



UNIVERSIDADE FEDERAL DO PARÁ
NÚCLEO DE DESENVOLVIMENTO AMAZÔNICO EM ENGENHARIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA DE INFRAESTRUTURA E
DESENVOLVIMENTO ENERGÉTICO

JOÃO LUCAS LOBATO SOARES

DESENVOLVIMENTO DE SISTEMA DE DIAGNÓSTICO DE FALHAS EM ROLETES DE
TRANSPORTADORES DE CORREIA

Tucuruí/Pa
2024

JOÃO LUCAS LOBATO SOARES

**DESENVOLVIMENTO DE SISTEMA DE DIAGNÓSTICO DE FALHAS EM ROLETES
DE TRANSPORTADORES DE CORREIA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Infraestrutura e Desenvolvimento Energético do Núcleo de Desenvolvimento Amazônico em Engenharia, da Universidade Federal do Pará, como requisito para a obtenção do título de Mestre em Infraestrutura e Desenvolvimento Energético. Área de concentração: Tecnologias Sustentáveis.

Prof. Dr. Alexandre Luiz Amarante Mesquita

Tucuruí/Pa

2024

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a)
autor(a)

S676d Soares, João Lucas Lobato.
Desenvolvimento de sistema de diagnóstico de falhas em
roletes de transportadores de correia / João Lucas Lobato
Soares. — 2024.
126 f. : il. color.

Orientador(a): Prof. Dr. Alexandre Luiz Amarante
Mesquita
Dissertação (Mestrado) - Universidade Federal do Pará,
Núcleo de Desenvolvimento Amazônico em Engenharia,
Programa de Pós-Graduação em Engenharia de
Infraestrutura e Desenvolvimento Energético, Tucuruí, 2024.

1. Roletes. 2. Diagnóstico de falha. 3. Wavelet
Packet Decomposition. 4. Seleção de features. 5.
Aprendizado de máquina. I. Título.

CDD 620.1054

JOÃO LUCAS LOBATO SOARES

**DESENVOLVIMENTO DE SISTEMA DE DIAGNÓSTICO DE FALHAS EM ROLETES
DE TRANSPORTADORES DE CORREIA**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia de Infraestrutura e Desenvolvimento Energético do Núcleo de Desenvolvimento Amazônico em Engenharia, da Universidade Federal do Pará, como requisito para a obtenção do título de Mestre em Infraestrutura e Desenvolvimento Energético. Área de concentração: Tecnologias Sustentáveis.

Prof. Dr. Alexandre Luiz Amarante Mesquita

Aprovada em 28 de Março de 2024.

BANCA EXAMINADORA:

Prof. Dr. Alexandre Luiz Amarante Mesquita / UFPA - Orientador

Prof. Dr. André Luiz Amarante Mesquita / UFPA

Prof. Dr. Júlio Apolinário Cordioli / UFSC

Dedico este trabalho a meus pais, fonte de inspiração na minha
jornada em busca de conhecimento

AGRADECIMENTOS

Agradeço à Deus, criador e regente do universo. Aos meus pais, João Gualberto e Prof.^a Joana Cláudia por todo apoio, conselhos e incentivo ao meu desenvolvimento pessoal e profissional. São a minha principal fonte de inspiração. À minha irmã Amanda de Cássia pela parceria desde sempre e, principalmente, nos últimos anos com conselhos e contribuições para o meu aperfeiçoamento acadêmico.

Aos meus avós Manoel Lobato e Maria de Fátima e à minha tia Natalina de Jesus por todo amor e acolhimento recebido antes e após minha graduação. À minha namorada Yanka Rodrigues pela compreensão e cumplicidade a todo momento, pessoalmente ou à distância.

Aos Professores Alexandre Mesquita, André Mesquita e Walter Sousa, pela oportunidade concedida para participar do projeto e do mestrado. Nunca esquecerei a ajuda, os conselhos e as críticas construtivas que recebi durante esse período. À equipe Dynamox, em especial, ao Danilo Braga e Jullyane Figueiredo, fundamentais no apoio técnico para o desenvolvimento do projeto.

E por último, a todos os participantes do projeto, como o Geovane Nascimento, Lis Moura, Elton Souza, Thiago Barroso e demais integrantes, pela troca de experiências e auxílio na execução das atividades.

"Em algum lugar, algo incrível está esperando para ser conhecido"

Carl Sagan

RESUMO

Os transportadores de correia são equipamentos essenciais na indústria de mineração e requerem monitoramento constante para manter uma boa confiabilidade. Com o objetivo de suportar a correia e o material transportado, os roletes são componentes que, constantemente, falham durante a operação, nos quais apresentam defeitos nos rolamentos e desgaste superficial na concha como modos de falha mais comuns. Assim, o monitoramento baseado em manutenção preditiva é essencial, e técnicas de aprendizado de máquina podem ser utilizadas como alternativa para detecção de falhas em equipamentos. Em diagnósticos utilizando aprendizado de máquina, a etapa de seleção de features é importante para evitar a perda de precisão na classificação da condição do equipamento. O presente estudo analisa o desempenho do algoritmo de árvore de decisão e Análise de Variância (ANOVA) como métodos alternativos para redução de dimensionalidade. Inicialmente, os sinais de vibração foram coletados nos roletes de uma bancada de transportador de correia e a *Wavelet Packet Decomposition* (WPD) foi aplicada aos sinais para obtenção das faixas de energia, que foram utilizadas como features para classificação. Após a determinação das melhores features, duas abordagens foram analisadas para seleção de características: uma com a aplicação do método sem redução de dimensionalidade e outra com a aplicação da árvore de decisão. Adicionalmente, foram empregados diferentes algoritmos de classificação: Máquina de Vetores de Suporte (SVM), k-ésimo Vizinho mais Próximo (kNN) e Rede Neural Artificial (ANN). Como resultados, constatou-se um desempenho superior de acurácia diagnóstica em todas as técnicas com redução de dimensionalidade das características selecionadas pela árvore de decisão. Além disso, SVM, kNN e ANN apresentaram aumentos de acuracidade dentre os modelos de diagnóstico de falha abordados.

Palavras-chave: Roletes. Diagnóstico de falha. *Wavelet Packet Decomposition*. Seleção de *features*. Aprendizado de máquina.

ABSTRACT

Belt conveyors are essential equipment in mining industry and require constant monitoring to maintain good reliability. In order to support the belt and the material being conveyed, rollers are components that constantly fail during operation, in which they present faults in bearings and surface wear in the shell as the most common failure modes. Thus, monitoring based on predictive maintenance is essential, and machine learning techniques can be used as an alternative for detecting equipment failures. In diagnostics using machine learning, the feature selection step is important to avoid loss of accuracy in the classification of the equipment's condition. The present study analyzes the performance of the decision tree algorithm and Analysis of Variance (ANOVA) as alternative methods for dimensionality reduction. Initially, the vibration signals were collected on the rollers of a belt conveyor bench and the Wavelet Packet Decomposition (WPD) was applied to the signals to obtain the energy ranges, which were used as features for classification. After the determination of the best features, two approaches were analyzed for the selection of features: one with the application of the method without dimensionality reduction and the other with the application of the decision tree. In addition, different classification algorithms were used: Support Vector Machine (SVM), k-Nearest Neighbor (kNN) and Artificial Neural Network (ANN). As a result, it was found a superior performance of diagnostic accuracy in all techniques with a reduction in the dimensionality of the characteristics selected by the decision tree. In addition, SVM, kNN and ANN showed increases in accuracy ranging among the fault diagnosis models approached.

Keywords: Rollers. Fault diagnosis. Wavelet Packet Decomposition. Feature selection. Machine learning.

LISTA DE ILUSTRAÇÕES

Figura 2. 1 – Roletes = rolos + cavalete.....	17
Figura 2. 2 – Estrutura do rolo	18
Figura 2. 3 – Fase de detecção por parâmetro físico	20
Figura 3. 1 – Tipos de dados não estacionários: (a) estacionários; (b) RMS variável no tempo; e (c) conteúdo de frequência variável no tempo	29
Figura 3. 2 – Resolução do sinal para diferentes técnicas de análise de sinais	31
Figura 3. 3 – Base wavelet.....	32
Figura 3. 4 – Aproximações ARM. a) Dois níveis. b) Multiníveis.....	35
Figura 3. 5 – Família Daubechies	37
Figura 3. 6 – Estrutura do a) DWT e b) WPD.....	38
Figura 3. 7 – Modelo de matriz de confusão.....	40
Figura 3. 8 – Hiperplano de separação ótima	41
Figura 3. 9 – Classificação para três vizinhos mais próximos.....	46
Figura 3. 10 – Evolução do aprendizado de máquinas	47
Figura 3. 11 – Estrutura do neurônio biológico	48
Figura 3. 12 – Diagrama de rede para uma camada oculta	50
Figura 3. 13 – Curva da função ReLU e sua derivada	51
Figura 3. 14 – Árvore de decisão	54
Figura 4. 1 – Etapas do trabalho a) Criação de falhas, b) Coleta e processamento de dados, c) Extração de features e seleção de features e d) Classificação.....	57
Figura 4. 2 – Desenho técnico da bancada de transportador de correia.....	58
Figura 4. 3 – Desenho técnico do rolo.....	59
Figura 4. 4 – Desenho técnico do cavalete	59
Figura 4. 5 – Torno mecânico.....	60
Figura 4. 6 – Rolos com desgaste artificial.....	61
Figura 4. 7 – Rolo desmontado	62
Figura 4. 8 – Comparativo entre rolamento saudável e defeituoso.....	62
Figura 4. 9 – Bancada do transportador de correia <i>as built</i>	63
Figura 4. 10 – Desenho técnico dos <i>data loggers</i> HF e HF+	65
Figura 4. 11 – Direção dos eixos de vibração	66

Figura 4. 12 – Configuração do rolo defeituoso instalado nas posições: Rolo com rolamento defeituoso no LA (a) e LOA (b) e rolo com desgaste superficial no LA (c) e LOA (d)	68
Figura 4. 13 – Modelo de árvore de decisão para seleção de <i>features</i>	71
Figura 5. 1 – <i>Boxplots</i> de classes na energia de banda ‘E1’ no eixo vertical	77
Figura 5. 2 – <i>Boxplots</i> de classes na energia de banda ‘E13’ no eixo vertical	77
Figura 5. 3 – <i>Boxplots</i> de classes na energia de banda ‘E11’ no eixo vertical	78
Figura 5. 4 – <i>Boxplots</i> de classes na energia de banda ‘E1’ no eixo horizontal	79
Figura 5. 5 – <i>Boxplots</i> de classes na energia de banda ‘E13’ no eixo horizontal	80
Figura 5. 6 – <i>Boxplots</i> de classes na energia de banda ‘E14’ no eixo horizontal	80
Figura 5. 7 – <i>Boxplots</i> de classes na energia de banda ‘E16’ no eixo axial	81
Figura 5. 8 – <i>Boxplots</i> de classes na energia de banda ‘E13’ no eixo axial	82
Figura 5. 9 – <i>Boxplots</i> de classes na energia de banda ‘E14’ no eixo axial	82
Figura 5. 10 – Árvore de decisão no eixo vertical.....	83
Figura 5. 11 – Árvore de decisão no eixo horizontal.....	84
Figura 5. 12 – Árvore de decisão no eixo axial.....	84
Figura 5. 13 – Acuracidade SVM (sem seleção de <i>features</i>)	86
Figura 5. 14 – Matriz de Confusão SVM (Sinal vertical).....	87
Figura 5. 15 – Matriz de Confusão SVM (Sinal horizontal).....	88
Figura 5. 16 – Matriz de Confusão SVM (Sinal axial).....	89
Figura 5. 17 – Acuracidade SVM (com seleção de <i>features</i>)	91
Figura 5. 18 – Matriz de Confusão SVM com seleção de <i>features</i> (Sinal vertical)....	92
Figura 5. 19 – Matriz de Confusão SVM com seleção de <i>features</i> (Sinal horizontal)93	
Figura 5. 20 – Matriz de Confusão SVM com seleção de <i>features</i> (Sinal axial)	94
Figura 5. 21 – Acuracidade kNN (sem seleção de <i>features</i>)	95
Figura 5. 22 – Matriz de Confusão kNN (Sinal vertical).....	96
Figura 5. 23 – Matriz de Confusão kNN (Sinal horizontal).....	97
Figura 5. 24 – Matriz de Confusão kNN (Sinal axial).....	98
Figura 5. 25 – Acuracidade kNN (com seleção de <i>features</i>)	100
Figura 5. 26 – Matriz de Confusão kNN com seleção de <i>features</i> (Sinal vertical)...	101
Figura 5. 27 – Matriz de Confusão kNN com seleção de <i>features</i> (Sinal horizontal)	102
Figura 5. 28 – Matriz de Confusão kNN com seleção de <i>features</i> (Sinal axial)	103

Figura 5. 29 – Acuracidade ANN-MLP (sem seleção de <i>features</i>)	105
Figura 5. 30 – Matriz de Confusão ANN-MLP (Sinal vertical)	106
Figura 5. 31 – Matriz de Confusão ANN-MLP (Sinal horizontal).....	107
Figura 5. 32 – Matriz de Confusão ANN-MLP (Sinal axial).....	108
Figura 5. 33 – Acuracidade ANN-MLP (com seleção de <i>features</i>).....	109
Figura 5. 34 – Matriz de Confusão ANN-MLP com seleção de <i>features</i> (Sinal vertical)	110
Figura 5. 35 – Matriz de Confusão ANN-MLP com seleção de <i>features</i> (Sinal horizontal).....	111
Figura 5. 36 – Matriz de Confusão ANN-MLP com seleção de <i>features</i> (Sinal axial)	112

LISTA DE TABELAS

Tabela 2. 1 – Levantamento de artigos publicados na área de diagnóstico de falhas em roletes assistidos por aprendizado de máquina.....	19
Tabela 4. 1 – Especificação técnica do sistema de acionamento e transmissão	58
Tabela 4. 2 – Especificações técnicas modelo HF.....	64
Tabela 4. 3 – Especificações técnicas modelo HF+	64
Tabela 4. 4 – Configuração padrão da espectral.....	66
Tabela 4. 5 – Perfil de coleta de sinais.....	67
Tabela 4. 6 – Rótulo da condição dos rolos	69
Tabela 4. 7 – Parâmetros para a árvore de decisão	70
Tabela 4. 8 – Configuração de parâmetros dos algoritmos de aprendizado	72
Tabela 5. 1 – Análise de Variância (Vertical)	74
Tabela 5. 2 – Análise de Variância (Horizontal).....	75
Tabela 5. 3 – Análise de Variância (Axial).....	76
Tabela 5. 4 – Desempenho geral do algoritmo SVM (sem seleção de <i>features</i>)	90
Tabela 5. 5 – Desempenho geral do algoritmo SVM (com seleção de <i>features</i>)	94
Tabela 5. 6 – Desempenho geral do algoritmo kNN (sem seleção de <i>features</i>).....	99
Tabela 5. 7 – Desempenho geral do algoritmo kNN (com seleção de <i>features</i>).....	103
Tabela 5. 8 – Desempenho geral do algoritmo ANN-MLP (sem seleção de <i>features</i>)	108
Tabela 5. 9 – Desempenho geral do algoritmo ANN-MLP (com seleção de <i>features</i>)	112

SUMÁRIO

1. INTRODUÇÃO	13
1.1. Contextualização.....	13
1.2. Objetivos.....	15
1.2.1. Objetivo Geral	15
1.2.2. Objetivos Específicos	15
1.3. Estrutura do trabalho	15
2. REVISÃO BIBLIOGRÁFICA	17
2.1. Diagnóstico de falha baseado em aprendizado de máquina para transportador de correia	18
2.2. Modos de falha	23
3. FUNDAMENTAÇÃO TEÓRICA	27
3.1. Processamento de sinais de vibração aplicado à manutenção industrial	28
3.1.1. Técnicas de tempo-frequência	29
3.1.2. Transformada Wavelet	31
3.1.3. Wavelet Packet Decomposition	37
3.2. Aprendizado de máquina.....	39
3.2.1. Máquina de vetores de suporte	40
3.2.2. K-ésimo vizinho mais próximo	45
3.2.3. Rede Neurais Artificiais	46
3.3. Redução de dimensionalidade.....	52
3.3.1. Árvore de decisão	53
3.3.2. Análise de Variância (ANOVA)	55
4. METODOLOGIA	57
4.1. Criação de falhas	59
4.1.1. Desgaste superficial	59
4.1.2. Defeito em rolamento	61
4.2. Coleta e processamento de dados	62
4.3. Extração e seleção de features	68
4.3.1. Bandas de energia wavelet normalizada	69
4.3.2. ANOVA	70
4.3.3. Árvore de decisão	70
4.4. Classificação.....	71

5. RESULTADOS	73
5.1. Seleção de features.....	73
5.1.1. ANOVA	73
5.1.2. Árvore de decisão	83
5.2. Máquina de Vetores de Suporte.....	85
5.2.1. Sem seleção de <i>features</i>	85
5.2.2. Com seleção de features	90
5.3. K-ésimo Vizinho mais Próximo.....	95
5.3.1. Sem seleção de <i>features</i>	95
5.3.2. Com seleção de features	99
5.4. Perceptron Multicamadas.....	104
5.4.1. Sem seleção de <i>features</i>	104
5.4.2. Com seleção de features	109
6. CONSIDERAÇÕES FINAIS	113
REFERÊNCIAS BIBLIOGRÁFICAS	115

1. INTRODUÇÃO

1.1. Contextualização

Considerada um dos principais equipamentos de transporte de granéis, o transportador de correia desempenha papel fundamental na mineração. Além disso, os avanços tecnológicos ao longo do tempo levaram a uma melhoria significativa no desempenho operacional, tanto no transporte de longa distância como em altas capacidades de carga. No entanto, o estado operacional resultou em problemas recorrentes de manutenção, como danos significativos ao equipamento. Portanto, foi necessária a análise de falhas dos componentes do transportador, que ocorrem, majoritariamente, em regimes não estacionários devido às variações de carga do material durante a operação do transportador (LODEWIJKS, 2002).

Dentre esses componentes, o rolete é um dos principais alvos de estudo, pois envolve a avaliação das condições de operação, estimativas de confiabilidade e monitoramento de dados para melhorar a precisão da tomada de decisões (Liu *et al*, 2019). Devido à dificuldade de acesso seguro às atividades de inspeção e manutenção, os rolos geralmente não são priorizados durante o tempo de inatividade programado, o que resulta na redução da funcionalidade do componente e no potencial de danos irreparáveis aos rolos inoperantes enquanto outros sistemas continuam inicialmente a funcionar (SWINDERMAN *et al*, 2016).

A manutenção preditiva é uma ferramenta que garante a integridade do sistema transportador por meio do monitoramento de parâmetros físicos como vibração e temperatura (LIU *et al*, 2019). A análise de vibração é uma das principais ferramentas de detecção de falhas que utiliza técnicas para monitorar sinais de máquinas rotativas. Os padrões de sinal são estabelecidos quando o equipamento está em condições normais e, posteriormente, são observadas alterações no domínio do tempo, frequência ou domínio tempo-frequência para diagnosticar possíveis falhas de componentes (POPESCU *et al*, 2021).

Devido ao comportamento não estacionário e não linear do sinal, técnicas tradicionais como a transformada de Fourier apresentam baixa eficiência na análise das características nos domínios do tempo e da frequência individualmente. Assim, métodos alternativos, como *Wavelet Packet Decomposition* (WPD), que aborda sinais

de vibração no domínio tempo-frequência a partir de dados não estacionários, podem analisar efetivamente sinais com essas características (LI *et al*, 2013).

Além disso, a manutenção preditiva tornou-se um dos principais paradigmas da Indústria 4.0, permitindo maior comunicação entre sistemas de monitoramento remoto com capacidades aprimoradas de processamento e memória (MUNIZ *et al*, 2023). Para soluções industriais mais competitivas, a manutenção preditiva é combinada com aprendizado de máquina para garantir que o equipamento não sofra falhas, ao mesmo tempo que reduz os custos de reparo (DA SILVA *et al*, 2017; LOBATO *et al*, 2020; EFEOĞLU *et al*, 2022). Assim, muitos estudos estão sendo realizados para diagnosticar falhas em rolos de correias transportadoras utilizando diferentes algoritmos de aprendizado de máquina supervisionados e não supervisionados.

O aprendizado de máquina também surge como uma alternativa para diagnosticar falhas através da análise de vibrações. Dessa forma, são criados modelos capazes de classificar características do sinal e determinar a integridade dos componentes com alta precisão (ALHARBI *et al*, 2023). Entre as principais técnicas de aprendizado de máquina já aplicadas na literatura para usuários, exemplos notáveis incluem Máquina de Vetores de Suporte (SVM) (LI *et al*, 2013), k-ésimo Vizinhos mais Próximos (kNN) (YANG *et al*, 2020) e Rede Neural Artificial (ANN) (RAVIKUMAR *et al*, 2020).

Para aplicação do modelo, é necessário selecionar as características mais significativas para reduzir a redundância de informações e melhorar a eficácia da classificação. A seleção de *features* auxilia na redução da dimensionalidade do banco de dados por meio de diversas técnicas que aumentam o poder discriminatório do modelo, preservando um subconjunto de características com maior relevância (RAUBER *et al*, 2014).

Ravikumar *et al* (2020) aplicaram árvores de decisão para identificar características estatísticas com maior capacidade de diagnosticar falhas em roletes de correias transportadoras. As árvores de decisão destacam regras que dividem o conjunto de amostras em classes com base em certas condições de um recurso. Enquanto, Peña *et al* (2018) usaram a Análise de Variância (ANOVA) para ranquear e selecionar as *features* mais significativas para o diagnóstico de falhas em máquinas rotativas.

No trabalho presente, os sinais de vibração dos roletes da correia transportadora foram decompostos em bandas de energia *wavelet*, a partir das quais as bandas de energia mais relevantes foram identificadas como seleção de *features* usando árvores de decisão e ANOVA. Posteriormente, foi feita uma comparação entre modelos de aprendizado de máquina com e sem seleção de recursos. As técnicas selecionadas para análise foram SVM, kNN e RNA, para identificar qual técnica apresenta maior performance de acuracidade com e sem a seleção de *features*.

1.2. Objetivos

1.2.1. Objetivo Geral

Desenvolver um sistema de diagnóstico de falhas em roletes de transportador de correia com diferentes modos de falha, graus de severidade, bem como o posicionamento do rolo defeituoso, a partir de diferentes técnicas de seleção de *features* e algoritmos de classificação.

1.2.2. Objetivos Específicos

- Avaliar a eficácia do diagnóstico de falha em roletes a partir do tratamento de sinais de cada eixo coletados em sensores de vibração triaxiais;
- Avaliar ANOVA e árvore de decisão como técnicas para seleção de *features*, bem como compará-las na análise para separabilidade entre classes;
- Realizar diagnóstico de falhas em roletes, com e sem seleção de *features*, por meio de diferentes técnicas de aprendizado de máquina;
- Identificar o posicionamento dos rolos defeituosos por meio de classificação do sinal de vibração medido no cavalete.

1.3. Estrutura do trabalho

A dissertação foi dividida em seis capítulos. O capítulo 2 apresenta a revisão bibliográfica, na qual apresenta o estado da arte da área relacionada a aplicação de diagnóstico de falha em roletes, bem como dos principais modos de falha desses componentes.

O capítulo 3 aborda o processamento de sinais aplicadas à manutenção industrial, direcionando a técnicas tempo-frequência para análise preditiva de vibração. Além disso, apresenta os conceitos das principais técnicas de aprendizado de máquina tradicionais e redes neurais para classificação. Também elucida a redução de dimensionalidade e a aplicação de seleção de *features* para algoritmos de aprendizado.

