Press, 2012.

MARSH, B. **Multivariate Analysis of the Vector Boson Fusion Higgs Boson**. University of Missouri, 2016.

M, H.; M.N, S. A review on evaluation metrics for data classification evaluations. **International Journal of Data Mining & Knowledge Management Process**, v. 5, n. 2, p. 01–11, 2015.

PECO CHACÓN, A. M.; SEGOVIA RAMÍREZ, I.; GARCÍA MÁRQUEZ, F. P. K-nearest neighbour and K-fold cross-validation used in wind turbines for false alarm detection. **Sustainable Futures**, v. 6, n. 100132, p. 100132, 2023.

RANDALL, R. B. **Vibration-based condition monitoring: Industrial, aerospace and automotive applications**. Hoboken, NJ, USA: Wiley-Blackwell, 2011.

ROLIM V. B., CORDEIRO F. R., FERREIRA R. Reconhecimento de Padrões Aplicados a Comentários de Fóruns Educacionais. 2014.

ROQUE, A. S. et al. An analysis of machine learning algorithms in rotating machines maintenance. **IFAC-PapersOnLine**, v. 55, n. 2, p. 252–257, 2022.

SARKER, I. H. Machine learning: Algorithms, real-world applications and research directions. **SN computer science**, v. 2, n. 3, 2021.

SHAHRIARI, B. et al. Taking the human out of the loop: A review of Bayesian optimization. **Proceedings of the IEEE. Institute of Electrical and Electronics Engineers**, v. 104, n. 1, p. 148–175, 2016.

SUN, Y. Iterative RELIEF for feature weighting: Algorithms, theories, and applications. **IEEE transactions on pattern analysis and machine intelligence**, v. 29, n. 6, p. 1035–1051, 2007.

VERBERT, K.; DE SCHUTTER, B.; BABUŠKA, R. Timely condition-based maintenance planning for multi-component systems. **Reliability Engineering & System Safety**, v. 159, p. 310–321, 2017.