O capítulo 4 apresenta a metodologia do trabalho como a preparação do aparato experimental, criação de falhas, coleta e processamento de dados, extração e seleção de *features* e o configuração dos métodos de aprendizado de máquinas para o diagnóstico por classificação de falhas dos roletes.

O capítulo 5 exhibe os resultados acerca da pré-análise de seleção de *features* com ANOVA e árvore de decisão, bem como do desempenho dos algoritmos aplicados e os principais modelos de diagnóstico obtidos.

Por último, o capítulo 6 apresenta as considerações finais do trabalho desenvolvido e sugere oportunidades de otimização.

2. REVISÃO BIBLIOGRÁFICA

Os transportadores de correia têm a capacidade de deslocar quantidades significativas de material por longas distâncias por meio de uma correia de borracha tensionado por um sistema de tambores (SWINDERMAN *et al*, 2016).

Os principais subconjuntos do transportador podem ser divididos em:

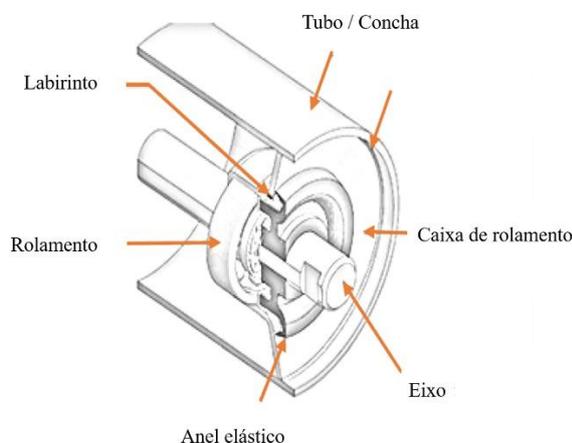
- Correia: Superfície de carregamento do material;
- Polias: Componentes de auxílio para movimentação e tensionamento da correia;
- Acionamento: Sistema de transmissão de potências para as polias;
- Estrutura: Elementos de suporte para fixação e alinhamento dos componentes rolantes;
- Sistema de transferência: Chute de Carga e Descarga do material do transportador;
- Sistema de suporte da correia: Conjunto de roletes;

Geralmente, os roletes correspondem ao conjunto de rolos instalados sobre um cavalete, conforme ilustrado na Figura 2.1, com a finalidade de auxiliar no suporte de carga de material sobre a correia (MORALES *et al*, 2009). No que diz respeito à falha em transportadores de correia, os rolos possuem componentes como o rolamento, eixo, concha e um sistema de vedação (como ilustrado na Figura 2.2) e, conseqüentemente, apresentam defeitos como desgastes e travamentos (ZIMROZ e KRÓL, 2009).

Figura 2. 1 – Roletes = rolos + cavalete



Fonte: Morales *et al* (2009)

Figura 2. 2 – Estrutura do rolo

Fonte: Adaptado de Alharbi *et al* (2023)

Sendo assim, ao longo do tempo surgiu diversos estudos científicos que exploram métodos de diagnóstico de falha em roletes, além de mapear os principais modos de falha.

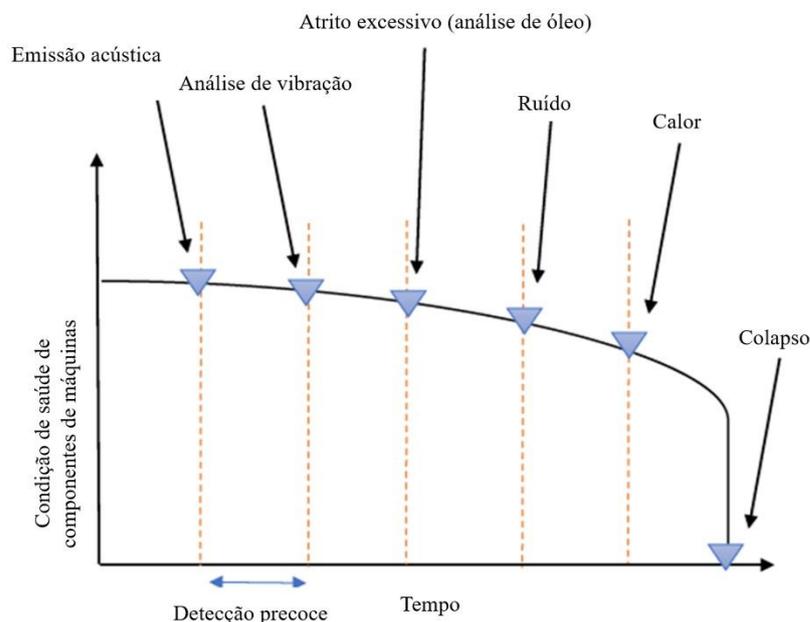
2.1. Diagnóstico de falha baseado em aprendizado de máquina para transportador de correia

As principais técnicas tradicionais para o monitoramento de falhas em transportadores de correia são realizadas a partir de inspeções manuais dos mantenedores periodicamente. Devido à dificuldade de monitoramento dos roletes por conta da distribuição ao longo do transportador de correia, o que pode comprometer a confiabilidade do equipamento devido a ruídos e perturbações externas responsáveis por falhas nas operações do subsistema de suporte de carga do transporte, a comunidade científica iniciou pesquisa para a implementação do aprendizado de máquina para detectar defeitos em roletes em tempo hábil para evitar paradas não planejadas, à exemplo dos trabalhos publicados, presentes na Tabela 2.1. A medição de parâmetros físicos como sinais acústicos e vibração são, comumente, aplicados como recurso para o monitoramento do estado dos roletes devido a detecção de falhas ainda em estágio inicial, como ilustrado na Figura 2.3, comparando-os com outros parâmetros físicos (ALHARBI *et al*, 2023).

Tabela 2. 1 – Levantamento de artigos publicados na área de diagnóstico de falhas em roletes assistidos por aprendizado de máquina

Autores	Ano	Extração de dados de entrada / features	Aprendizado de máquina
Li et al	2013	<i>Wavelet Packet Decomposition</i>	<i>Support Vector Machine (SVM)</i>
Muralidharan et al	2014	Análise estatística dos sinais	Algoritmo J48
Ravikumar et al	2019	Análise estatística dos sinais	Algoritmo k-star e Perceptron Multicamadas
Bortnowski et al	2022	Transformada de Fourier de Curta-Duração (STFT); Espectograma; Autocorrelação; Pico de frequência; Média do pico de frequência	LSTM Autoencoder
Roos e Heyns	2021	<i>Wavelet Packet Decomposition</i>	<i>Support Vector Machine (SVM)</i> e Rede Neural Artificial (RNA)
Wijaya et al	2022	Transformada Rápida de Fourier; Análise de Envelope	Isolation forest
Wijaya et al	2021	Transformada Discreta de Fourier; Domínio do tempo	Rede Neural Artificial (RNA)
Liu et al	2020	Coeficientes Mel-cepstrais	Gradient Boosting Decision Tree
Rocha et al	2021	Transformada de Fourier; Média das magnitudes	Florestas aleatórias
Yang et al	2020	Domínio do tempo; Stacked sparse auto encode: Coeficientes Mel-cepstrais	Rede Neural Deconvolucional (DNN); Rede Neural Convolucional Profunda (DCNN); <i>Support Vector Machine (SVM)</i> ; k-ésimo Vizinho mais próximo (kNN)
Peng et al	2020	<i>Wavelet Packet Decomposition</i>	Rede Neural Convolucional (CNN)

Figura 2. 3 – Fase de detecção por parâmetro físico



Fonte: Alharbi *et al* (2023)

Li *et al* (2013) publicaram um trabalho de aplicação de aprendizado de máquinas para desenvolver um programa de diagnóstico de falhas em roletes a partir da medição de vibração por acelerômetros. A extração de *features* foi realizada a partir da Wavelet Packet Decomposition (WPD) e para o modelo de aprendizado foi utilizado a Máquina de Vetores de Suporte (SVM). Além disso, houve a aplicação de cargas diferentes para roletes em condições de saúde diferentes e foi criado um sistema de monitoramento *online* para diagnosticar e localizar a posição de defeito dos três roletes mais próximos do sensor, reduzindo número de sensores necessários no transportador. Apesar disso, por se tratar de um sistema *online*, o tempo de compilação da classificação gera atraso no diagnóstico de falhas.

Muralidharan *et al* (2014) também apresentaram um sistema de diagnóstico em roletes por meio da medição dos sinais de vibração. Esses sinais foram tratados mediante análise estatística para extração de *features* e foi implementado a árvore de decisão a partir do algoritmo J48 para seleção de *features* e classificação. A seleção de *features* a partir da árvore de decisão simplifica a interpretação e entendimento das principais características para a classificação. Porém, houve baixa generalização de modelos, não podendo ser aplicáveis para todos os casos.

Ravikumar *et al* (2019) vieram com uma abordagem análoga em relação a árvore de decisão para seleção de *features*, porém, aplicando o algoritmo k-star e Perceptron Multicamadas para a classificação. Ao todo, foram selecionados menos de 50% das características do sinal previamente escolhidas. Algoritmo k-star apresentou melhoria na acuracidade com a redução das *features*. No entanto, o Perceptron Multicamadas não apresentou resultados significativos com acuracidade abaixo de 90%, mesmo com melhorias de acuracidade com a árvore de decisão.

Roos e Heyns (2021) aplicaram WPD para extrair as *features* dos sinais e criou modelos de aprendizado para detecção de falhas com algoritmos de SVM e Redes Neurais Artificiais (RNA). Foi aplicado um único acelerômetro fixado na correia afim de avaliar o sinal dos rolamentos dos roletes no decorrer do percurso. Apesar da característica não-estacionária dos sinais em relação as frequências características de defeitos dos rolamentos, foi possível classificar os roletes defeituosos. Foram propostas melhorias e a incorporação do equipamento de aquisição de sinal na correia para proteção.

Bortnowski *et al* (2022) implementaram um sistema *wireless* de medição sem fio para a coleta dos sinais de vibração. Além disso, aplicou a Transformada de *Fourier* de Curta-Duração (STFT), o espectograma e parâmetros como autocorrelação, pico de frequência e a média do pico de frequência para a extração de *features*. Por fim, aplicou o aprendizado profundo com a técnica LSTM Autoencoder para classificação de falha. Foi realizado um comparativo entre testes laboratoriais e testes de campo, ao qual foi instalado um sensor móvel ao longo do transportador para detectar o local de dano do rolete. Também, foi proposto uma melhoria no algoritmo de medição, haja vista que os sinais podem captar mais fenômenos do que foi proposto experimentalmente, devido à má cooperação entre correia e rolete.

Wijaya *et al* (2021) usaram o sensor de fibra ótica para coletar de sinais acústico para aprendizado de máquina. Foi aplicada a Transformada Rápida de *Fourier* (FFT) e características no domínio do tempo para extração de *features*, além de RNA para a classificação. Foi proposto etapas de retreino e atualização do modelo quando abaixo do limite de confiança. Também houve alta acuracidade para diferentes velocidades da correia. Ademais, foi aplicado um número limitado de falhas e não foi possível simular o efeito de carregamento nas assinaturas de frequência do transportador.

Wijaya *et al* (2022) utilizaram a mesma instrumentação para coleta de dados, todavia, além do FFT também aplicou a análise envelope para extrair *features* dos sinais acústicos. Foi criada uma técnica de detecção de dano e acompanhamento da progressão da falha. Uma das desvantagens foi segmentação dos dados captados, ao qual, foram divididos ao meio ao invés de serem divididas randomicamente.

Liu *et al* (2020) instalaram um microfone para captação de sinais acústicos dos roletes para detecção de defeitos. Além disso, extraíram os coeficientes mel-cepstrais dos sinais como *features* para algoritmo de aprendizado, ao qual, aplicaram a técnica Gradient Boosting Decision Tree (GBDT). O método aplicado possibilitou a detecção de falhas em roletes baseado em sinais de sons. Além disso, o Gradient Boost auxilia na determinação das etapas e limites das *features*. Ainda assim, houve limitações referentes ao quantitativo de dados obtidos, além de propor testes futuros com ruídos de fundo.

Rocha *et al* (2021) utilizaram um sistema robótico de inspeção automatizada para coleta de sinais acústicos e de vibração ao longo dos roletes do transportador de correia. Da coleta foi utilizada a Transformada de Fourier e a média das magnitudes dos sinais para extração de *features*. Com a utilização de um robô de inspeção com diferentes tipos de sensores, foi aplicado o algoritmo de florestas aleatórias, também conhecida como *random forest* para classificar o estado dos roletes, ao qual simplifica e apresenta resultados satisfatórios.

Yang *et al* (2020) aplicaram sensores acústicos para captação dos sinais, ao qual, extraíram *features* no domínio do tempo além da aplicação de rede neurais como *stacked sparse auto encode* e os coeficientes mel-cepstrais. Para seleção das melhores *features* foi utilizado algoritmos de aprendizado não supervisionados como o algoritmo *k-means* e o agrupamento espectral. E para classificação foram escolhidas as seguintes técnicas: Rede Neural Deconvolucional (DNN), Rede Neural Convolutiva Profunda (DCNN), Máquina de Vetores de Suporte (SVM) e k-ésimo vizinho mais próximo (kNN). A abordagem do diagnóstico de falhas aplicando redes neurais de aprendizagem profunda produz melhores resultados em relação às técnicas tradicionais. Porém, aplicar redes neurais profundas exigem um alto quantitativo de dados e modelos de aprendizado mais robustos.

Peng *et al* (2020) também instalaram sensores acústicos, mas com a aplicação de WPD para extração de *features*. Já para modelo de classificação de

falhas aplicaram a Rede Neural Convolutacional (CNN). Assim como as outras técnicas de redes neurais profundas, o CNN é vantajoso quando lida com uma grande quantidade de dados. No entanto, por conta do WPD as energias das bandas de frequência mais baixas possuem proporções maiores o que aumenta a influência no diagnóstico de falhas. Além disso, foi observado a dificuldade de medição de dados de campo para cada condição.

De modo geral, nota-se que técnicas tradicionais de classificação como a SVM, o kNN e as redes neurais etc, são comumente aplicadas para diagnóstico de falhas em roletes de transportador de correia, no entanto com maior variabilidade nos métodos de extração e seleção de *features*. Para análise de vibração, a WPD torna-se uma das alternativas viáveis para extração de *features* de sinais não-estacionários, em detrimento de técnicas estatísticas e/ou no domínio de tempo-frequência (ALHARBI *et al*, 2023).

Ainda assim, técnicas de seleção de *features* como a árvore de decisão são, majoritariamente, aplicadas à parâmetros globais como *features* estatísticas e/ou dos domínios do tempo ou frequência, isoladamente. Outrossim, é possível notar que o ranqueamento de *features* por meio da Análise de Variância (ANOVA), ainda não foi, notavelmente, aplicada para diagnóstico de falhas em roletes de transportador de correia por aprendizado de máquina (ALHARBI *et al*, 2023). Portanto a combinação da WPD com técnicas de seleção de *features*, como a árvore de decisão e ANOVA, aplicada para a classificação de falhas em roletes, por meio de técnicas tradicionais de aprendizado de máquinas, será o foco desse trabalho.

2.2. Modos de falha

Por conta da necessidade de uma base de dados robusta para aprimorar o aprendizado dos modelos de classificação de falhas, demanda-se uma alta quantidade de medições de vibração/acústica no menor tempo possível, de modo que haja a diferenciação de diferentes estados dos componentes do transportador de correia. Por conta disso, os defeitos podem ser criados artificialmente, em diferentes estágios da falha, para simular o sinal de vibração em estados defeituosos (ALHARBI *et al*, 2023).

De modo geral, deve-se levantar os principais modos de falhas que motivam a quebra do transportador de correia, baseado no grau de severidade, probabilidade

de ocorrência do evento e a capacidade de detecção do defeito antes da falha. Assim, o método de análise de falha deve ser realizado com abordagens qualitativas e quantitativas, para possibilitar o mapeamento completo dos riscos de paradas não planejadas (ÖZFIRAT *et al*, 2022).

Gurjar (2012) apresentou um levantamento de falhas e danos dos componentes de um transportador de correia para análise de falha. Inicialmente, relatou-se que mais de 58% das falhas foram referentes a falhas mecânicas. Dentre essas falhas, 74,1% das falhas foram relacionadas a rolos/roletes, enquanto 7,2% estão relacionadas a problemas em tambor, como presente na Figura 2.15. Além disso, o mesmo autor levantou os principais tipos de falhas para componentes críticos do transportador, dentre eles, rolos/roletes e tambores.

Além de expor as similaridades entre rolete e tambor, dado a existência da concha e rolamento para ambos componentes, também expôs as individualidades de cada um. Para um sistema de suporte com três rolos para cada cavalete, os rolos laterais estão sujeitos a maior parte dos danos. Problemas de rolamentos nos roletes, podem criar danos na correia e sobrecarga no sistema de acionamento, incluindo tambores. Para tambores, os problemas primários estão relacionados aos rolamentos, concha e revestimento, ao qual mais de 50% dos casos estão relacionados aos rolamentos.

Zimroz e Król (2009) mostraram que os tambores mais afetados por problemas de danos e falhas são os tambores de acionamento. Além disso, dentro do levantamento as principais falhas também estão relacionadas a problemas de rolamento, porém defeitos de desgaste da concha e revestimento também são descritos. Em relação aos roletes, foi descrito a diferença de tamanho do rolete comparado ao tambor, ao qual exige maior trabalho para ser reparado. No entanto, essa diferença é compensada ao fato de os roletes apresentarem uma quantidade bastante superior no transportador, em relação aos tambores.

Fernandez *et al* (2016) identificaram que os rolamentos de roletes e tambores são as principais causas de incêndio em 5 dos 9 países investigados. Além disso, as principais causas de falhas em rolamentos são: lubrificação inadequada (36%), fadiga (34%) e contaminação (14%). Ainda assim, apenas 0,5% dos rolamentos são substituídos durante a vida útil do transportador devido à dificuldade de detecção das condições do componente. Normalmente, com o sobreaquecimento do rolete, pode

ocorrer o vazamento da graxa aquecida, ao qual surge o risco de contato entre o lubrificante com materiais inflamáveis.

Reicks (2008) descreveu os principais modos de falhas em roletes como defeitos nos rolamentos, selos e desgaste nos rolos. A vida dos rolamentos é limitada pela fadiga devido o contato entre os elementos rolantes e as pistas interna e externa. Essa fadiga é intensificada mediante sobrecargas impostas sobre o rolamento. Dentre as sobrecargas, as principais estão relacionadas as cargas axiais, radiais e angulares causadas pela dinâmica do equipamento.

Além disso, problemas de lubrificação podem ser um outro agravante, haja vista que causam deformações como amolgamentos e brinelagem nas pistas, deslizamento entre os componentes e arranhões por contaminação do lubrificante que causam início de falhas por trinca. Outro fator que afeta a vida dos rolos é o contato com a superfície da correia, ao qual pode ser influenciado pela instalação dos rolos, design e espaçamento dos cavaletes e curvatura da correia. Portanto, diâmetro do tubo, espessura e material do rolo devem ser selecionados para evitar o deslizamento do rolo sobre a correia e fuga de materiais.

Diniz *et al* (2016) aplicaram a ferramenta FMEA (*Failure Mode and Effect Analysis*) sobre o Transportador de Correia, incluindo componentes como tambores e roletes. Dentre os modos de falhas analisados, destacam-se: o revestimento do tambor solto ou desgastado devido à falta de atrito entre tambor e correia de acionamento, dano nos tambores por conta da falta de tensionamento da correia e danos nos cavaletes e roletes devido à queda de minério. Os modos de falhas citados representam um RPN (Risk Priority Number) equivalente a mais de 18% de todo o levantamento feito no trabalho.

Vasić *et al* (2020) Este trabalho explorou a análise de falha, especificamente, em rolamento de roletes e as principais causas para cada modo de falha por meio de uma árvore de falhas. Dentre as falhas, 71,2% encontram-se nos roletes laterais com inclinação de 10° a 45°. Enquanto 28,8% estão presentes nos roletes horizontais, devido à ausência de cargas axiais nos rolamentos na horizontal, se exposto somente a cargas radiais.

Dentre os modos de falhas, os principais são deformações plásticas, desgaste corrosivo, corrosão por atrito, fraturas, desgaste por fadiga superficial e desgaste adesivo e abrasivo. Em relação às fraturas (correspondente a 5% das falhas), pode

ocorrer por conta de mal manuseio, erro de montagem ou projeto e por fadiga durante a operação. Normalmente as fraturas ocorrem nos elementos rolantes ou na gaiola, ao qual pode travar o rolete e causar superaquecimento.

Domnick (2002) apresentou, individualmente, as principais condições para a falha de roletes como o travamento do rolo, furo na concha, disco lateral solto, rolamento com rangido inadequado, ruptura ou deflexão excessiva do cavalete como limitador da funcionalidade do rolete. Em relação ao tambor, Kumar e Mandloi (2013) apresentou as principais causas de falhas em tambores como falha na solda do cubo, falha no eixo ou rolamentos do tambor, desgaste ou rachadura da concha, flexão da concha e variações de carga. Desgaste na concha e falha em rolamentos representam cerca de 20% e 6% das falhas em tambores, respectivamente.

Seguindo o estado da arte acerca dos principais modos de falhas em roletes, nota-se que desgastes superficiais na concha/revestimento dos rolos, bem como, defeitos em rolamentos dos rolos, destacam-se entre os tipos de falhas mais recorrentes. Também, é importante salientar que os rolos laterais dos cavaletes apresentam maior taxa de falha, sendo então, os objetos de estudo do trabalho.

3. FUNDAMENTAÇÃO TEÓRICA

Segundo a NBR 5462 (1994) acerca confiabilidade e manutenibilidade, a manutenção pode ser conceituada como a “combinação de todas as ações técnicas e administrativas, incluindo as de supervisão, destinadas a manter ou recolocar um item em um estado no qual possa desempenhar uma função requerida”. O objetivo imprescindível da manutenção é conservar os equipamentos e máquinas em ótimas condições de execução, além de possuir uma ampla gama de funções. Admite-se, portanto, que todo equipamento apresenta desgaste e, conseqüentemente, acarretará no rompimento ou quebra de um ou mais componentes (NEPOMUCENO, 1989).

Com o passar do tempo, a visão acerca da manutenção como uma atividade a ser executada após a falha de um item ou equipamento mudou e novas estratégias foram criadas para aumentar a funcionalidade do mesmo. Devido ao aumento da competitividade industrial, as políticas de manutenção influenciam, diretamente, na produtividade e métodos e ferramentas são criados para otimizar as estratégias de manutenção, sem comprometer a gestão de custos (DE JONGE; SCARF, 2020; RUSCHEL *et al*, 2017; ALSYOUF, 2007; LUXHØJ *et al*, 1997).

Sendo assim, o tipo de manutenção ideal a ser aplicado em uma empresa é aquele que melhora a disponibilidade das máquinas ou equipamentos para a produção a um custo adequado (WEBER *et al*, 2009). Devido a diversidade de elementos e métodos aplicáveis, geralmente, há dificuldades na classificação do tipo de manutenção. Por isso, é de indubitável, classificar de modo objetivo os diversos tipos de manutenção. De modo geral, há os seguintes tipos de manutenção: Corretiva (Planejada e Não Planejada), Preventiva, Preditiva, Detectiva e Engenharia de Manutenção (KARDEC; NASCIF, 2009).

Conhecida como uma técnica de manutenção baseada na condição, a preditiva possibilita a atuação sobre a máquina dependendo dos dados atuais e futuros da máquina. Por conta disso, requer um investimento inicial elevado, ao qual é compensado a longo por evitar gastos em reparos e substituição de componentes (SAHA *et al*, 2019). A NBR 5462 também define a manutenção preditiva como a “manutenção que permite garantir uma qualidade de serviço desejada, com base na aplicação sistemática de técnicas de análise, utilizando-se de meios de supervisão centralizados ou de amostragem, para reduzir ao mínimo a manutenção preventiva e diminuir a manutenção corretiva”.

A supervisão da máquina por meio da manutenção preditiva é dada a partir do monitoramento dos parâmetros físicos que determinam a alteração do desempenho e condição, para ser tomada a decisão de intervenção sobre alguma falha ou defeito específico (KARDEC; NASCIF, 2009). De modo geral, procura-se reduzir paradas e desmontagens quando submetidos em inspeções, aumentar a disponibilidade do equipamento e atenuar as paradas por corretivas não planejadas, reduzir o agravamento dos defeitos, conseqüentemente, aumentar a vida útil do equipamento e aumentar a confiabilidade, por meio de planejamento e prioridade estratégica de manutenção sob o sistema produtivo (CYRINO, 2017).

Com o desenvolvimento da inteligência artificial interligada à equipamentos industriais, é possível desenvolver sistemas com autoconsciência, autoaprendizagem e com tomadas de decisão independentes a partir de dispositivos de detecção como: inspeções sensitivas, câmeras térmicas, sensores acústicos e de vibração (ZHANG *et al*, 2021). Dentre os parâmetros coletado por esses dispositivos, a vibração mecânica das máquinas destaca-se pela facilidade de captação por diversas fontes como rolamentos, engrenagens, desbalanço etc. Portanto, o monitoramento de vibração de máquinas é a técnica mais aplicada na manutenção preditiva (LACEY, 2011).

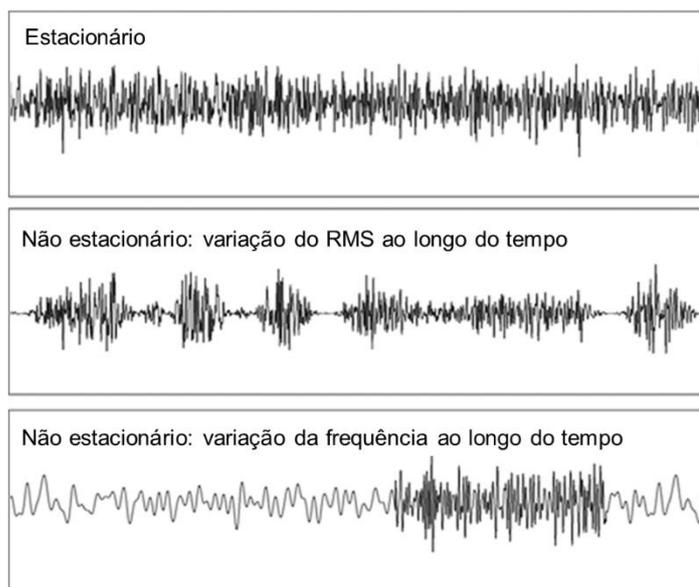
3.1. Processamento de sinais de vibração aplicado à manutenção industrial

O monitoramento de vibração torna-se um excelente indicador em ambiente industrial para avaliar a condição da máquina a partir de influências como carregamento dinâmico, comportamento da fundação e as todas influências externas. Por conta da complexidade do sinal, foram desenvolvidos parâmetros que auxiliam na identificação de anormalidades no sinal. A partir de técnicas de análise e processamento de sinais de vibração no domínio do tempo e no domínio da frequência, é possível detectar, localizar e categorizar diversos defeitos em componentes rotativos (TANDON; CHOUDHURY, 1999).

No entanto, sinais não estacionários são muito presentes na área de diagnóstico de falha em máquinas. Logo, técnicas de tempo-frequência são mais eficazes na extração de informações sobre a condição do maquinário devido a capacidade de identificar os componentes da frequência do sinal, além de revelar suas

características variantes no tempo (FENG *et al*, 2013). A Figura 3.1 apresenta as principais diferenças entre sinais estacionários e não-estacionários, na qual, os sinais não-estacionários podem ser classificados a partir da variação da frequência ou da média quadrática do sinal ao longo do tempo (ROUILLARD, 2014).

Figura 3. 1 – Tipos de dados não estacionários: (a) estacionários; (b) RMS variável no tempo; e (c) conteúdo de frequência variável no tempo



Fonte: Rouillard (2014)

3.1.1. Técnicas de tempo-frequência

As técnicas de tempo-frequência é uma das ferramentas mais novas e eficientes na área da análise de vibração. Sua eficácia está relacionada em lidar com dados não-estacionário, presentes em casos de condição operacionais rápidas como na partida do sistema de acionamento de um equipamento ou na presença de uma falha que causa descontinuidade no sinal de vibração que está sendo monitorado (ALBADOUR *et al*, 2011).

De modo geral, essas técnicas preocupam-se com a análise e processamento de sinais com a variação do conteúdo da frequência em relação ao tempo, a partir da energia do sinal distribuída no espaço de tempo-frequência (PENG *et al*, 2012). Basicamente, a análise do tempo-frequência identifica os componentes no domínio da frequência e revela as *features* no domínio do tempo, possibilitando a extração de informações acerca da condição do maquinário (FENG *et al*, 2013; XIANG; HU, 2012).

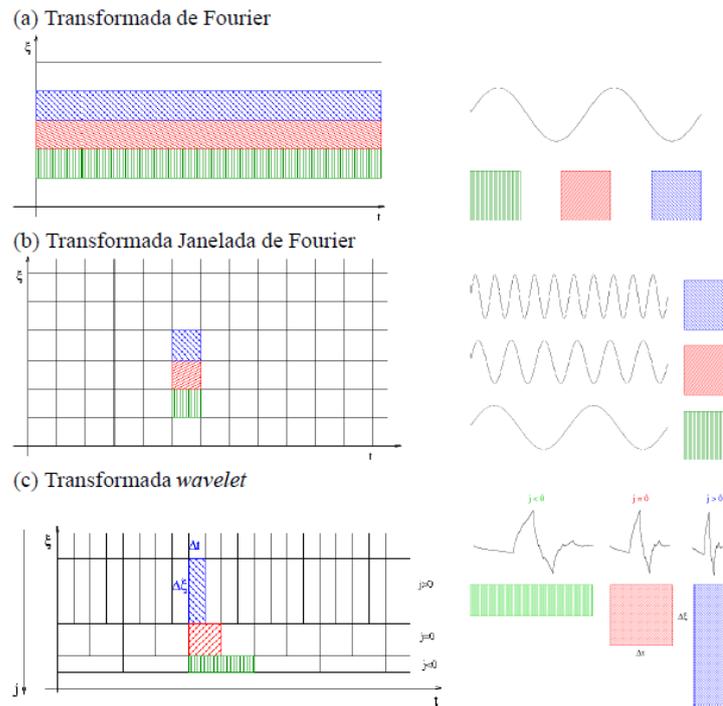
Como *features* estatísticas são pouco eficazes no domínio do tempo-frequência, haja vista que mudam com o tempo, o desenvolvimento de métodos de análise desses sinais tornou-se cada vez mais importante para o diagnóstico de falhas. Alguns dos principais métodos são as representações lineares do tempo-frequência. Essas representações são, essencialmente, processos de decomposição dos sinais a partir de somas ponderadas de bases localizadas no tempo e frequência. A resolução do tempo-frequência é governada por meio do Princípio da incerteza de *Heisenberg*, de modo que as soluções em ambos domínios não podem ser obtidas simultaneamente (FENG *et al*, 2013).

Segundo Castilho *et al* (2012), no Princípio da incerteza de *Heisenberg* “existe um balanço entre a resolução em tempo e a resolução em frequência. Se f é uma função do $\mathbb{L}^2(\mathbb{R})$, então as variações em tempo Δx e em frequência $\Delta \xi$ são inversamente proporcionais”. Essa afirmativa é corroborada pela Eq. (3.1), onde c é uma constante.

$$\Delta x^2 \Delta \xi^2 \geq c \tag{3.1}$$

Dentre as principais técnicas de tempo frequência, destacam-se a Transformada de Fourier de Curta Duração ou *Short Time Fourier Transform* (STFT) e a Transformada *Wavelet* (PENG *et al*, 2012; FENG *et al*, 2013; XIANG; HU, 2012). A Figura 3.2 apresenta a resolução nos domínios de tempo e frequência para diferentes técnicas de análise de sinais.

Figura 3. 2 – Resolução do sinal para diferentes técnicas de análise de sinais



Fonte: Castilho *et al* (2012)

3.1.2. Transformada Wavelet

Segundo Daubechies (1992), *wavelet* pode ser descrito, analiticamente, a partir de um conjunto de funções como representado abaixo na Eq. (3.2), simplificado por uma única função $\psi(x)$ (também conhecida como *wavelet* mãe), ao qual, possui um conjunto de operação de dilatações e translações (a partir dos parâmetros a e b), ao qual, formam famílias *wavelet*. Essas famílias são selecionadas como ferramenta matemática nas mais diversas áreas.

$$\psi_{a,b}(x) = |a|^{-1/2} \psi\left(\frac{x-b}{a}\right), a, b \in \mathbb{R}, a \neq 0 \quad (3.2)$$

Para aplicação em análise de sinais, os valores dos parâmetros a e b são restritos à passos de dilatação e translação (ou escala e tempo, respectivamente), de modo que $a_0 > 1$ e $b_0 \neq 0$. Então, a família *wavelet* é representada pela Eq. (3.3). Nota-se que a e b correspondem, respectivamente, às expressões a_0^m e $nb_0 a_0^m$.

$$\psi_{k,n}(x) = a_0^{-j/2} \psi(a_0^{-j} x - nb_0), m, n \in \mathbb{Z} \quad (3.3)$$

Uma função pode ser considerada *wavelet*, quando satisfaz duas condições. A primeira condição é a condição de admissibilidade viabiliza a invertibilidade da transformada *wavelet*, como descrito pela Eq. (3.4). Além disso, uma outra condição, majoritariamente, equivalente é a exigência da integral da função *wavelet* ser zero, de modo que a forma oscilatória da *wavelet* tenha média zero, como apresentado pela Eq. (3.5). A segunda condição é que a função tenha energia unitária, tal que *wavelet* possua suporte compacto ou com decaimento rápido de amplitude, como presente na Eq. (3.6) (CASTILHO *et al*, 2012).

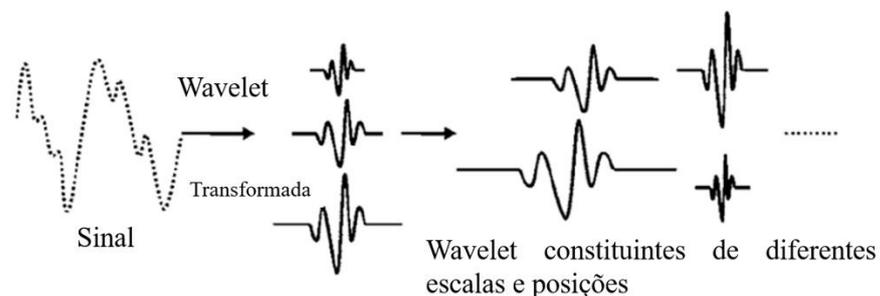
$$c_h = 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi < \infty \quad (3.4)$$

$$\int_{-\infty}^{\infty} \psi(x) dx = 0 \quad (3.5)$$

$$\int_{-\infty}^{\infty} |W(t)|^2 dt = 1 \quad (3.6)$$

De modo geral, a transformada *wavelet* aplica a base *wavelet*, de modo análogo as funções senoidais aplicadas para a Transformada de Fourier. A principal diferença é a forma das ondas senoidais, aos quais são funções periódicas com amplitude constante para além do domínio, enquanto a base *wavelet* são funções de curta duração com valor linear zero fora do domínio (SILVA, 2019; TORRENCE; COMPO, 1998). A forma da base *wavelet* pode ser vista na Figura 3.3.

Figura 3. 3 – Base wavelet



Fonte: Adaptado de Sygouni; Tsakiroglou; Payatakes (2006)

A família de funções de base wavelet são compostas de funções de alta frequência e curta duração e funções de baixa frequência e longa duração. Essas funções são determinadas a partir do domínio da frequência, as quais podem contrair e expandir a wavelet mãe por meio do parâmetro de dilatação a , para demonstrar características de alta e baixa frequência, respectivamente. Para evitar redundância de dados, bem como, cálculos excessivos em todas as escalas possíveis, pode-se discretizar os parâmetros a e b de modo que a análise de sinais continue, eficientemente, precisa (YEN; LIN. 2000).

Além disso, para Castilho *et al* (2012), a Transformada Wavelet Discreta (DWT) é uma análise de sinais não estacionários para a detecção das estruturas dos domínios espaciais e/ou temporais, além da extração de informações por meio das variações de frequência. Dessa forma, a DWT é implementada a partir de valores discretos de escala nas expressões com o objetivo de dilatar e/ou transladar a *wavelet*. Ao contrário de outras técnicas, como a STFT, a DWT oferece maior flexibilidade na representação de um sinal no domínio tempo-frequência devido ao uso de janelas de tamanho variável com intervalos de tempo mais longos para baixas frequências e intervalos mais curtos para altas frequências.

Para a aplicação da DWT, deve ser associado à *wavelet* discreto que mapeia funções f em \mathbb{Z}^2 , como descrito na expressão da Eq. (3.7), considerando $k = nb_0$.

$$(Tf)_{jk} = \langle \psi_{jk}, f \rangle = a_0^{-j/2} \int dx \overline{\psi(a_0^{-j}x - k)} f(x) \quad (3.7)$$

Na implementação da DWT, foi introduzida a Análise de Multirresolução (MRA), afim de adequar dos sinais de tempo discreto de comprimento finito (MALLAT, 1989; YEN; LIN. 2000). Conseqüentemente, foi desenvolvido um algoritmo, computacionalmente eficiente, conhecido como Transformada Rápida *Wavelet* ou *Fast Wavelet Transform* (FWT).

Para muitas aplicações, as primeiras formas *wavelet* criadas eram altamente redundantes, isto é, eram dependentes devido ao fato de cada 'onduleta' residir na região de valor linear zero da outra. Logo, um subespaço criado pela extensão da transformada *wavelet* era menor, conforme o aumento da redundância da banda. Para reduzir a redundância, uma alternativa é escolher os parâmetros a_0 e b_0 , de modo que $h_{m,n}$ constitua uma base ortonormal. Por conseguinte, indica-se que a base *wavelet*, a

partir dos pressupostos acima, podem ser concebidos mediante a Análise de Multirresolução (DAUBECHIES, 1988).

A MRA foi uma alternativa criada para criar funções *wavelet* a partir de determinadas condições escolhidas de acordo com as características dos sinais a serem analisados. Sendo assim, é possível criar famílias de *wavelets* que obedecem determinadas condições. Dentre essas condições a transformação feita pelo MRA, as principais são: a inversibilidade da operação do DWT, ao qual as informações de escalas mais finas devem ser recuperáveis a partir dos dados de multiescala. Esse processo reverso é a Transformada Discreta Wavelet Inversa (IDWT); tanto o DWT quanto o IDWT devem ser operados por algoritmos eficientes; e por fim, os algoritmos devem ser estáveis assumindo baixa amplificação de perturbações de dados (CASTILHO *et al*, 2012).

Considerando $(V_j)_{j \in \mathbb{Z}}$ uma sequência de subespaços fechados de $\mathbb{L}^2(\mathbb{R})$, então pode-se definir que é uma aproximação de multirresolução de $\mathbb{L}^2(\mathbb{R})$, desde que obedeça às seguintes propriedades, exibidas pelas Eqs. (3.8) a (3.12) (MALLAT, 1988; DOMINGUES; KAIBARA, 2012).

$$V_j \subset V_{j+1}, \forall j \in \mathbb{Z} \quad (3.8)$$

$$\bigcup_{i=-\infty}^{+\infty} V_j = L^2(\mathbb{R}) \text{ e } \bigcap_{j=-\infty}^{\infty} V_j = \{0\} \quad (3.9)$$

$$f(x) \in V_j \stackrel{n}{\Leftrightarrow} f(2x) \in V_{j+1}, \forall j \in \mathbb{Z} \quad (3.10)$$

$$f(x) \in V_0 \Leftrightarrow f(x - k) \in V_0, \forall k \in \mathbb{Z} \quad (3.11)$$

$\phi(x - k)_{k \in \mathbb{Z}}$, é uma base de Riesz de V_0 , tal que

$$A \sum_k |c_k|^2 \leq \|c_k \phi(x - k)\|^2 \leq B \sum_k |c_k|^2, \text{ sendo} \quad (3.12)$$

$$0 < A, B < \infty, \forall (c_k)_{k \in \mathbb{Z}} \in l^2$$

A partir das propriedades de multirresolução expressadas, a aproximação de uma função $f(x)$ e $\mathbb{L}^2(\mathbb{R})$ com uma resolução 2^j é definida como a projeção ortogonal

de $f(x)$ sobre V_j . Para calcular esta projeção ortogonal mostramos que existe uma única função, conhecida como função escala, $\phi(x) \in \mathbb{L}^2(\mathbb{R})$ tal que, para qualquer $j \in \mathbb{Z}$ a função é uma base ortonormal de V_j , como descrito pela Eq. (3.13) (MALLAT, 1988).

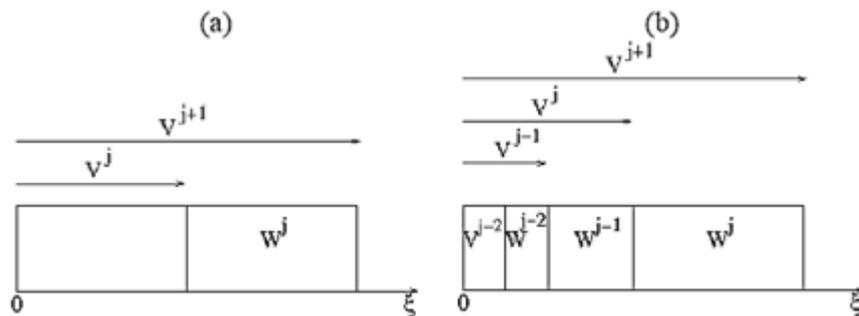
Analogamente, dado W_j como um complemento ortogonal de V_j , a função $h(x) \in \mathbb{L}^2(\mathbb{R})$ tal que, para qualquer $j \in \mathbb{Z}$, a função é um base *wavelet* ortonormal de W_j , como descrito pela Eq. (3.14). Considerando a propriedade exibida pela Eq. (3.8) e os pressupostos já apresentados, pode-se afirmar, portanto, que $V_{j-1} = V_j \oplus W_j$ (MALLAT, 1988).

$$\phi_k^j(x) = \sqrt{2^j} \phi(2^j x - k), k \in \mathbb{Z} \quad (3.13)$$

$$\psi_k^j(x) = \sqrt{2^j} \psi(2^j x - k), (k, j) \in \mathbb{Z} \quad (3.14)$$

Para $j = 1$, a Eq. (3.13) pode ser descrita pela Eq. (3.15), onde a_k é descrita pela Eq. (3.16). A Eq. (3.15) é chamada de relação de escala, onde h_k é conhecida como coeficiente de escala ou coeficiente de filtros (DOMINGUES & KAIBARA, 2012; DA SILVA, 2022). Castilho *et al* (2012) exemplifica a MRA a partir das aproximações entre os subespaços V_j , considerando W_j como o espaço complementar na diferença entre os subespaços. A Figura 3.4 representa o funcionamento das resoluções múltiplas a partir dos subespaços.

Figura 3. 4 – Aproximações ARM. a) Dois níveis. b) Multiníveis



Fonte: Adaptado de Castilho *et al* (2012)

$$\phi(x) = 2 \sum_{k \in \mathbb{Z}} h_k \phi(2x - k), \quad (3.15)$$

$$h_k = \sqrt{2} \langle \phi, \phi_{1,k} \rangle_{\mathbb{L}^2(\mathbb{R})} = 2 \int_{-\infty}^{\infty} \phi(x) \overline{\phi(2x - k)} dx \quad (3.16)$$

De modo geral, Daubechies (1988) afirma que as propriedades acima possibilitam a criação de uma base ortonormal para a criação de sucessivas aproximações de funções f em L^2 , cada qual como uma versão suavizada de f , criando um conjunto de funções concentradas com diferentes resoluções, que definem a ARM.

A partir da MRA, é possível a construção da multirresolução biortogonal (MRAb), ao qual os pares $\{V_j, \phi\}$ e $\{V_j^*, \phi^*\}$ satisfazem a seguinte relação, explícita pela Eq. (3.17). De forma análoga, temos que $V_{j-1}^* = V_j^* \oplus W_j^*$. Por conseguinte, são definidas as Eqs. (3.18) e (3.19) (DOMINGUES; KAIBARA, 2012).

$$\langle \phi(y - k), \phi^*(y - k) \rangle := \int \phi(y - k) \phi^*(y - k) dx = \delta_{k,l} \quad (3.17)$$

$$\psi(x) = 2 \sum_{k \in \mathbb{Z}} g_k \phi(2x - k) \quad (3.18)$$

$$\psi^*(x) = \sqrt{2} \sum_{k \in \mathbb{Z}} g_k^* \phi^*(2x - k) \quad (3.19)$$

Sendo g_k e g_k^* correspondentes às Eq. (3.20) e (3.21). Tanto g_k quanto h_k serão aplicadas na DWT, ao qual, o sinal é decomposto, iterativamente, a partir do processo de filtragem digital como filtros passa-alta e passa-baixa, associados a uma função *wavelet* dentro de um espaço detalhado D_j e a uma função de escala dentro um espaço de aproximação A_j , respectivamente (WANG *et al*, 2013).

$$g_k = (-1)^{k+1} h_{1-k}^* \quad (3.20)$$

$$g_k^* = (-1)^{k+1} h_{1-k} \quad (3.21)$$

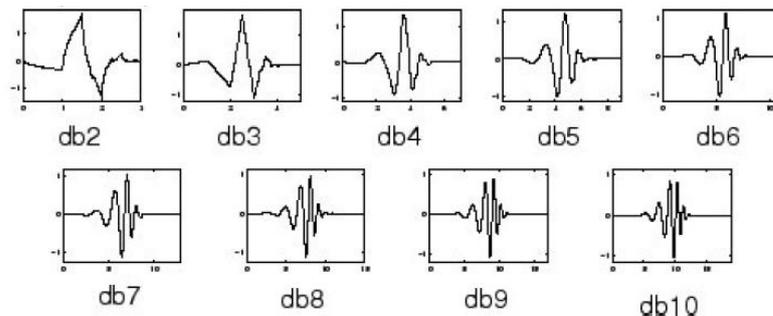
Esse mesmo processo de filtragem continua decompondo as bandas de baixa frequência até finalizar as iterações. Por fim, coeficientes wavelet de aproximação $A_{(j)}$ e de detalhe $D_{(j)}$ são extraídos a partir do produto interno das funções de escala e *wavelet*, respectivamente, com o sinal $f(x)$, expressos pelas Eq. (3.22) e (3.23) (WANG *et al*, 2013).

$$A_{(j)}(x) = f(x) \cdot \phi_k^j(x) \quad (3.22)$$

$$D_{(j)} = f(x) \cdot \psi_k^j(x) \quad (3.23)$$

Para atender as condições expostas acima, havia restrições acerca sobre as funções permitidas para a MRA. Como uma alternativa, foi criada a família de funções *wavelet* de *Daubechies*, como ilustrado na Figura 3.5 (ROWE; ABBOTT, 1995). A família *Daubechies*, bem como outras famílias *wavelet*, não possui expressão analítica e é obtida a partir de algoritmos criadas de acordo com as propriedades exigidas para a análise (DA SILVA, 2019).

Figura 3. 5 – Família Daubechies



Fonte: Silva (2019)

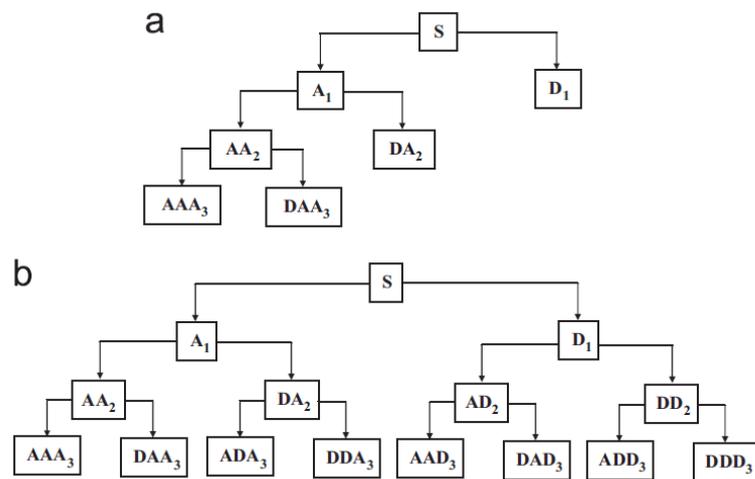
3.1.3. Wavelet Packet Decomposition

Apesar da flexibilidade das propriedades de resolução da DWT, uma das desvantagens é a fraca resolução na alta frequência, o que dificulta a discriminação entre componentes dos sinais nas frequências mais altas. Logo, a criação de *Wavelet Packet Decomposition* (WPD) a partir da generalização das bases wavelet, surgem

como bases alternativas que herdaram propriedades de ortonormalidade e localização de tempo-frequência para as funções wavelet correspondentes (YEN; LIN. 2000).

Enquanto o WPD oferece uma análise mais rica, a partir da decomposição de sinais por filtragem digital em todas as bandas de frequência, ao qual cria-se um conjunto de sub-bandas de frequência aplicadas para melhor discriminação de componentes em todo domínio da frequência (WANG *et al*, 2013). A Figura 3.6 representa a diferença entre as estruturas de decomposição do DWT e WPD.

Figura 3. 6 – Estrutura do a) DWT e b) WPD



Fonte: Wang *et al* (2013)

Por conseguinte, o princípio básico do WPD é a decomposição do sinal a partir da transformada em bandas de baixa a alta frequência. Para cada banda, era extraída a energia do espectro. Energias de diferentes bandas de frequência de um sinal de vibração podem ser usadas como recursos para permitir a identificação de falhas usando um algoritmo classificador inteligente. As Equações (3.24), (3.25) e (3.26) simplificam a função *wavelet*, coeficientes *wavelet* e energia de banda, respectivamente. O n representa o nível de decomposição, j é o fator de escala, k é o fator de translação e $f(t)$ é o sinal no domínio do tempo (LI *et al*, 2013).

$$W_{j,k}^n(t) = 2^{\frac{j}{2}} W^n(2^j t - k) \quad (3.24)$$

$$w_{j,k}^n = \int_{-\infty}^{+\infty} f(t) W_{j,k}^n dt \quad (3.25)$$

$$E(J, i) = \|w_{j,k}^n\|^2 \quad (3.26)$$

3.2. **Aprendizado de máquina**

O aprendizado de máquina é a ciência que estuda algoritmos e modelos estatísticos com o objetivo de sistemas computacionais executarem determinadas atividades, sem programações com comandos explícitos. Dentre as atividades, o algoritmo pode atuar na mineração de dados, processamento de imagens, análise preditiva etc. Sendo assim, com o modelo gerado a partir do aprendizado, o computador pode executar essas tarefas automaticamente (MAHESH, 2020).

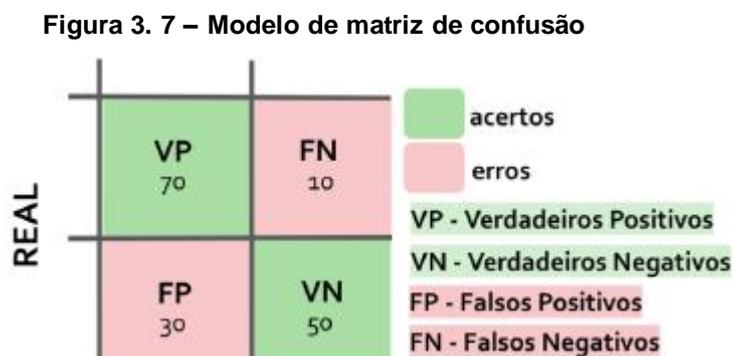
De acordo com Géron (2019), dentre os tipos de aprendizado, pode-se classifica-los nas seguintes categorias:

- **Supervisionado:** Apresenta algoritmos que rotulam as informações sobre os dados, de modo que o computador assimile as soluções expostas e consiga predizer novos dados;
- **Não supervisionado:** As informações não possuem rótulos dos dados e tentam criar regras sem supervisão;
- **Semi-supervisionado:** Possuem características de aprendizado Supervisionado e Não Supervisionado em diferentes etapas do algoritmo;
- **Reforço:** Algoritmo ao qual fornece ‘recompensas’ ou ‘penalidades’ a partir das regras e ações criadas.

Uma outra abordagem acerca do Aprendizado de Máquina define quais os tipos de dados que o computador irá predizer. Segundo Izbicki e Santos (2020), a predição é o resultado de uma variável $Y \in \mathbb{R}$, ao qual, é dependente de um vetor de entrada $x = (x_1, \dots, x_d) \in \mathbb{R}^d$. Esse vetor representa as principais características dos dados, também conhecidas como *features*. Quando a predição fornece Y como uma variável quantitativa, define-se que o aprendizado está solucionando um problema de Regressão. Enquanto, se Y for uma variável qualitativa, o problema em questão é denominado de Classificação.

A avaliação do nível de precisão (também conhecido como acurácia) de um modelo de aprendizado de máquina pode ser feita a partir da matriz de confusão. O modelo é criado com dados de treino como *features* e então o modelo realizará

predições a partir de novos dados de entrada (dados de teste), ao qual o número de predições corretas e incorretas serão expressos conforme a Figura 3.7 (Ruback *et al*, 2021)



Fonte: Ruback *et al* (2021)

Há diversos algoritmos que podem ser aplicados para o aprendizado de máquina, ao qual, cada um se adequa, mais efetivamente, para solucionar determinado problema considerando, também, o número de variáveis disponíveis para a base de dados (MAHESH, 2020). Para os algoritmos de classificação, existem técnicas como a Máquina de Vetor de Suporte Machine, Método do k -ésimo vizinho mais próximo e, as Redes Neurais Artificiais.

3.2.1. Máquina de vetores de suporte

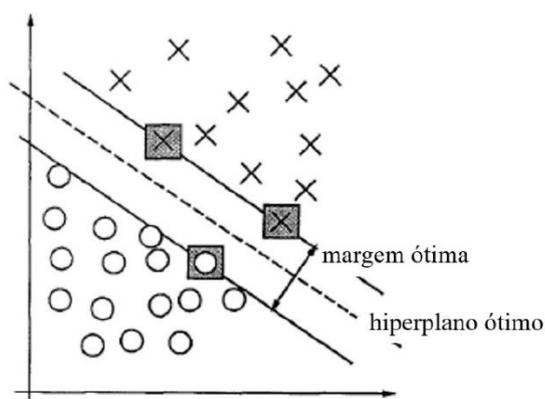
Para problemas de classificação a partir da obtenção de uma superfície de decisão, há dois problemas: um conceitual e um técnico. Na abordagem conceitual é preciso encontrar um hiperplano de separação para espaços de *features* de alta dimensionalidade que generalize bem o modelo de aprendizado, ao qual, nem todos os hiperplanos obtidos classificam bem os dados. Em relação a abordagem técnica, o tratamento computacional precisa lidar com espaços de alta dimensionalidade. Sendo assim, a solução proposta para o problema conceitual foi a criação de uma técnica conhecida como *Support Vector Machine* ou Máquina de Vetores de Suporte (SVM), na qual define-se um hiperplano ótimo com vetores de separação de classes (CORTES; VAPNIK, 1995).

A SVM é considerada uma das principais ferramentas na ciência e indústria, devido a efetividade de classificação às aplicações dadas em relação as outras

técnicas (BRUNTON; KUTZ, 2019). De modo geral, a SVM é uma técnica de aprendizado de máquina, ao qual, é criado um hiperplano com separação ótima de vetores de entrada mapeados, não linearmente, em um espaço \mathbb{Z} de *features* de alta dimensionalidade. Para o hiperplano ótimo (Figura 3.8) garantir uma boa generalização das classes, são construídas margens de máxima distância entre o vetor mais próximo.

Sendo assim, é possível generalizar modelos que classifiquem dado linearmente separáveis e não separáveis (VAPNIK, 1999). As margens são criadas a partir de uma pequena parcela de dados denominados Vetores de Suporte (CORTES; VAPNIK, 1995).

Figura 3. 8 – Hiperplano de separação ótima



Fonte: Adaptado de Cortes; Vapnik (1995)

Considerando um conjunto de dados com l classes,

$$(x_1, y_1), \dots, (x_l, y_l), x \in \mathbb{R}^n, y \in \{+1, -1\} \quad (3.27)$$

Dessa forma, o hiperplano de separação de dados, segue a seguinte notação na Eq. (3.28), sendo o vetor ω e o escalar b pesos na equação do hiperplano. Sendo o hiperplano de separação ótimo descrito pela Eq. (3.29). Além disso, a distância máxima que determina a margem máxima entre as classes é dada na Eq. (3.30).

$$y_l[(\omega \cdot x) - b] \geq 1, \quad i = 1, \dots, l \quad (3.28)$$

$$\omega_0 \cdot x - b_0 = 0 \quad (3.29)$$

$$\rho(\omega_0, b_0) = \frac{2}{|\omega_0|} = \frac{2}{\sqrt{\omega_0 \cdot \omega_0}} \quad (3.30)$$

O hiperplano ótimo pode ser checado em duas condições: satisfazendo as restrições da Eq. (3.28) e sendo mapeada não linearmente pela Eq. (3.31), ao qual constrói-se um Lagrangiano com multiplicador α_i como técnica de otimização padrão, regido pela Eq. (3.32).

$$\Phi(\omega) = \omega \cdot \omega \quad (3.31)$$

$$L(\omega, b, \alpha) = \frac{1}{2}(\omega \cdot \omega) - \sum_{i=1}^l \alpha_i \{[(x_i \cdot \omega) - b]y_i - 1\} \quad (3.32)$$

Tendo isso, é possível encontrar os pontos críticos a partir da derivação da Eq. (3.32) nas Eq. (3.33) e Eq. (3.34).

$$\left. \frac{\partial L(\omega, b, \alpha)}{\partial \omega} \right|_{\omega=\omega_0} = \left(\omega_0 - \sum_{i=1}^l \alpha_i y_i x_i \right) = 0 \quad (3.33)$$

$$\left. \frac{\partial L(\omega, b, \alpha)}{\partial b} \right|_{b=b_0} = \sum_{\alpha_i} \alpha_i y_i = 0 \quad (3.34)$$

Os vetores de suporte são os únicos vetores que consideram $\alpha_i \neq 0$ na expansão de ω_0 , além de serem restritos as condições da Eq. (3.28). Logo, é obtida a Eq. (3.35) abaixo.

$$\omega_0 = \sum_{\text{Vetores de suporte}} \alpha_i y_i x_i \quad (3.35)$$

A partir das condições do teorema de Kuhn-Tucker, o ω_0 obtido em (3.35) é integrado ao Lagrangiano, ao qual é dada pela Eq. (3.36). Por fim, a norma do vetor ω_0 é dada pela Eq. (3.37), sendo $\alpha_0 = (\alpha_1, \dots, \alpha_l)$.

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (3.36)$$

$$[\omega_0] = 2W(\alpha_0) = \sum_{\text{Vetores de suporte}} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \quad (3.37)$$

De forma geral, a regra de separação baseada no hiperplano ótimo é dada pelo indicador representado pela Eq. (3.38), ao qual, x_i são os vetores de suporte e b_0 é a constante presente na Eq. (3.39), sendo $x(1)$ um vetor de suporte pertencente a primeira classe, enquanto $x(-1)$ é um vetor de suporte pertencente a segunda classe.

$$f(x) = \text{sign} \left(\sum_{\text{Vetores de suporte}} y_i \alpha_i (x_i \cdot x) - b_0 \right) \quad (3.38)$$

$$b_0 = \frac{1}{2} [(\omega_0 \cdot x(1)) + (\omega_0 \cdot x(-1))] \quad (3.39)$$

Para solucionar o problema técnico, foi demonstrado a possibilidade de trocar a ordem de operação para construir a função de decisão, ao qual, deve-se primeiro comparar os dois vetores de entrada no espaço, a partir do produto escalar e então transformar não-linearmente o valor do resultado. Logo, é possível construir uma maior variedade de superfícies de decisão, incluindo superfícies não lineares (CORTES; VAPNIK, 1995).

Para isso, não é necessário considerar um espaço \mathbb{Z} de *features* de alta dimensionalidade de forma explicitar, sendo somente preciso calcular o produto interno entre os vetores de suporte e os vetores do espaço de *features*. Assim, a Eq. (3.40) representa a expressão geral do produto interno, sendo K , uma função simétrica.

$$(z_i \cdot z) = K(x_i \cdot x) \quad (3.40)$$

Logo, é possível criar uma função de decisão a partir do produto interno, representada pela Eq. (3.41). De forma análoga, encontra-se os parâmetros α_i a partir da Eq. (3.42). A Eq. (3.41) é função de decisão linear.

$$f(x) = \text{sign} \left(\sum_{\text{Vetores de suporte}} y_i \alpha_i K(x_i \cdot x) - b \right) \quad (3.41)$$

$$W(\alpha) = \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j K(x_i \cdot x_j) \quad (3.42)$$

Para o caso de superfície de decisão não lineares há variações na convolução produto interno. Para o caso de superfícies polinomiais de grau d , o produto interno é definido pela Eq. (3.43). Logo, a função de decisão polinomial é dada pela Eq. (3.44).

$$K(x_i, x) = [(x \cdot x_i) + 1]^d \quad (3.43)$$

$$f(x) = \text{sign} \left(\sum_{\text{Vetores de suporte}} y_i \alpha_i [(x \cdot x_i) + 1]^d - b \right) \quad (3.44)$$

Outro tipo de superfície não linear são as Funções de Base Radial (RBF), ao qual a função de decisão é dada pela Eq. (3.45). Sendo $K_\gamma(|x - x_i|)$ expressa pela Eq. (3.46) e γ é um parâmetro não-negativo de largura da função.

$$f(x) = \text{sign} \left(\sum_{i=1}^N \alpha_i K_\gamma(|x - x_i|) - b \right) \quad (3.45)$$

$$K_\gamma(|x - x_i|) = \exp\{-\gamma|x - x_i|^2\} \quad (3.46)$$

De forma geral, é possível mapear diferentes tipo de função de decisão quando se conhece a fórmula de K , também conhecida como *Kernel*, muito úteis para lidar em casos de dados não separáveis linearmente (BOSWELL, 2002).

3.2.2. K-ésimo vizinho mais próximo

Assim como o SVM, o k-ésimo vizinho mais próximo ou *k-nearest neighbors* (kNN) é umas técnicas tradicionais de aprendizado de máquina, amplamente utilizadas, devido a simplicidade de aplicação nas mais diversas áreas (BATISTA; SILVA, 2009). A técnica consiste na classificação de um conjunto de dados baseado na classe dos dados ou vizinhos mais próximos, ao qual a técnica é dividida, basicamente, em dois estágios: a determinação dos k-ésimo vizinho mais próximo e a determinação das classes dada aos vizinhos (CUNNINGHAM; DELANY, 2021).

Assumindo um espaço n-dimensional \mathfrak{X}^n , os vizinhos mais próximos são definidos a partir da distância padrão Euclidiana. Essa distância é arbitrada pelo vetor $\langle x, f(x) \rangle$, aonde f é a função que fornece o valor da classe para cada instância x . Além disso, $\langle a_1(x), a_2(x), \dots, a_n(x) \rangle$ é vetor das *features* que descreve a instância x . Sendo assim, considerando a r-ésima *feature*, a distância Euclidiana padrão entre duas instâncias $d(x_i, x_j)$ é descrita pela Eq. (3.47). A função f pode apresentar valores discretos, de forma que, $f: \mathfrak{X}^n \rightarrow \mathcal{V}$, onde \mathcal{V} é um conjunto finito de valores de classe $\{v_1, \dots, v_s\}$ (BATISTA; SILVA, 2009).

$$d(x_i, x_j) = \sqrt{\sum_{r=1}^n (a_r(x_i) - a_r(x_j))^2} \quad (3.47)$$

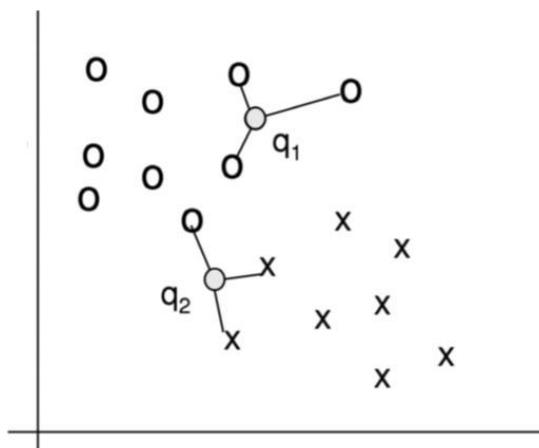
Com o cálculo e a ordenação de todas as distâncias, desde a menor até a maior, pode ser determinado os k pontos mais próximos, a partir dos dados de treinamento. Logo, a função preditora que classifica um dado desconhecido, a partir de k rótulos estabelecidos, pode ser descrito a partir da Eq. (3.48) (WANG *et al*, 2016).

$$f(x_q) = \arg \max_{v \in \mathcal{V}} \sum_{i=1}^k \delta(v, f(x_i)) \quad (3.48)$$

Sendo x_q a instância de teste, cuja distância é denotada por k instâncias do conjunto de x_i . Além disso, $\delta(v, f(x_i)) = 1$ se $v = f(x_i)$ e $\delta(v, f(x_i)) = 0$ se $v \neq f(x_i)$. Por fim, $f(x_q)$ retorna a moda da função f para todos os k treinos mais próximos de x_q (BATISTA; SILVA, 2009). De modo geral, a Figura 3.9 ilustra um exemplo de como

a técnica vai atuar para casos de classificação binária para um espaço de *features* em 2D. Considerando os três vizinhos mais próximos, os dados devem ser segregados entre as classes O e X . Destarte, q_1 pertencerá a classe O , por possuir os três vizinhos mais próximos pertencentes a mesma classe. Enquanto, q_2 pertencerá a classe X , apesar de um dos vizinhos ser da classe oposta, devido a moda do conjunto de vizinhos mais próximos pertencerem a classe X (CUNNINGHAM; DELANY, 2021).

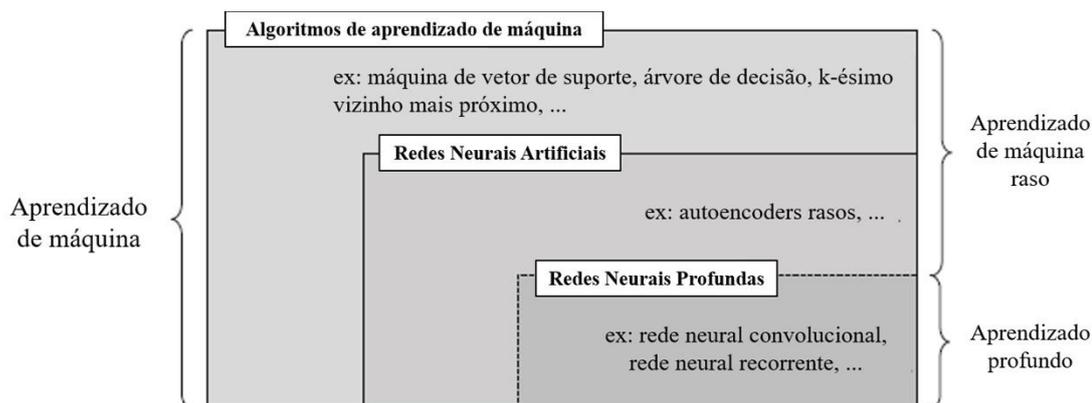
Figura 3. 9 – Classificação para três vizinhos mais próximos



Fonte: Adaptado de Cunningham; Delany (2021)

3.2.3. Rede Neurais Artificiais

O aprendizado de máquinas possui limitações, haja vista que o nível de eficiência do algoritmo é melhorado, conforme aumenta a quantidade de treinos dos modelos. Por conta disso, com o passar do tempo, novas abordagens de aprendizado. A Figura 3.10 apresenta o diagrama de Venn com o nível de evolução do aprendizado de máquinas (JANIESCH *et al*, 2021).

Figura 3. 10 – Evolução do aprendizado de máquinas

Fonte: Adaptado de Janiesch *et al* (2021)

A Redes Neurais Artificiais é uma técnica de aprendizado de máquinas que aplica a comunicação dos dados a partir de camadas, ao qual, camadas intermediárias são denominadas de camadas ocultas. O aprendizado profundo é caracterizado como um aprendizado por redes neurais artificias com mais de uma camada oculta (JANIESCH *et al*, 2021).

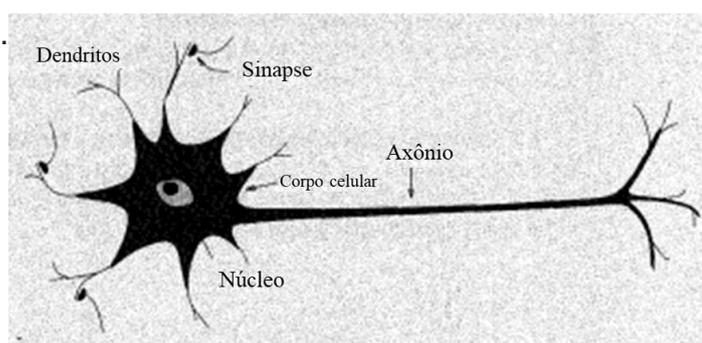
As redes neurais artificiais (RNA) são um conjunto de técnicas de aprendizado de máquina, cujo, diferencial é o aprendizado de sistemas complexos de entrada-saída de dados. As RNAs têm como paradigma o funcionamento biológico do sistema nervoso para processamento de informações, reconhecimento de padrões e classificação de dados. Este funcionamento ocorre mediante um conjunto de neurônios interconectados, cuja comunicação entre os neurônios, também conhecida como sinapse, é criada a partir de algoritmos não-lineares que mapeam e extraem as principais características dos dados de entrada (BASU *et al*, 2010).

Dentre as principais vantagens das RNAs, destaca-se a representação distribuída e computação, a capacidade de aprendizagem e generalização, a adaptabilidade, a tolerância a falhas e o baixo consumo de energia. E apesar dos computadores modernos superarem os humanos na execução de determinadas atividades, os seres humanos resolvem com maior agilidade problemas de percepções complexas, como o reconhecimento de uma pessoa no meio da multidão apenas observando o rosto. Este desempenho ocorre devido a complexa arquitetura das redes neurais biológicas, diferente dos convencionais modelos de aprendizados computadorizadas (JAIN *et al*, 1996).

O neurônio é uma célula pertencente ao sistema nervoso, responsável por processar informações (ilustrado pela Figura 3.11). É basicamente formado pelo corpo

celular, além de dois tipos de ramificações: os axônios e os dendritos. Além disso, o corpo celular possui um núcleo, ao qual armazena informações acerca das características hereditárias da célula. O neurônio recebe sinais de outros neurônios por meio dos dendritos que funcionam como receptores e emite sinais por meio de impulsos gerados pelo axônio ao longo do corpo celular. O sinal é ramificado em filamentos cada vez menores, até chegar em terminais, denominados de sinapses, que são unidades funcionais entre dois neurônios (entre dendrito de um neurônio e o axônio de outro). Nas sinapses substâncias químicas, conhecidas como neurotransmissores, são difundidas no espaço entre sinapses, possibilitando a geração de impulsos elétricos. A eficácia da sinapse pode ser ajustada por meio dos sinais transmitidos pela célula, mediante o aprendizado das atividades recebidas (JAIN *et al*, 1996).

Figura 3. 11 – Estrutura do neurônio biológico



Fonte: Adaptado de Jain *et al* (1996)

De modo geral, o conjunto de neurônios executam atividades responsáveis por tomadas de decisão complexas, como o reconhecimento facial, na escala de milissegundos, o que é facilitado devido a distribuição entre as interconexões das sinapses, por isso é conhecido por redes neurais. Por conta disso, as RNAs se inspiram nas interconexões para um melhor aprendizado e processamento de informações (JAIN *et al*, 1996).

Por conseguinte, os modelos de RNAs apresentam modelos responsáveis pelo reconhecimento estatísticos de padrões, sendo um dos principais modelos as redes neurais *Feed-forward*, também conhecido como Perceptron Multicamadas (MLP). Para entender esse modelo, deve-se, inicialmente, considerar a forma funcional da rede, com inclusão da parametrização específica das funções de base. Além disso, deve-se determinar os parâmetros de rede que fornecem melhor

desempenho de aprendizado, incluindo a otimização de sistemas não lineares. Para solucionar esse problema, é aplicada a Retropropagação, com o objetivo de reduzir o erro dos parâmetros de rede, e tornar o modelo com máxima verossimilhança (BISHOP; NASRABADI, 2006).

Considerando um conjunto de variáveis de entrada x_1, \dots, x_D , é construído outro conjunto com M combinações lineares, a partir da Eq. (3.49), onde $j = 1, \dots, M$. Nessa equação há dois tipos de parâmetros: ω_{ji} são os pesos, enquanto ω_{j0} é conhecido como *bias* ou viés. Os parâmetros são pertencentes a primeira camada da rede neural.

$$a_j = \sum_{i=1}^D \omega_{ji} x_i + \omega_{j0} \quad (3.49)$$

Ademais, a_j é conhecida como ativação, a qual é transformada a partir de uma função de ativação não linear diferenciável $h(\cdot)$ representada pela Eq. (3.50), onde z_j representa as saídas dos dados de entrada, também conhecidas como unidades ocultas.

$$z_j = h(a_j) \quad (3.50)$$

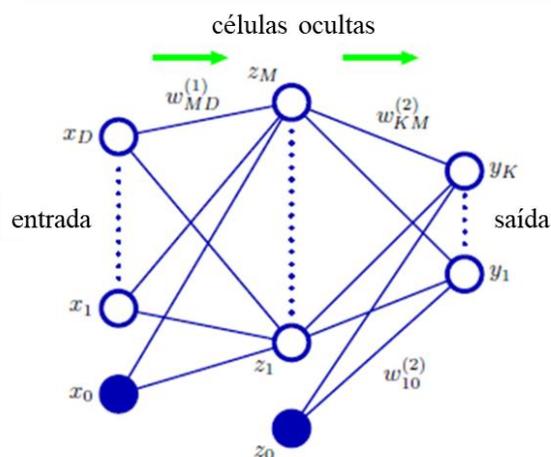
Por conseguinte, z_j será combinada linearmente na Eq. (3.51), que corresponde a ativação das unidades de saída. Sendo $k = 1, \dots, K$ o número de saídas e n a n -ésima camada oculta ou a $(n+1)$ -ésima camada da rede neural.

$$a_k = \sum_{j=1}^M \omega_{kj}^{(n)} z_j + \omega_{k0} \quad (3.51)$$

Por fim, na última camada oculta haverá a transformação, por meio da função de ativação, para a estimação das unidades da camada de saída da rede neural, como presente na Eq. (3.52). A Figura 3.12 ilustra, de modo geral, o diagrama da rede neural artificial (BISHOP; NASRABADI, 2006).

$$y_k = h(a_k) \quad (3.52)$$

Figura 3. 12 – Diagrama de rede para uma camada oculta



Fonte: Adaptado de Bishop e Nasrabadi (2006)

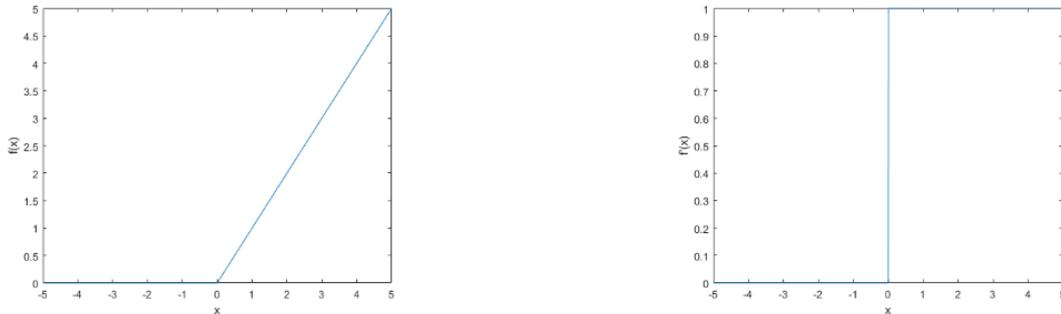
As funções de ativação associam os neurônios entre si, possibilitando a comunicação entre camadas. Sendo assim, é possível mapear e processar os dados, além de criar sistemas de aprendizado complexos, capazes de representar modelos e saídas inteligentes. Uma das principais características de uma função de ativação é a diferenciabilidade, responsável pela implementação de estratégias de otimização, a partir da Retropropagação, a qual calcula os erros e perdas, em relação aos pesos, por meio do Gradiente Descendente, a fim de aumentar a acuracidade do modelo. Dentre as principais funções de ativação, destacam-se as funções: Linear, Sigmóide, Tangente hiperbólica (Tanh) e a Unidade Linear Retificada (ReLU) (FENG; LU, 2019; SHARMA *et al*, 2017).

A função de ativação ReLU, tem sido amplamente aplicada, recentemente, nas redes neurais. Considerada uma função não linear, é descrita pela Eq. (3.53) e sua derivada pela Eq. (3.54). Sendo assim, com a função ReLU é possível desativar alguns neurônios da rede quando $a_j < 0$ o que, conseqüentemente, acarretará na ‘morte’ dos neurônios que apresentarem gradiente igual a zero, devido a impossibilidade de reajuste de peso. A Figura 3.13 ilustra as curvas da função de ativação e derivada da ReLU.

$$h(a_j) = \max(0, a_j) = \begin{cases} a_j, & a_j > 0 \\ 0, & a_j < 0 \end{cases} \quad (3.53)$$

$$h'(a_j) = \begin{cases} 1, & a_j > 0 \\ 0, & a_j < 0 \end{cases} \quad (3.54)$$

Figura 3. 13 – Curva da função ReLU e sua derivada



Fonte: Feng & Lu (2019)

Com as respostas dos neurônios na camada de saída da rede neural, é avaliado a acuracidade do modelo a partir da função de perda. Há diversos tipos de função de perda, dentre as mais conhecidas, destacam-se as funções de erro dos mínimos quadrados e a entropia cruzada, correspondentes às Eq. (3.55) e (3.56), respectivamente. Sendo t_n um conjunto de vetores dos valores rotulados do aprendizado, x_n um conjunto de vetores das variáveis de entrada e y_n um conjunto de vetores de saída do modelo (BISHOP; NASRABADI, 2006).

$$E(w) = \frac{1}{2} \sum_{n=1}^N \|y(x_n, w) - t_n\|^2 \quad (3.55)$$

$$E(w) = - \sum_{n=1}^N \{t_n \ln y_n + (1 - t_n) \ln(1 - y_n)\} \quad (3.56)$$

A partir da função de perda é possível otimizar o modelo com a Retropropagação. A Retropropagação aplica o gradiente descendente mediante a regra da cadeia para diferenciação e, conseqüentemente, é possível modificar os pesos. A minimização do erro é baseada na Eq. (3.57) (BRUNTON; KUTZ, 2019).

$$\frac{\partial E}{\partial w} = \frac{\partial E}{\partial y} \frac{\partial y}{\partial a} \frac{\partial a}{\partial w} \quad (3.57)$$

Sendo um processo iterativo, o peso é calculado a partir da Eq. (3.58). Onde $k + 1$ é o número da iteração aplicada, j refere-se ao j -ésimo componente do vetor w e δ é a taxa de aprendizado do modelo). A otimização começa no cálculo dos novos pesos da última camada e termina no cálculo dos pesos das primeiras camadas, reiniciando o modelo com os novos pesos e estimando novos resultados na camada de saída, reiniciando uma nova iteração da Retropropagação (BRUNTON; KUTZ, 2019).

$$w_{k+1}^j = w_k^j + \delta \frac{\partial E}{\partial w_k^j} \quad (3.58)$$

3.3. Redução de dimensionalidade

Por se tratar do aprendizado de máquinas, com as mais diversas técnicas, para lidar com diferentes abordagens, é necessário analisar quais *features*, de fato, contribuirão efetivamente para um bom modelo de aprendizado. Sendo assim, a redução de dimensionalidade surge com a proposta de representar dados de alta dimensionalidade, sem prejudicar sua estrutura, em espaços de baixa dimensionalidade, com inclusão de *outliers* e *clusters*. Com isso, é possível visualizar, graficamente, dados de alta dimensão e perdendo informações em contrapartida (OBAID *et al*, 2019; WENSKOVITCH *et al*, 2017).

De modo geral, a redução de dimensionalidade filtra as amostras de dados dentro de um conjunto de dados, a partir do mapeamento de entradas, o que reduz a complexidade do tempo do modelo, e tornando o algoritmo do aprendizado, menos robusto (REDDY *et al*, 2020). As principais classes de redução de dimensionalidade resumem-se a: classes lineares como Análise de Componentes Principais (PCA), Análise Probabilística de Componentes Principais (PPCA), Análise de Fator e *Projection Pursuit* (PP); classes não lineares como *Glimmer*, *Isomap*. Alocação Latente de *Dirichlet* (LDA), *t-Distributed Stochastic Neighbor Embedding* e *t-SNE*; também há algoritmos de redução de dimensionalidade que implementam ambas classes como a Análise de Componentes Independentes (ICA), *Multidimensional*

Scaling (MDS), *Weighted Multidimensional Scaling* (MDS) e a Seleção de *Features* (WENSKOVITCH *et al*, 2017).

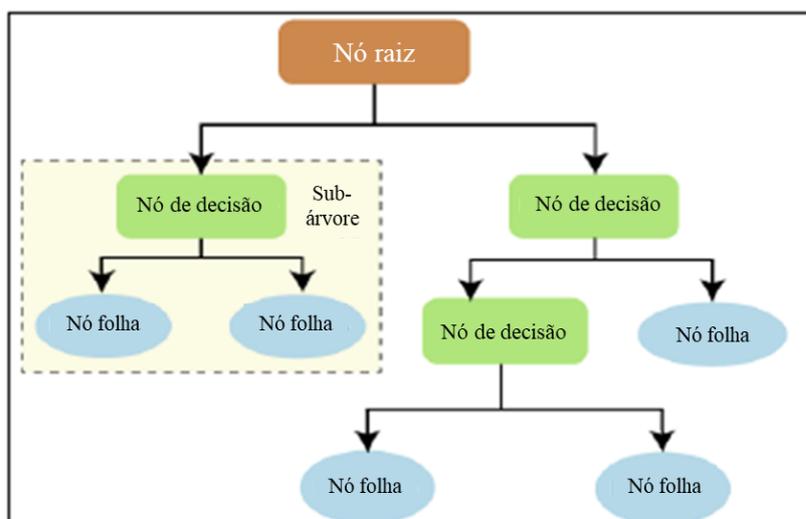
Em relação a seleção de *features*, algumas técnicas visam identificar as *features* mais significativas, em detrimento, de *features* mais ruidosas que comprometem o processo de aprendizado dos algoritmos (CHEN, 2014). Dentre algumas técnicas exemplos já aplicadas, pode-se mencionar a seleção de *features* por meio de técnicas como árvores de decisão, apresentam quais *features* contribuem mais significativamente aos modelos de aprendizado (RAVIKUMAR *et al*, 2020), bem como, a aplicação de testes estatísticos como a Análise de Variância (ANOVA), ao qual ranqueiam as *features* mais significativas, para serem priorizadas no aprendizado de máquina (PEÑA *et al*, 2018).

3.3.1. Árvore de decisão

A árvore de decisão é uma das ferramentas mais aplicadas para o aprendizado de máquina, principalmente em problemas de classificação, sendo abordada em áreas que exigem o estudo de máquinas, reconhecimento de padrões e estatística (JIJO; ABDULAZEEZ, 2021). Basicamente, é criada uma partição do espaço amostral orientado pela base de dados na forma de árvore (JANIKOW, 1998). Além disso, são consideradas metodologias não paramétricas, com resultados interpretados por meio de particionamentos denominados nós.

Inicialmente, cada particionamento define uma condição sobre a distribuição dos *dataset* por classes e ramifica as amostras de dados para outros nós, aos quais serão condicionados sucessivas vezes para outros nós, conforme os parâmetros estabelecidos no aprendizado. Os últimos nós da árvore são chamados de folhas, correspondentes as predições (DE VILLE, 2013; IZBICKI; SANTOS, 2020). A Figura 3.14 ilustra a formação de uma árvore de decisão.

Figura 3. 14 – Árvore de decisão



Fonte: Adaptado de Jijo & Abdulazeez (2021)

Uma das maiores vantagens é a fácil interpretação das previsões, a partir de recursos gráficos. Além do mais, podem trabalhar com altas escalas de dados e também podem ser otimizados a partir de métodos de combinação de mais de uma árvore, afim de melhorar a acurácia da previsão (LIU *et al*, 2020). Dentre as configurações de árvore de decisão, as mais conhecidas são: a Árvore de Classificação e Regressão (CART); Dicotomias Iterativas 4 (ID4); Detecção Automática de Interação Qui-quadrado (CHAID); e Árvore Estatística Imparcial e Eficiente (QUEST) (JIJO; ABDULAZEEZ, 2021).

A configuração da CART é obtida a partir de um método binário ao qual divide um conjunto de dados em dois grupos, a partir de uma regra estabelecida. Para isso é necessário a aplicação de um conjunto de dados de aprendizado, constituído por variáveis dependentes ou variáveis de saída, além das variáveis de entrada, responsáveis pela previsão do modelo a partir das regras criadas pela árvore (ZAMAN *et al*, 2011).

Em cada regra, é utilizada uma única variável de entrada. Para encontrar a melhor divisão para cada nó, são aplicados critérios de medida de impureza do nó. A impureza expressa a distribuição de dados pertencentes a cada classe em um nó, isto é, quanto maior o conjunto de dados pertencentes a uma única classe em detrimento das demais classe, maior a pureza do nó. Um dos principais critérios de medida de impureza é o índice de Gini, que pode ser expresso pela Eq. (3.59). Onde p é

probabilidade do conjunto de dados do nó t pertencer a h -ésima classe c (ZAMAN *et al*, 2011).

$$i(t) = 1 - \sum_{h=1}^H P(t_{c_h}) \quad (3.59)$$

Conforme o algoritmo da árvore de decisão é escolhido, é feito a seleção de um conjunto de *features* para um modelo de aprendizado. Logo, é possível visualizar no diagrama da árvore as *features* mais relevantes e eliminar as menos significativas para um modelo de classificação. As técnicas de aprendizado nas etapas posteriores apresentarão maior acuracidade com um conjunto menos robusto de informações (MURALIDHARAN *et al*, 2014; RAVIKUMAR *et al*, 2019; RAVIKUMAR *et al*, 2020).

3.3.2. Análise de Variância (ANOVA)

A ANOVA é uma excelente técnica para testes estatísticos aplicados para rejeitar a hipótese nula H_0 , ao qual é declarada como $H_0: \mu_1 = \mu_2 = \dots = \mu_K$, onde μ_i , $i = 1, 2, \dots, K$ são as médias para K diferentes classes, identificando a separabilidade entre classes. Além de comparar as médias das classes, a ANOVA as compara a partir de diferentes níveis de fatores. Como ferramenta auxiliar, é aplicado o teste-F, ao qual, é representada pela Eq. (3.60) (PEÑA *et al*, 2017; PEÑA *et al*, 2018).

$$F = \frac{\sum_{i=1}^K (\bar{Y}_i - \bar{Y})^2}{K - 1} \quad (3.60)$$

$$\frac{\sum_{i=1}^K \sum_{j=1}^{n_i} (\bar{Y}_{ij} - \bar{Y})^2}{N - K}$$

Sendo \bar{Y}_i representa a média amostral no i -ésimo grupo, n_i é a quantidade de amostras no i -ésimo grupo. Já \bar{Y} representa a média geral das amostras, \bar{Y}_{ij} é a j -ésima amostra no i -ésimo grupo, e K representa o número de grupos. Portanto, para rejeitar H_0 é necessário que F seja alto.

No entanto, só o cálculo de F não cria uma boa interpretação para o ranqueamento de *features*, sendo necessário outro valor de referência como o valor- p . O valor- p é a probabilidade de F seja alto o suficiente dentro da análise observado, de modo que a hipótese H_0 seja rejeitada. De modo geral, H_0 é rejeitada se $p \leq 0,05$.

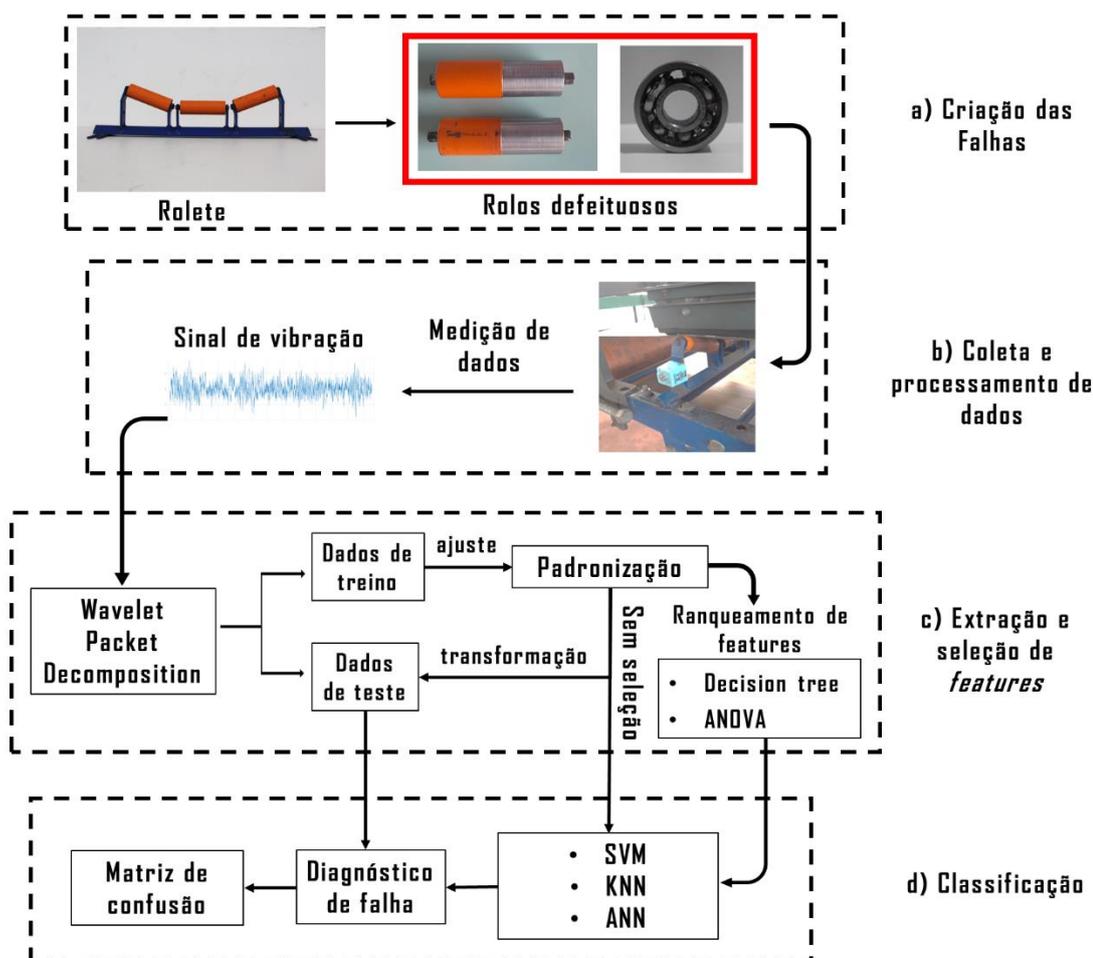
Essa técnica pode ser aplicada para o ranqueamento de features nos modelos de aprendizado para classificação de falhas em máquinas rotativas (PEÑA *et al*, 2017; PEÑA *et al*, 2018).

Baseado no conhecimento acerca da análise e processamento de sinais não-estacionários e não-lineares por meio da WPD para extração de *features*, aplicadas para a criação de uma base de dados para classificação de falhas por meio de técnicas tradicionais de aprendizado de máquinas como a SVM, kNN e MLP-ANN, aos podem ser otimizados por métodos de seleção de features como a árvore de decisão e ANOVA, foi desenvolvido este trabalho.

4. METODOLOGIA

A metodologia do trabalho foi desenvolvida a partir de abordagens experimentais e computacionais, ao qual, foi passado pelas seguintes etapas: a) Criação de falhas, b) Coleta e processamento de dados, c) Extração de features e seleção de features e d) Classificação (ilustrado pela Figura 4.1). O projeto foi desenvolvido no Parque Tecnológico Tecnolago no Laboratório de Fluidodinâmica e Particulado (FluidPar) no campus da Universidade Federal do Pará (UFPA) do município de Tucuruí-Pa.

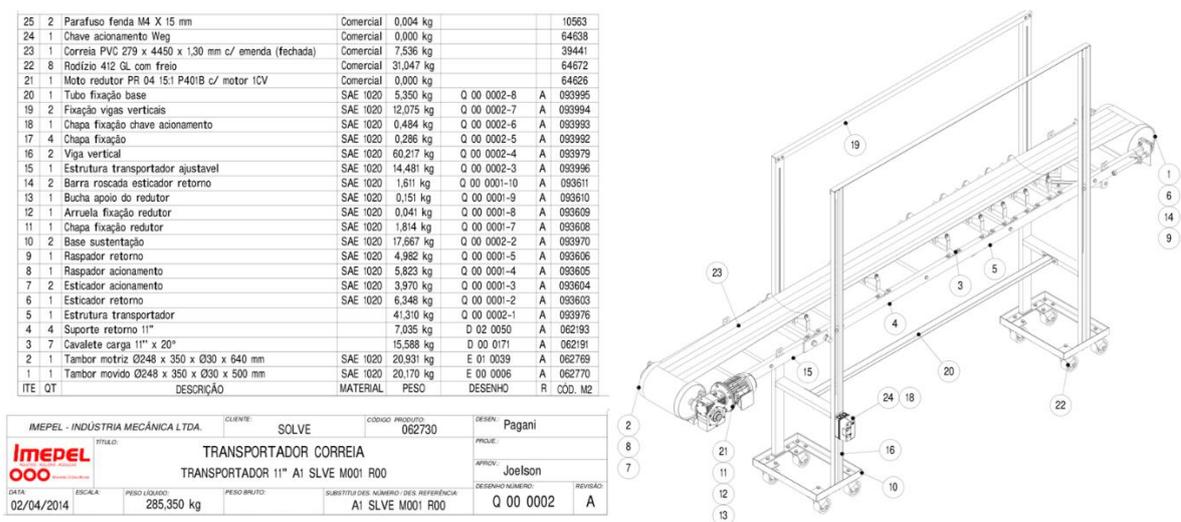
Figura 4. 1 – Etapas do trabalho a) Criação de falhas, b) Coleta e processamento de dados, c) Extração de features e seleção de features e d) Classificação



Fonte: O autor

O equipamento utilizado para aplicação da análise de vibração e estudo de classificação de falhas foi a bancada do transportador de correia de fabricante Imepel – Indústria Mecânica LTDA e de modelo de origem Transportador 11" A1 SLVE M001 R00, conforme presente na Figura 4.2. O transportador é composto pelos seguintes subconjuntos: Sistema de acionamento (Motorreductor), Sistema de transmissão (Tambores de acionamento e retorno, roletes), Estrutura e Correia. Além disso, o equipamento foi ajustado para operar com inclinação de 8° e com velocidade de 90 rpm, controlada por um inversor de frequência. As informações acerca da frequência de rotação estão representadas na Tabela 4.1 abaixo. O transportador foi acionado sem carga de material durante a experimentação, de modo análogo as bancadas apresentadas durante a revisão bibliográfica, porém, é importante ressaltar que a aplicação de testes em bancadas com carga fornece maior similaridade nos sinais de vibração laboratoriais, em relação aos sinais medidos em campo.

Figura 4. 2 – Desenho técnico da bancada de transportador de correia



Fonte: O autor

Tabela 4. 1 – Especificação técnica do sistema de acionamento e transmissão

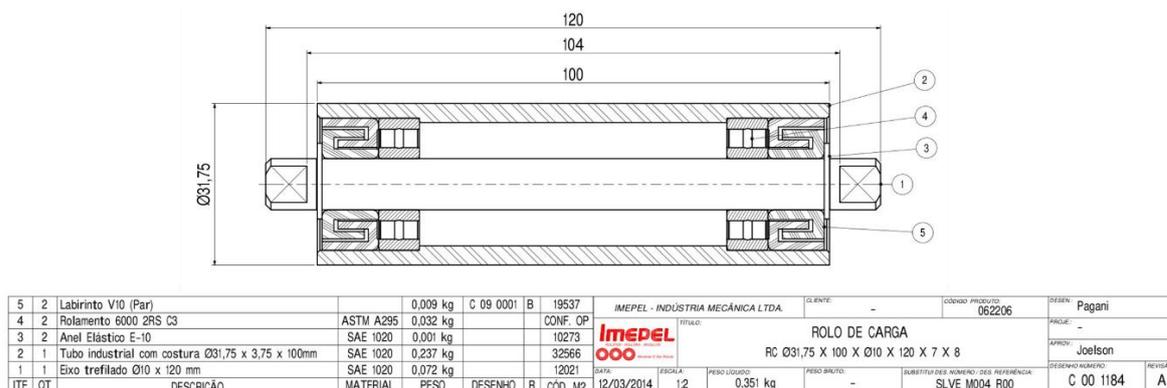
Item	Fabricante	Frequência (r/min)	Redução
Motor (1 hp)	WEG®	1350	-
Redutor	Pierine®	90	1:15
Tambor	Imepel®	90	-
Rolete	Imepel®	690	-

Fonte: O autor

4.1. Criação de falhas

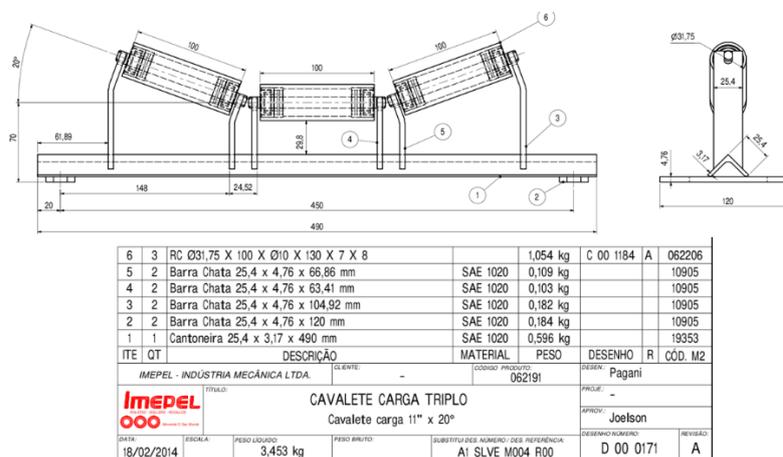
A criação de falhas nos rolos dos roletes foi baseada na simulação em dois dos principais modos de falha nos rolos: desgaste superficial na concha e defeitos nos rolamentos. O rolete utilizado para avaliação da classificação de defeitos foram o rolo e cavalete de carga, cujos desenhos são representados pelas Figura 4.3 e 4.4, respectivamente.

Figura 4.3 – Desenho técnico do rolo



Fonte: O autor

Figura 4.4 – Desenho técnico do cavalete



Fonte: O autor

4.1.1. Desgaste superficial

Inicialmente, foram implementados defeitos artificiais na superfície dos rolos com dois diferentes graus de severidades, por meio de usinagem em torno mecânico de modelo CD6241x1500. A Figura 4.5 mostra o torno utilizado no desgaste do rolo.

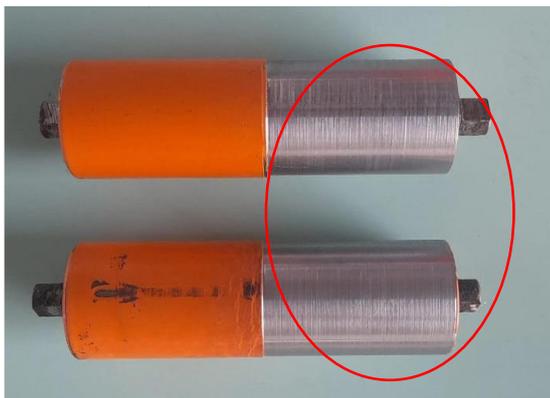
Figura 4. 5 – Torno mecânico

Fonte: O autor

Possibilitando uma margem de controle dos defeitos, a partir dos passos realizados no processo de faceamento da superfície do rolo, foram analisados modos de desgaste específicos em rolos. Destacando as causas, consequências e possíveis estratégias de prevenção, com base nas informações contidas nos artigos revisados, é possível uma compreensão mais profunda acerca do desgaste recorrente na indústria.

Por ser o elemento de contato direto com a correia e em constante movimento, o rolo está suscetível ao desgaste natural e a diferentes tipos de falhas que reduzem seu ciclo de vida. O revestimento está sujeito a rasgos e fissuras causados pela degradação devido ao uso, atrito com um ponto específico da correia quando um rolamento trava ou pela queda de materiais perfurantes sobre ele.

Ao todo, dois rolos foram usinados, um com desgaste de 0.5 mm (definido como grau 1) e outro com desgaste de 1 mm (definido como grau 2), como presente na Figura 4.6. A região usinada simula o desgaste da concha na área de contato direto com a correia. Os níveis de desgaste foram baseados nas normas brasileiras relacionadas a espessura mínima de parede do tubo dos rolos, como as NBRs 6678 (2017) e 6591 (2008) relacionadas, respectivamente a: Transportadores contínuos - Transportadores de correias - Roletes - Projeto, seleção e padronização, e Tubos de aço-carbono com solda longitudinal de seção circular, quadrada, retangular e especial para fins industriais – Especificação. Estas normas indicam uma tolerância para espessura nominal de 1 mm, aproximadamente, para laminados revestidos.

Figura 4. 6 – Rolos com desgaste artificial

Fonte: O autor

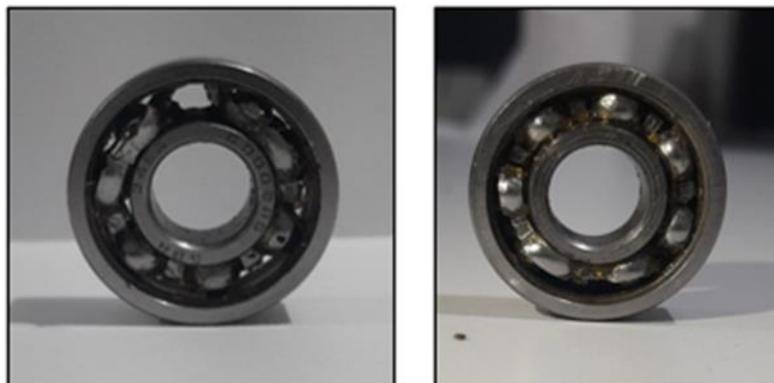
4.1.2. Defeito em rolamento

Em relação a fabricação de defeitos artificiais nos rolamentos dos rolos, foi desmontado o rolo, a partir dos seguintes passos: primeiramente, remove-se o anel elástico com o auxílio de um alicate de bico curvo, em ambas laterais do rolo, em seguida usa-se um martelo de borracha para realizar batidas leves em uma das laterais do eixo, sem comprometer a integridade do rolamento, até o labirinto da outra extremidade ficar visível e de fácil acesso para remoção do eixo, esse passo é repetido do lado oposto, e então remove os rolamentos do interior da concha.

A Figura 4.7 mostra o rolo de carga desmontado. Com a remoção dos rolamentos, foi realizado dois níveis de defeitos para simular dois diferentes graus de severidade do defeito, ao qual, foi feito um furo passante para o rompimento da gaiola por meio de uma furadeira da fabricante Bosch®, modelo GSR 1000 Smart com uma broca de 2,25 mm. No primeiro grau de severidade foi fabricado a falha em apenas um dos rolamentos do rolo (definido como grau 1), enquanto no segundo grau de severidade foi realizado um furo em cada um dos dois rolamentos (definido como grau 2). A Figura 4.8 apresenta um comparativo entre um rolamento saudável e um rolamento com defeito artificial na gaiola. A criação de defeitos na gaiola do rolamento dos roletes é uma adaptação da indução de falha já encontrado na literatura (LIU, 2016, FERNANDEZ *et al*, 2013). Segundo esses autores, o grau de severidade aumentava conforme havia o aumento de perfurações na gaiola do rolamento do rolo.

Figura 4. 7 – Rolo desmontado

Fonte: O autor

Figura 4. 8 – Comparativo entre rolamento saudável e defeituoso

Fonte: O autor

4.2. Coleta e processamento de dados

Para realizar a coleta de dados, foi necessário avaliar as condições de funcionamento do transportador de correia e instalação dos sensores de vibração. Por conta disso, analisou-se as condições de funcionamento, por meio de inspeção visual detalhada do transportador de correias. Verificando a integridade dos componentes, como: motor elétrico, redutor, correia, tambores e roletes, além da ausência de desgastes excessivos em componentes rolantes. Conseqüentemente foi realizado ajustes como o abastecimento de óleo no redutor para evitar superaquecimento no

sistema de acionamento, alinhamento de cavaletes e limpeza de correia e raspadores para evitar obstruções.

Além do transportador montado conforme o projeto original, conforme a Figura 6.1 foi instalado um chute de carregamento de material, como é possível visualizar na Figura 6.8. Para a instalação dos sensores de vibração, foi escolhida a região do primeiro cavalete abaixo do chute, por proporcionar maior tensão sobre a correia e, conseqüentemente, maior carga sobre os roletes. Sendo assim, foram instalados dois sensores nas laterais do cavalete, conforme destacado na Figura 4.9.

Figura 4. 9 – Bancada do transportador de correia *as built*



Fonte: O autor

Os sensores selecionados para o monitoramento dos roletes foram os sensores da fabricante @Dynamox S.A, dos modelos *DynaLogger HF (High Frequency)* durante os períodos iniciais da coleta de dados e sendo substituído, posteriormente, por sensores HF+. Ambos modelos são *data loggers Bluetooth* com sensores sem-fio de temperatura e sensores triaxiais de vibração para monitoramento e análise espectral da máquina e, cuja especificações técnicas estão disponíveis nas Tabela 4.2 e 4.3. Outro ponto a ser observado é que não houve variabilidade dos rolos medidos nas diferentes condições de saúde, o que pode prejudicar na generalização dos modelos de diagnósticos. Por isso a importância da medição do cavalete com sensores em dois pontos distintos e dois modelos distintos também.

Tabela 4. 2 – Especificações técnicas modelo HF

Informações gerais	
Modelo	HF
Dimensão	39 mm x 39 mm x 35 mm
Peso	67 g
Material	LEXAN™
Cor	Azul
Fixação	Parafusado ou colado
Memória	7 116 amostras
Bluetooth	BLE 5.0 / 2 400 – 2 483,5 MHz
Tensão	3,0 V
Autonomia	3 a 5 anos (dependendo do uso)
Alcance	60 m (campo aberto)
Comunicação com App	Android e iOS
Temperatura	
Temperatura de operação	-10 °C ≤ T ≤ 84 °C
Resolução de temperatura	0,01 °C (-10 °C a 80 °C)
Vibração	
Faixa de amplitude	± 16 g
Taxa de amostragem	12,8 kHz
Linhas de resolução (LOR)	16.384 (uniaxial) e 4.096 (triaxial)
Frequência máxima	800 Hz a 6400 Hz
Tempo máx, de aquisição	20,48 s (uniaxial) e 5,12 s (triaxial)

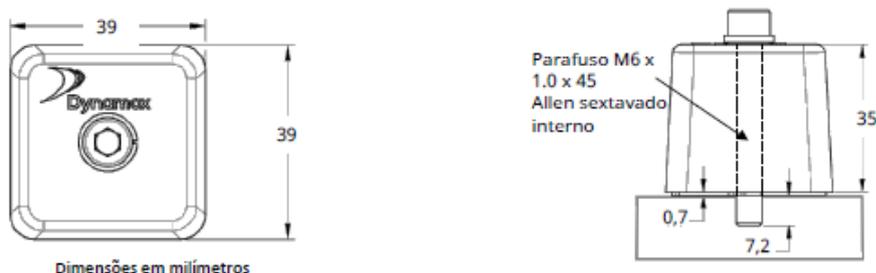
Tabela 4. 3 – Especificações técnicas modelo HF+

Informações gerais	
Modelo	HF+
Dimensão	39 mm x 39 mm x 35 mm
Peso	67 g
Material	LEXAN™
Cor	Azul
Fixação	Parafusado ou colado
Memória	51 200 amostras (configurável)
Bluetooth	BLE 5.3 / 2400 – 2483.5 MHz
Tensão	3 V
Autonomia	3 a 5 anos

Alcance	100 m
Comunicação com App	Android e iOS
Temperatura	
Temperatura de operação	$-20^{\circ}\text{C} \leq T \leq 84^{\circ}\text{C}$
Resolução de temperatura	0,01°C
Vibração	
Acelerômetro	MEMS triaxial
Bandas de frequência	2 Hz a 13 kHz (configurável)
Resposta em frequência (± 3 dB)	7,6 kHz
Faixa de amplitude	Até ± 16 g
Taxa de amostragem	Até 26 kHz
Resolução mín. em frequência	0,006 Hz (8 bits) e 0,012 Hz (16 bits)
Resolução mín. em amplitude	16 mg (8 bits) e 61 μg (16 bits)
Linhas de resolução (LOR)	98 304 (uniaxial) e 32 768 (triaxial)
Frequência máxima	571 Hz a 13 kHz (configurável)
Tempo máx. de aquisição	172,2 s (uniaxial) e 57,3 s (triaxial)

Sobre a instalação dos sensores nos cavaletes, decidiu-se que o tipo de fixação seria por aparafusamento, recomendada para instalações permanentes. Logo, foi utilizado um parafuso M6 com comprimento que permita um mínimo de 7,2 mm de profundidade em relação a superfície perfurada, conforme visto pela Figura 4.10.

Figura 4. 10 – Desenho técnico dos *data loggers* HF e HF+



Fonte: O autor

A configuração da coleta de sinal dos sensores foi realizada na plataforma *DynaPredict Web*. *DynaPredict Web* é um ambiente em nuvem, cujos sinais registrados pelos *data loggers* são coletados por *app mobile* e são centralizados para análise, interpretação e prognóstico de falhas. A partir da plataforma, os espectrais

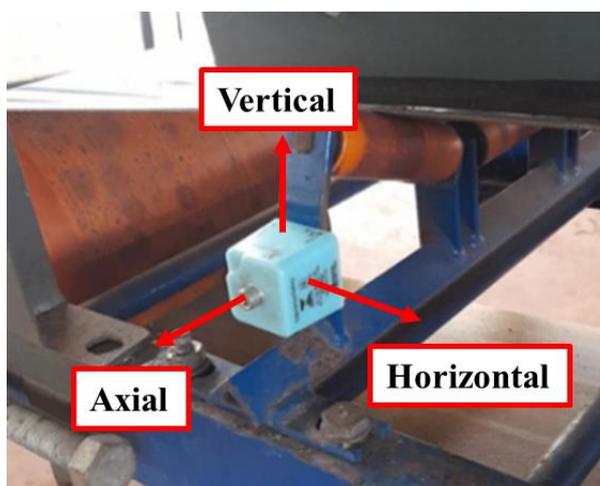
padrões do sinal são configurados, conforme a Tabela 4.4. A figura 4.11 apresenta a direção dos eixos de vibração no cavalete

Tabela 4. 4 – Configuração padrão da espectral

Parâmetros	Status
Faixa dinâmica	8g
Eixos	Vertical, Horizontal e Axial
Frequência máxima	6572,5 Hz
Duração da coleta	0,311373 s
Taxa de amostragem	13151,4 Hz

Fonte: O autor

Figura 4. 11 – Direção dos eixos de vibração



Fonte: O autor

Após a fabricação de falhas artificiais dos rolos e a configuração das espectrais para o registro da forma de onda dos sinais, foi realizado a coleta de dados de aceleração de vibração onde a Tabela 4.5 expressa a quantidade de dados coletados dos sensores fixados em cada lado do cavalete, além de indicar a posição do rolo defeituoso, em relação ao rolete. Os rolos defeituosos foram instalados somente nas posições laterais e inclinadas do cavalete (excluindo a posição centralizada, com o rolo na horizontal), cuja referência será a posição em relação ao motor de acionamento do transportador de correia. Portanto, LA representará o Lado do Acionamento, enquanto LOA representará o Lado Oposto ao Acionamento.

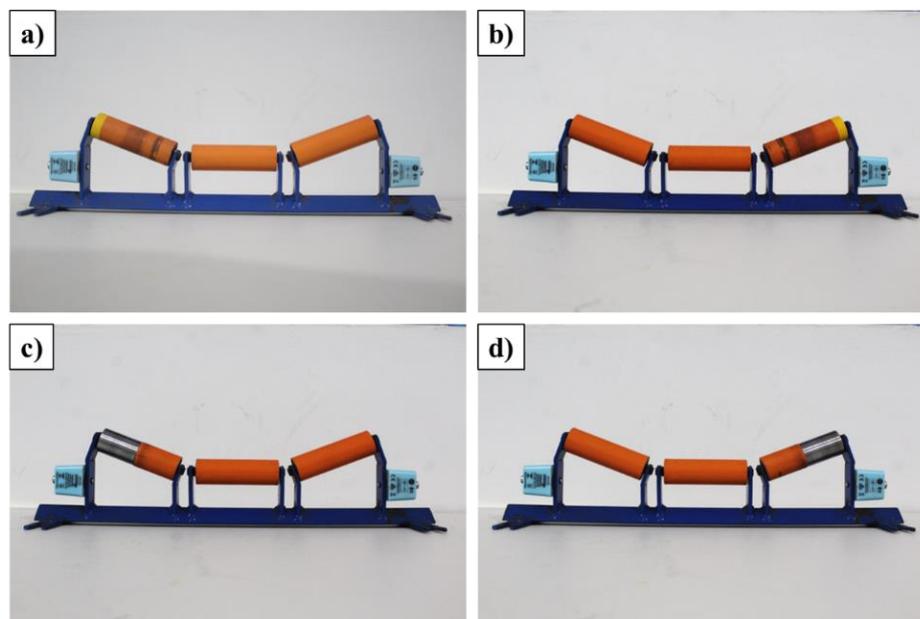
Tabela 4. 5 – Perfil de coleta de sinais

Condição do rolo	Posição do acelerômetro		Posição do rolo	
	LA	LOA	LA	LOA
	Número de coletas da vibração			
Sem defeitos	416	416	X	X
Desgaste superficial (grau 1)	416	416	X	
Desgaste superficial (grau 1)	416	416		X
Desgaste superficial (grau 2)	416	416	X	
Desgaste superficial (grau 2)	416	416		X
Defeito em rolamento (grau 1)	416	416	X	
Defeito em rolamento (grau 1)	416	416		X
Defeito em rolamento (grau 2)	416	416	X	
Defeito em rolamento (grau 2)	416	416		X

Fonte: O autor

A Figura 4.12 ilustra as quatro configurações de posicionamento dos rolos defeituosos em relação aos sensores instalados. Dentro de cada configuração serão coletados os sinais dos rolos com os dois diferentes modos de falhas, cada qual com dois graus de severidade.

Figura 4. 12 – Configuração do rolo defeituoso instalado nas posições: Rolo com rolamento defeituoso no LA (a) e LOA (b) e rolo com desgaste superficial no LA (c) e LOA (d)



Fonte: O autor

Todos os sinais coletados são processados e sobem para a nuvem, ao qual, são armazenados em um sistema de *bucket* criado pelo *Google Cloud*. Esses sinais armazenados são organizados em arquivos de formato *json* e distribuídos em pastas conforme a data e local de coleta. Destarte, os arquivos foram baixados da nuvem para a criação de um banco de dados para o tratamento dos sinais. Durante o tratamento de dados, foi trabalhado cada eixo de vibração da máquina, isoladamente, para averiguação da direção de vibração mais eficaz para o diagnóstico de falhas.

4.3. Extração e seleção de features

O banco de dados foi aquisitado e tratado mediante a criação de um algoritmo para *Machine Learning* com linguagem em Python® desenvolvido no Google Colaboratory®. Dentre as principais bibliotecas para a execução do algoritmo, destacam-se: Scikit-Learn (Aplicação de técnicas de *Machine Learning*), NumPy (Funções de computação numérica), Pandas (Aquisição e análise de banco de dados), PyWavelets (Utilização de funções wavelet) e Matplotlib (Recursos gráficos visuais).

4.3.1. Bandas de energia wavelet normalizada

A partir da criação de um banco de dados de aceleração de vibração no domínio do tempo, foi realizado o processo de decomposição dos sinais em bandas de energia por meio do WPD, ao qual, foi configuração 15 níveis de decomposição, possibilitando a formação de 16 bandas de frequência para cada amostra. Inicialmente, com a seleção do membro 'db 8', extraída da família Daubechies, como função *wavelet* selecionada foi feito o cálculo dos coeficientes e da energia wavelet para cada banda, representado, respectivamente, pelas Eq. (3.25) e (3.26) já apresentadas.

Ademais, as bandas de energia *wavelet* foram pré-processadas de modo que as amostras se ajustassem em novos valores z dentro de uma faixa normalizada, por meio da Eq. (4.1). Sendo x o valor inicial da amostra, u a média amostral e s o desvio padrão. Com a energia wavelet normalizada calculada, foi extraído um novo banco de dados com cada banda de energia representando uma *feature* da amostra.

$$z = \frac{x - u}{s} \quad (4.1)$$

Outrossim, para cada amostra foi identificado o estado do sinal como 'condição normal' ou 'condição defeituosa' para ambos modos de falha e graus de severidade. Além disso, o algoritmo também classifica o estado de rolos defeituosos localizados na posição oposta do cavalete em relação a posição do acelerômetro, viabilizando a classificação de falha dos rolos nas posições laterais com apenas um acelerômetro por cavalete. Sendo assim, foi aplicado nove rótulos diferentes para o aprendizado supervisionado de classificação de falhas, como apresentado na Tabela 4.6.

Tabela 4. 6 – Rótulo da condição dos rolos

Condição do rolo	Rótulo de classificação
Sem defeitos	0
Desgaste superficial (grau 1)	1
Desgaste superficial (grau 1) – Lado oposto	2
Desgaste superficial (grau 2)	4
Desgaste superficial (grau 2) – Lado oposto	3

Defeito em rolamento (grau 1)	5
Defeito em rolamento (grau 1) – Lado oposto	6
Defeito em rolamento (grau 2)	7
Defeito em rolamento (grau 2) – Lado oposto	8

Fonte: O autor

4.3.2. ANOVA

Após a formatação adequada da nova base de dados com a energia wavelet normalizada, as *features* foram tratadas por meio do processo de seleção de *features* por meio da ANOVA, para comparar médias entre os nove diferentes rótulos dos sinais dos rolos. Visando determinar diferenças, estatisticamente, significativas foram calculados o F-estatístico para compreender a variabilidade dos dados dentro da classe e entre classes, bem como o valor- p ao qual indica a probabilidade de uma classe ser separável. Como ferramenta auxiliar, foram aplicados *boxplots* nas três *features* que apresentaram maior separabilidade entre classes (menores valor- p) para visualizar padrões e similaridades para cada eixo de direção.

4.3.3. Árvore de decisão

Além da pré-análise com Análise de Variância, as *features* também foram tratadas por meio da seleção de *features* com a árvore de decisão. Inicialmente, separou-se 75% dos dados, de cada classe, para evitar o *overfitting* dos modelos durante a classificação de falhas. Então foi aplicado o algoritmo de aplicação da árvore de decisão a partir da base de dados com as *features* obtidas. A configuração dos parâmetros escolhidos para o algoritmo da árvore de decisão é exibida pela Tabela 4.7.

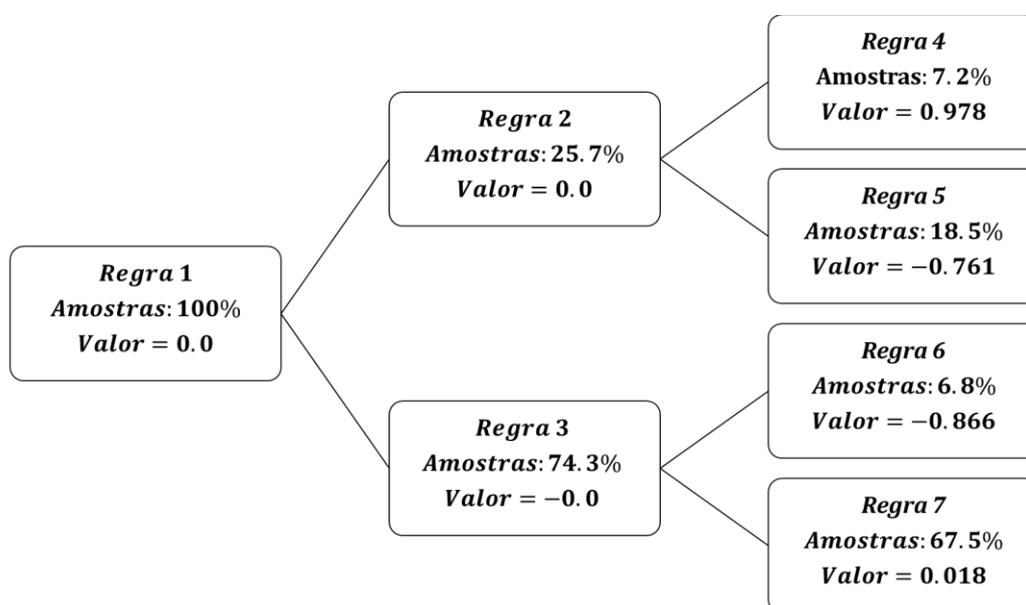
Tabela 4. 7 – Parâmetros para a árvore de decisão

Parâmetro	Status
Critério (Função de qualidade de divisão)	Gini
Máxima profundidade	3
Número mínimo de amostras para divisão de um nó interno	2
Número mínimo de amostras para estar contido em uma folha	1

Fonte: O autor

Com a exibição da árvore de decisão, como ilustrado pela Figura 4.13, é possível visualizar as regras criadas pelo modelo, a partir das *features* criadas pelas bandas de energia, para a classificação das condições do rolo. Logo, as *features* mais significativas escolhidas nas regras pelo modelo foram selecionadas para a aplicação das técnicas que criam os modelos de aprendizado de máquinas para o diagnóstico de falhas.

Figura 4. 13 – Modelo de árvore de decisão para seleção de *features*



Fonte: O autor

4.4. Classificação

Após a extração e pré-análise de *features*, foram selecionadas as seis *features* mais significativas para comparativo dos modelos de diagnóstico com e sem a seleção de *features*. Antes disso, foi aplicado o balanceamento de dados de modo, a equilibrar o quantitativo de dados por classe para o treinamento de dados, com o objetivo de otimizar o modelo de classificação a ser usado. Por conseguinte, as amostras foram divididas em uma proporção de 75%/25% para dados de treino e de teste, para cada classe, aleatoriamente. Sendo assim, os 25% dos dados restantes serão aplicados para o conjunto de teste para verificar a capacidade de generalização do modelo.

Foram aplicadas três diferentes técnicas de aprendizado: SVM, kNN e ANN-MLP. Para cada técnica foram criados três algoritmos de classificação para cada eixo de direção da vibração captada pelo acelerômetro, com e sem a seleção de *features*. Além disso, foi realizado um comparativo para cada técnica mediante a análise de acuracidade a partir da variação de determinados parâmetros, conforme exibido na Tabela 4.8.

Tabela 4. 8 – Configuração de parâmetros dos algoritmos de aprendizado

SVM	
Parâmetro C	1 – 199
Kernel	Radius Basis Function (RBF)
kNN	
Número de vizinhos (n)	1 – 25
Peso	Uniforme
ANN – MLP	
Máximas iterações (i)	1 – 970
Número de camadas ocultas	25
Função de ativação	ReLU
Solver	Adam

Fonte: O autor

Para avaliação do diagnóstico de falha da base de dados, foi analisada a acurácia total dos modelos classificadores bem como alguns indicadores como: o falso negativo (correspondente a taxa de erro na detecção de sinais defeituosos), o falso positivo (correspondente a taxa de erro na detecção de sinais saudáveis) e taxa de acerto do posicionamento do rolo defeituoso no cavalete. Além de apresentar as curvas da acurácia com a variação do parâmetro C, também foram apresentadas as matrizes de confusão, para maior entendimento acerca da predição das classes.

5. RESULTADOS

O tempo de resposta computacional para a importação de sinais, extração de *features* e criação da base de dados foi de 24 minutos. Esses dados são salvos e não necessitam de execução periódica, mas somente na necessidade de renovação dos dados de treino. As etapas de seleção de *features*, criação dos modelos de classificação e plotagem das matrizes de confusão obtiveram um tempo de resposta de 10 segundos, o que torna o diagnóstico de falhas em roletes ágil e prático.

5.1. Seleção de *features*

A partir da criação de três bases de dados (cada base representando um eixo de direção do sinal de vibração coletado) com dezesseis *features*, cada qual representa uma energia de banda *wavelet* normalizada, e nove rótulos de classes diferentes representando os estados dos rolos laterais do rolete do transportador de correia, foi realizado a seleção de *features* por meio da ANOVA e árvore de decisão. Sendo assim, foi realizado um comparativo entre as duas técnicas para avaliar a convergência das *features* mais significativas.

5.1.1. ANOVA

As Tabelas 5.1, 5.2 e 5.3 apresentam a análise de variância das energias de banda *wavelet* normalizadas nos eixos de direção vertical; horizontal e axial; respectivamente. Nas tabelas foram apresentados o F estatístico e o valor *p* de cada *feature*, ao qual as energias de banda foram ordenadas em graus decrescentes de significância, segundo a ANOVA, de acordo com o valor-*p*. cuja as *features* com maior importância apresentam os menores valores.

Tabela 5. 1 – Análise de Variância (Vertical)

<i>Feature</i>	F estatístico	Valor- <i>p</i>
E1	5090,732871	0
E13	349,3078252	0
E11	219,7028143	1,70E-289
E12	130,678986	1,19E-186
E14	122,6946892	1,58E-176
E9	40,79249918	4,99E-62
E10	34,52651898	2,14E-52
E15	26,64856048	4,06E-40
E16	14,11578752	2,58E-20
E5	10,83869223	3,82E-15
E2	9,404818497	6,73E-13
E6	9,258081543	1,14E-12
E8	9,182964485	1,49E-12
E7	9,058377345	2,34E-12
E4	8,958285018	3,34E-12
E3	8,914194848	3,92E-12

Fonte: O autor

De modo geral, em todos os casos, as features apresentaram valor-*p* abaixo de 0.01, ao qual podem ser caracterizados como ‘altamente significante’ devido ao alto grau de diferenciabilidade entre os grupos de dados entre classes. Em relação às *features* do eixo vertical, as energias de banda ‘E1’ e ‘E13’ apresentaram valor-*p* igual à zero.

Tabela 5. 2 – Análise de Variância (Horizontal)

<i>Feature</i>	F Estatístico	Valor- <i>p</i>
E1	2727,062	0
E13	328,5547	0
E14	404,9886	0
E12	194,8019	1,17E-262
E16	169,8198	3,39E-234
E5	52,23314	2,58E-79
E15	51,06756	1,42E-77
E6	43,61359	2,42E-66
E4	41,81165	1,34E-63
E7	41,5644	3,19E-63
E8	41,28837	8,41E-63
E2	32,54955	2,42E-49
E3	32,21839	7,93E-49
E10	15,13817	6,16E-22
E9	13,01839	1,40E-18
E11	11,31686	6,75E-16

Fonte: O autor

No eixo horizontal, foi identificado três *features* foram identificados com valor-*p* igual à zero nas energias de banda ‘E1’, ‘E13’ e ‘E14’. No eixo axial, duas energias de banda (‘E14’ e ‘E16’) apresentaram valor-*p* igual à zero. Logo, verificou-se a similaridade entre as energias ‘E1’ e ‘E13’ entre as *features* com maior grau de significância nos eixos vertical e horizontal e a energia ‘E14’ com maior significância nos eixos horizontal e axial.

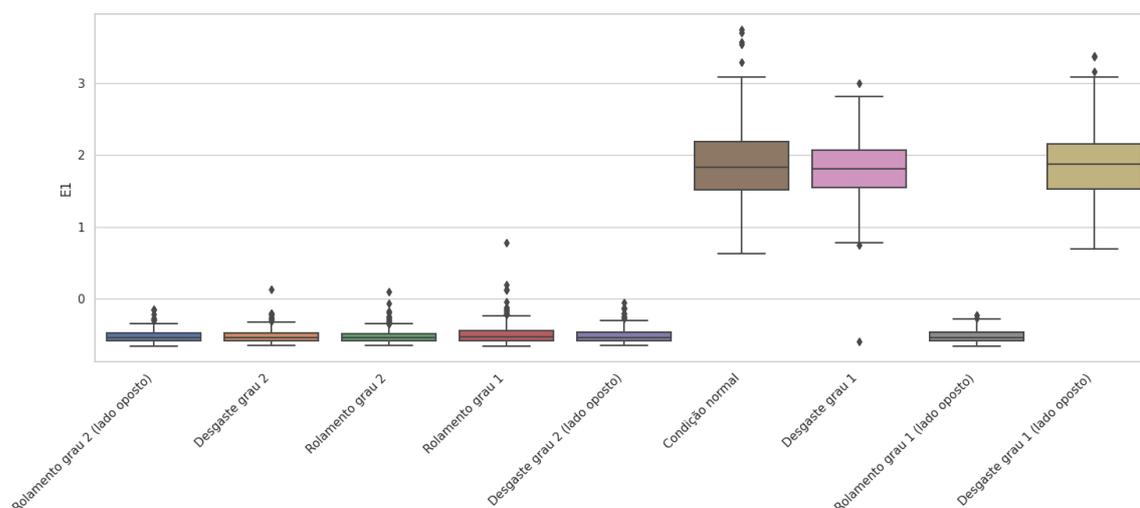
Tabela 5. 3 – Análise de Variância (Axial)

Feature	F Estatístico	Valor-p
E14	281,2792	0
E16	849,7101	0
E13	233,6291	2,85E-304
E11	111,6891	2,42E-162
E15	61,57778	4,38E-93
E12	39,56866	3,65E-60
E9	36,97626	3,50E-56
E10	35,27867	1,45E-53
E5	34,07923	1,04E-51
E6	26,3021	1,41E-39
E7	25,54433	2,19E-38
E4	24,73145	4,16E-37
E8	23,76072	1,40E-35
E2	23,0125	2,13E-34
E3	22,9613	2,56E-34
E1	10,95404	2,52E-15

Fonte: O autor

Mediante a ANOVA nos três eixos, foi identificado as três *features* em cada eixo de direção com os menores valores-*p* para serem analisadas, mais detalhadamente, a partir da representação gráfica de *boxplots* expostas nos eixos vertical, horizontal e axial. Para cada *feature* selecionada, é possível analisar a separabilidade entre os diferentes rótulos de classes e identificar possíveis dificuldades de classificar os estados dos sinais de vibração dos rolos. A Figura 5.1 apresenta os *boxplots* de classes na energia de banda 'E1' no eixo vertical.

Figura 5. 1 – *Boxplots* de classes na energia de banda ‘E1’ no eixo vertical



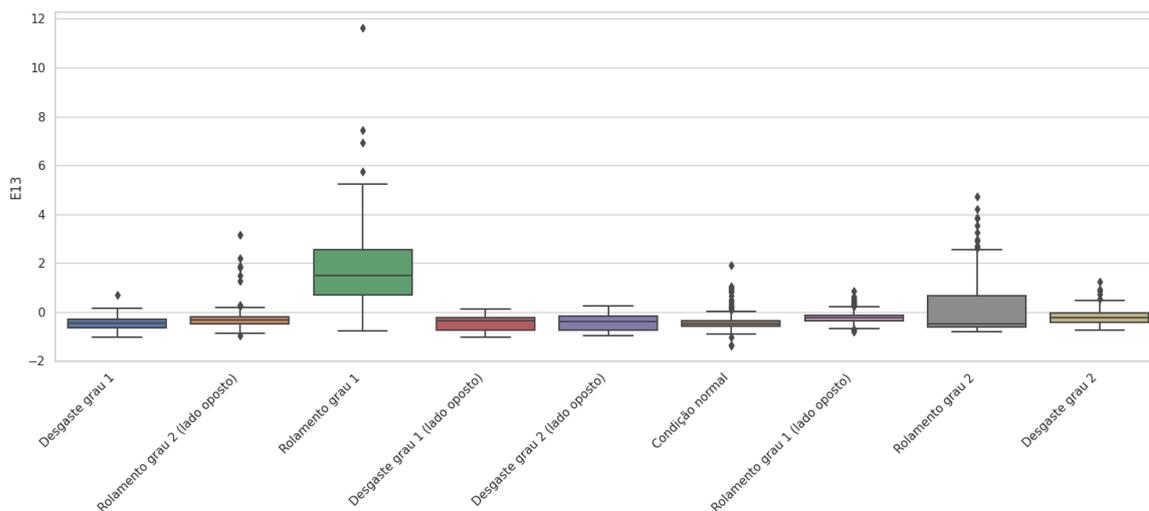
Fonte: O autor

Sobre a Figura 5.1, percebe-se uma ótima separabilidade do conjunto de dados com ‘condição normal’ em relação a seis de oito estados defeituosos. No entanto, esse estado apresentou similaridade em relação aos conjuntos de dados de rolos que apresentam desgaste superficial de grau 1, tanto no lado quando no lado oposto ao posicionamento do acelerômetro no cavalete.

Apesar disso, houve boa diferenciabilidade em rolos que continham desgaste mais acentuado na concha (grau 2) em relação aos rolos saudáveis, o que implica na dificuldade de detecção de desgaste na concha dos rolos, em graus mais prematuros. Além disso, defeitos em rolamentos apresentaram boa diferenciabilidade para os dois graus de severidade em relação a ‘condição normal’, o que agiliza o diagnóstico de falhas e evita detecções mais tardias de falhas.

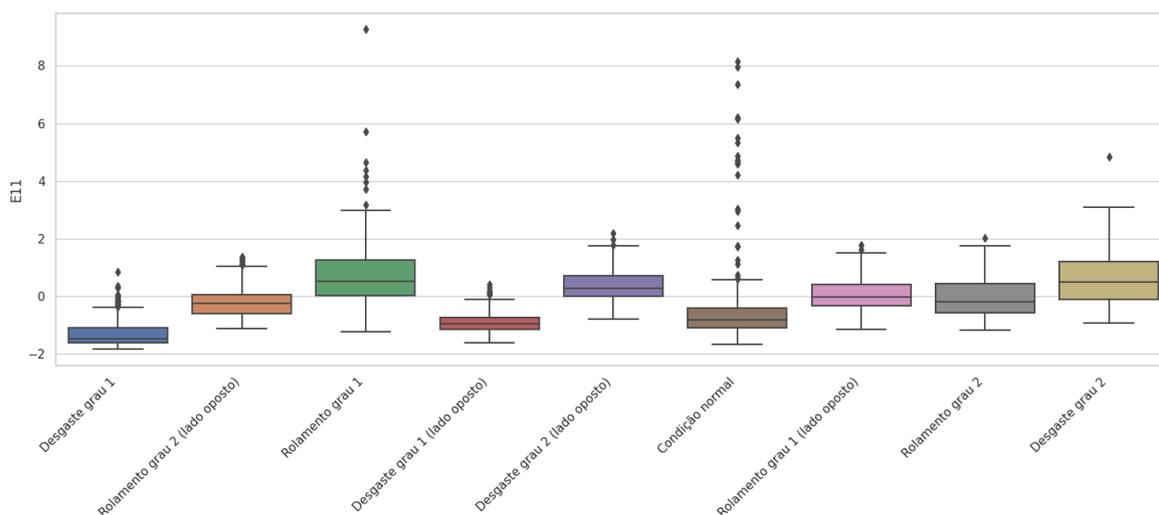
Todavia, apesar da energia de banda ‘E1’ ser um ótimo indicativo para detecção de estados binários de rolos em condições ‘normais’ ou ‘defeituosas’, é notável uma maior dificuldade em classificar os diferentes modos de falhas, graus de severidade, bem como, a posição dos rolos defeituosos em relação aos sensores. Logo, infere-se a necessidade da aplicação de mais *features* para um diagnóstico de falha mais completo em conjunto com a energia ‘E1’. As Figuras 5.2 e 5.3 apresentam os *boxplots* de classes na energia de banda ‘E13’ e ‘E11’ no eixo vertical, respectivamente.

Figura 5. 2 – *Boxplots* de classes na energia de banda ‘E13’ no eixo vertical



Fonte: O autor

Figura 5.3 – Boxplots de classes na energia de banda ‘E11’ no eixo vertical



Fonte: O autor

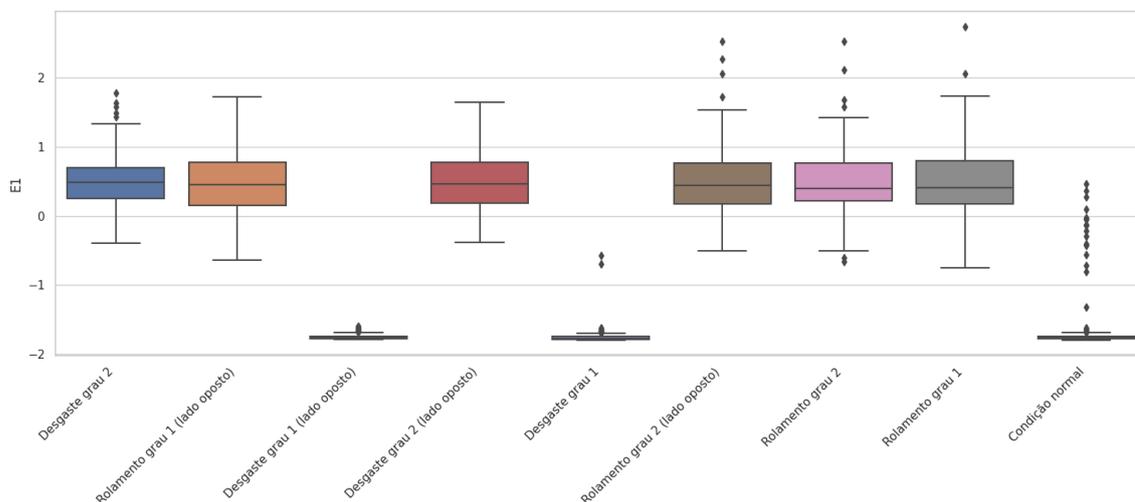
Na Figura 5.2 nota-se uma melhor diferenciação do estado de defeito em ‘rolamento grau 1’ dos demais estados do rolo. Já o conjunto de dados em ‘condição normal’ apresentou menor separabilidade em relação aos estados defeituosos, o que implica que a feature não é indicável para classificações binárias. Por fim, a *feature* ‘E13’ também não apresentou boa separabilidade para classes de mesmo modo de falha e com diferentes graus de severidade.

Sobre a Figura 5.3 as *boxplots* apresentaram comportamento análogo a Figura 7.1, ao qual expressa a similaridade entre os estados de ‘condição normal’ e os estados de ‘desgaste grau 1’, em contraste com melhor separabilidade em relação aos demais estados defeituosos dos conjuntos de dados. Todavia foi observado

melhor diferenciação entre estado com mesmo modo de falha e diferentes graus de severidade, o que o torna um ótimo indicador para o monitoramento da evolução dos defeitos.

Em relação ao eixo horizontal as três features com maior grau de significância foram as energias de banda ‘E1’, ‘E13’ e ‘E14’ conforme apresentadas, anteriormente, com os menores valores- p pela ANOVA. A Figura 5.4 apresenta os *boxplots* de classes na energia de banda ‘E1’ no eixo horizontal. Assim como na análise dos conjuntos de dados verticais, na energia de banda ‘E1’ é possível identificar excelente separabilidade, com exceção dos estados de ‘desgaste grau 1’, do estado dos sinais em ‘condição normal’ em relação aos demais estados defeituosos.

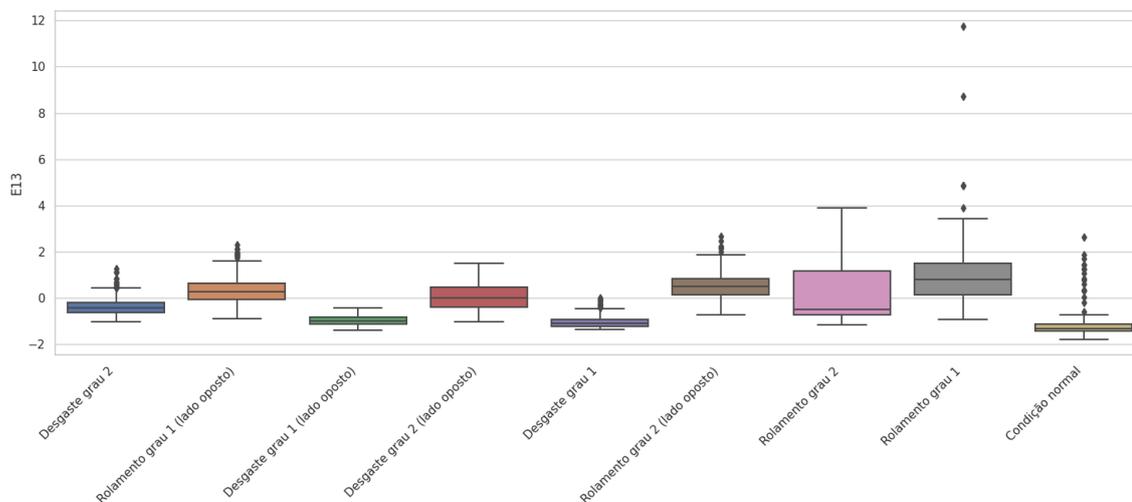
Figura 5. 4 – Boxplots de classes na energia de banda ‘E1’ no eixo horizontal



Fonte: O autor

A Figura 5.5 retrata os *boxplots* de classes na energia de banda ‘E13’ no eixo horizontal. É possível notar para esse caso, que há maior diferenciação entre os estados com diferentes modos de falhas e graus de severidade. Apesar disso, há maior dificuldade na identificação do lado que está posicionado o rolo defeituoso. Por fim, a Figura 8.6 retrata os *boxplots* de classes na energia de banda ‘E14’ no eixo horizontal.

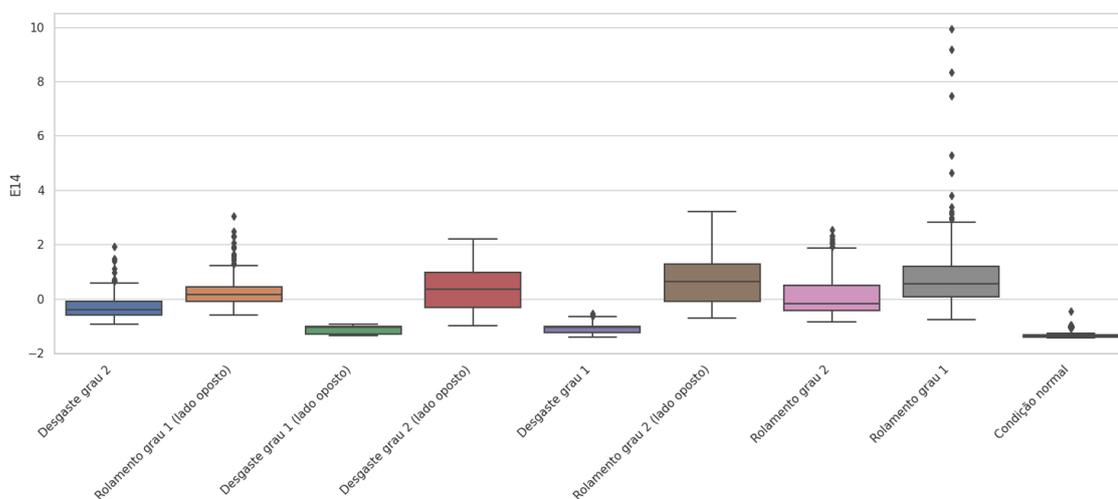
Figura 5. 5 – Boxplots de classes na energia de banda ‘E13’ no eixo horizontal



Fonte: O autor

Sobre a Figura 5.6, observa-se os conjuntos de sinais em ‘condição normal’ apresentam uma menor dispersão de dados em relação às demais condições já apresentadas. Além disso, apresentou boa separabilidade para classificação binária, sendo um excelente indicador para detecção da presença de falhas, incluindo desgastes em fase prematura, sendo um diferencial da *feature* e complementar em relação as demais já apresentadas para um diagnóstico mais completo de falhas.

Figura 5. 6 – Boxplots de classes na energia de banda ‘E14’ no eixo horizontal

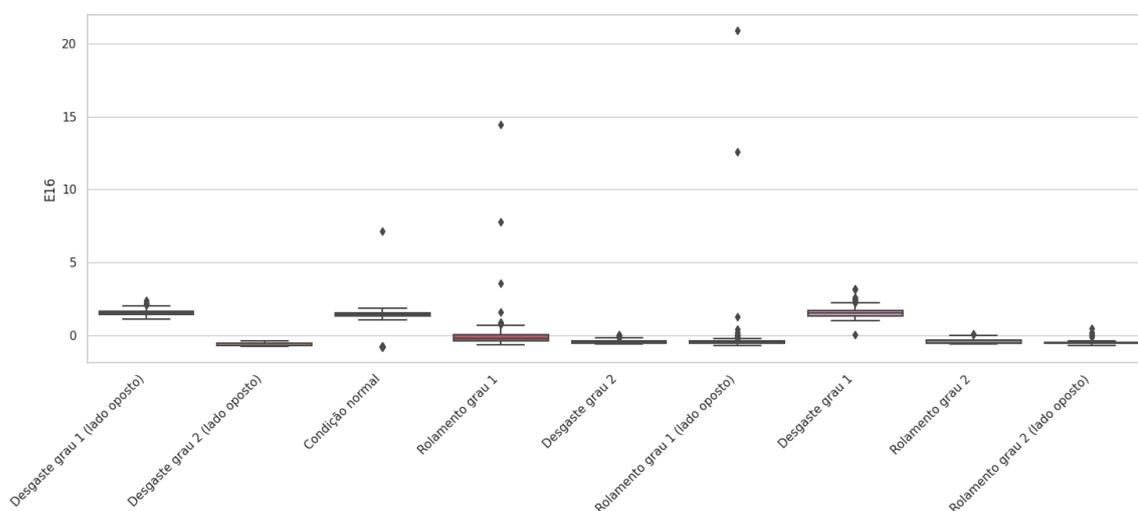


Fonte: O autor

Na direção axial, percebe-se diferenças significativa sobre as *features* selecionadas em relação as demais direções. A *feature* com maior grau de significância na direção axial foi a energia de banda ‘E16’ ao contrário das outras

direções, cujo *boxplots* são ilustrados pela Figura 5.7. Na Figura 5.7 percebe-se que os conjuntos de dados em todos os rótulos de classes apresentam baixa dispersão, o que facilita na identificação de *outliers*. Apesar disso, há baixa separabilidade entre classes, com exceção dos conjuntos com rótulos em ‘condição normal’ e com desgaste grau 1, assim como o comportamento das principal *feature* das direções vertical e horizontal. Já a Figura 5.8 trata-se dos *boxplots* de classes na energia de banda ‘E13’ no eixo axial.

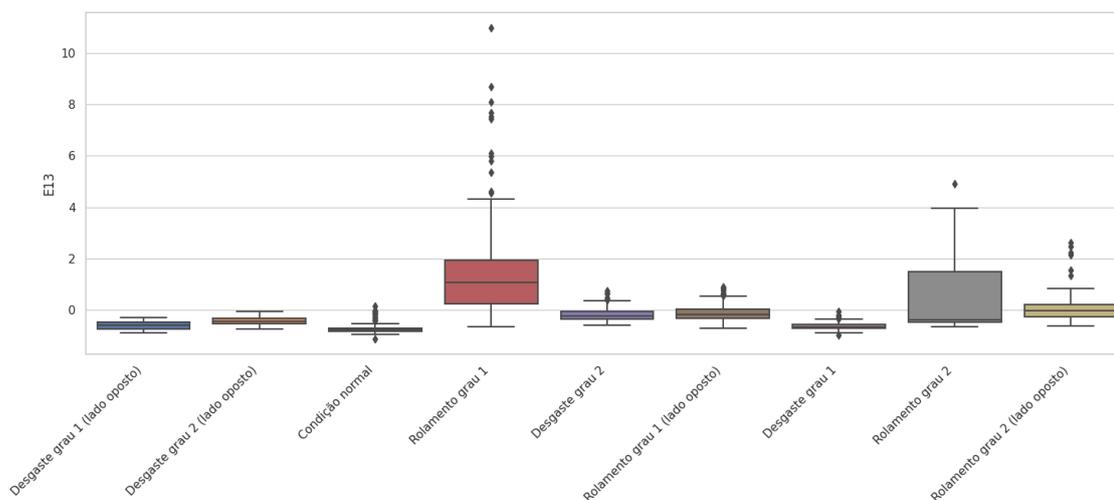
Figura 5.7 – *Boxplots* de classes na energia de banda ‘E16’ no eixo axial



Fonte: O autor

Na Figura 5.8 é notável os maiores níveis de dispersão entre os conjuntos dos dados que apresentam defeitos em rolamentos nos graus 1 e 2, e mais enfaticamente nos rolos que estavam presentes no mesmo lado do cavalete que o acelerômetro estava instalado. Portanto, infere-se que a energia de banda ‘E13’ na direção axial, apresenta maior dificuldade em encontrar os padrões de defeitos em rolamentos de rolos. A Figura 5.9 apresenta os *boxplots* de classes na energia de banda ‘E14’ no eixo axial.

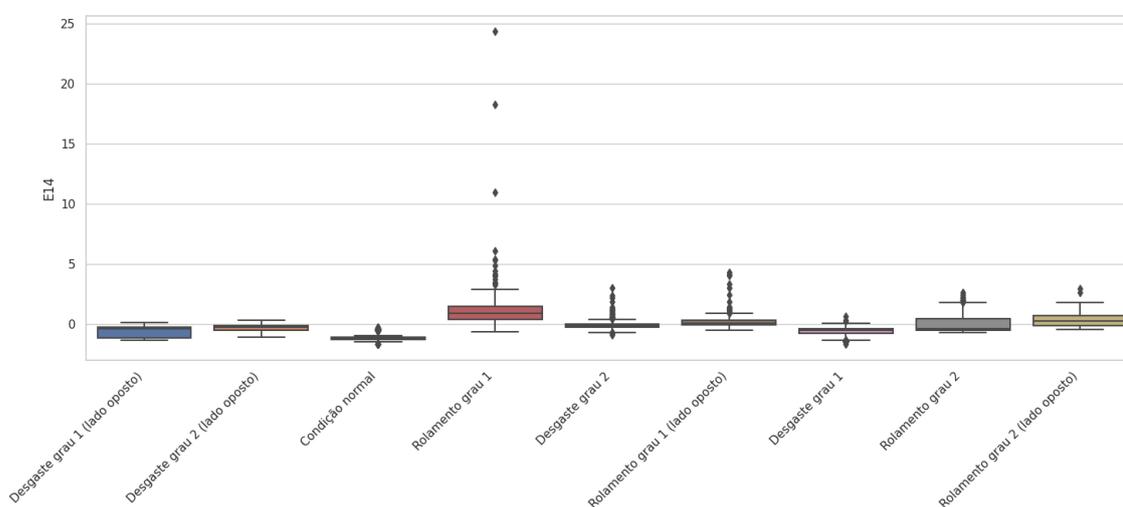
Figura 5. 8 – Boxplots de classes na energia de banda ‘E13’ no eixo axial



Fonte: O autor

Na Figura 5.9, assim como na energia de banda ‘E16’, os conjuntos apresentam baixa dispersão em todos os rótulos de classes da *feature* ‘E14’, o que as tornam *features* complementares para a detecção e diagnóstico de falhas dos rolos.

Figura 5. 9 – Boxplots de classes na energia de banda ‘E14’ no eixo axial



Fonte: O autor

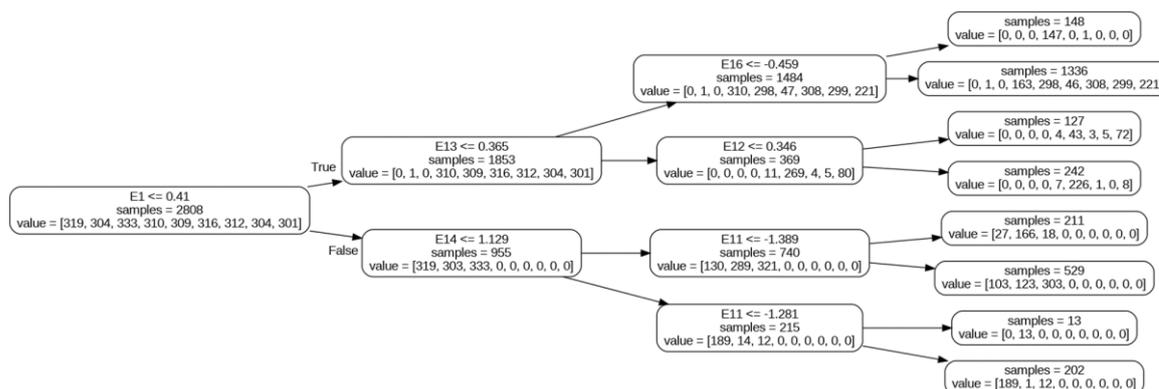
Com a avaliação das principais *features* nas três direções mediante a ANOVA e os *boxplots*, é possível verificar que é preciso aplicar mais de uma *feature* para um diagnóstico mais completo acerca do estado dos rolos. Além disso, apesar de apresentar as *features* com os maiores graus de significância, a partir dos valores-*p*, a ANOVA não fornece regras ou limites para determinação de cada rótulo, sendo

limitado a eliminar as *features* menos relevantes, para uma análise posterior com parametrização de falhas por análise humana ou por aprendizado de máquinas.

5.1.2. Árvore de decisão

Diferente da seleção de *features* com a ANOVA, a árvore de decisão identifica a combinação das *features* utilizadas para a classificação dos diferentes estados do rolo, por meio das regras criadas pelo modelo. Além disso, apresenta as *features* com maior significância por meio da regra sobre a energia de banda presente no nó raiz, bem como das primeiras ramificações. Assim como a ANOVA foram criados um modelo de árvore de decisão para as direções vertical, horizontal e axial, as quais são representadas pelas Figuras 5.10, 5.11 e 5.12, respectivamente.

Figura 5. 10 – Árvore de decisão no eixo vertical

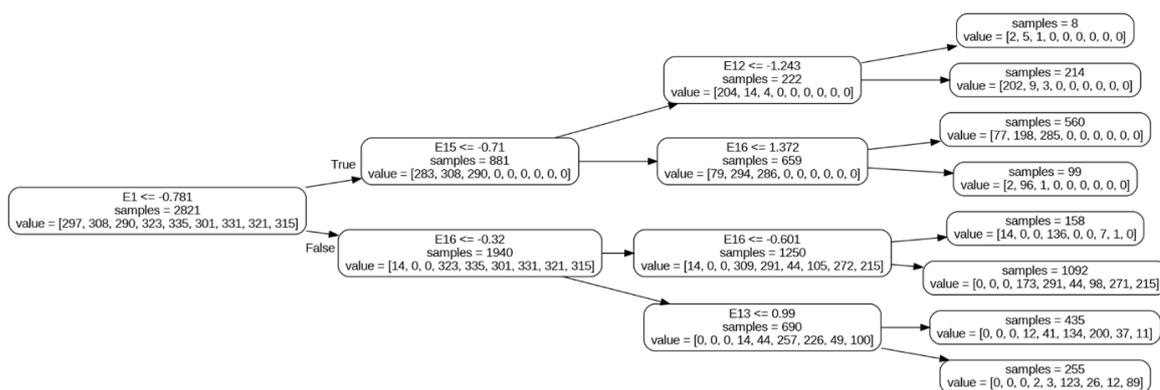


Fonte: O autor

Na Figura 5.10, assim como na ANOVA, percebe-se que a energia de banda 'E1' foi considerada a *feature* mais relevante na árvore de decisão, devido a regra do nó raiz ter sido tomada pelo modelo a partir dessa *feature*. Outra semelhança com a Análise de Variância foi a seleção da energia de banda 'E13', aplicada como regra na primeira ramificação do modelo, dentre as mais significantes.

Apesar disso, a energia de banda 'E14' foi priorizada na árvore de decisão, enquanto a ANOVA considerou a *feature* 'E11' dentre as três mais significativas. Ainda assim, a energia 'E11' foi aplicada na segunda ramificação do modelo, assim como as energias 'E12' e 'E16', totalizando em seis *features* selecionadas na árvore de decisão na direção vertical.

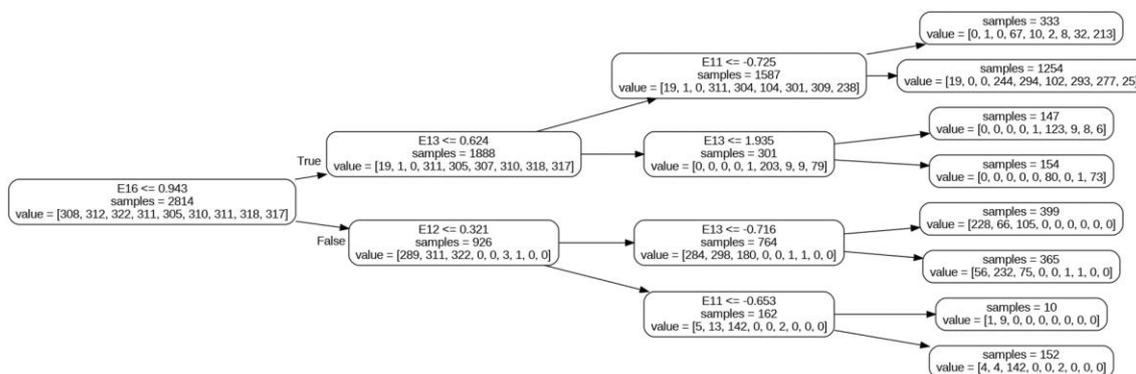
Figura 5. 11 – Árvore de decisão no eixo horizontal



Fonte: O autor

Na Figura 5.11, também foi selecionada a energia ‘E1’ como *feature* principal tal como a avaliação da ANOVA na direção horizontal. Todavia, as energias de banda ‘E15’ e ‘E16’ foram selecionadas na primeira ramificação da árvore, diferentemente da Análise de Variância que priorizou *features* como as energias ‘E13’ e ‘E14’. Apesar disso, a energia ‘E13’ foi aplicada na segunda ramificação, embora a ‘E14’ não tenha sido incluída no modelo. Por fim, a *feature* ‘E12’ também foi selecionada para compor as regras do modelo.

Figura 5. 12 – Árvore de decisão no eixo axial



Fonte: O autor

Na Figura 5.12, é notável que, assim como na ANOVA, a árvore de decisão não priorizou a energia de banda ‘E1’ na direção axial para seleção de *features*, bem como classificou a energia ‘E16’ na regra do nó raiz e, portanto, como a *feature* mais significativa do modelo. Na primeira ramificação foram selecionadas as energias ‘E12’ e ‘E13’, sendo a última *feature* também presente dentre as três *features* com maior grau de significância na ANOVA.

Por último, a *feature* 'E11' foi aplicada nas regras da segunda ramificação, assim como a 'E13'. Destarte, é perceptível que o modelo de árvore de decisão das energias de banda *wavelet* na direção axial selecionaram uma quantidade menor de *features* em relação as demais direções. Logo, infere-se que na direção axial, houve maior dificuldade na classificação das regras do modelo, ao qual pode prejudicar na acuracidade das técnicas de aprendizado para o diagnóstico de falhas.

De modo geral, devido à similaridade entre a seleção de *features* reproduzidas pela ANOVA e árvore de decisão, optou-se por selecionar as seguintes *features* para os modelos de aprendizado com redução de dimensionalidade: na direção vertical optou-se as energias de banda 'E1', 'E11', 'E12', 'E13', 'E14', 'E15' e 'E16'; na direção horizontal foram escolhidas as energias 'E1', 'E11', 'E12', 'E13', 'E14', 'E15' e 'E16'; e para a direção a direção axial foram selecionadas as *features* 'E10', 'E11', 'E12', 'E13', 'E14', 'E15' e 'E16'.

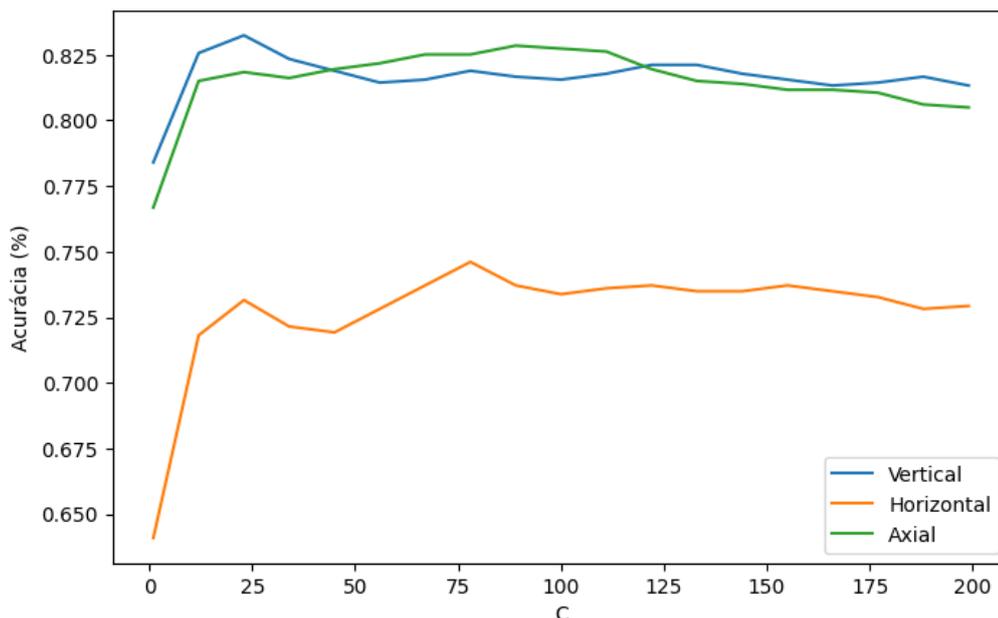
5.2. Máquina de Vetores de Suporte

5.2.1. Sem seleção de *features*

O algoritmo de aprendizado, que analisou todo o conjunto de dados sem a eliminação de nenhuma *feature*, criou 19 diferentes modelos, para cada direção, por meio da técnica SVM, com variações do parâmetro C entre 1 a 199. A Figura 5.13 apresenta a acuracidade de cada direção dos modelos SVM para diferentes valores de C. Nota-se que os melhores resultados foram nas direções vertical e axial.

Ambas direções apresentaram acuracidades superiores a 75% ao longo da variação do parâmetro C e superiores a 80% para $C > 12$. De modo geral as *features* na direção vertical apresentaram os melhores desempenhos, alcançando maior precisão de acerto no diagnóstico de 83,2% quando $C = 23$. No entanto para $45 < C < 122$, os modelos na direção axial se sobressaíram alcançando a máxima precisão de 82,8% para $C = 89$. Com desempenhos inferiores, os modelos de diagnóstico dos sinais na direção horizontal apresentaram taxas de acerto do diagnóstico de falha dos rolos que variaram entre 64,1% a 74,6%, sendo a acuracidade máxima alcançada para o modelo SVM com parâmetro $C = 78$.

Figura 5. 13 – Acuracidade SVM (sem seleção de *features*)



Fonte: O autor

Em relação ao melhor modelo de aprendizado SVM criado na direção vertical, foi plotado uma matriz de confusão para avaliar, mais detalhadamente, o quantitativo de acertos por cada classe. Na matriz de confusão (Figura 5.14) é possível observar que em 7 dos 89 diagnósticos do modelo foram classificados com o estado de ‘condição normal’ quando, na verdade, o rolo apresentava desgaste de grau 1 em ambas laterais do cavalete.

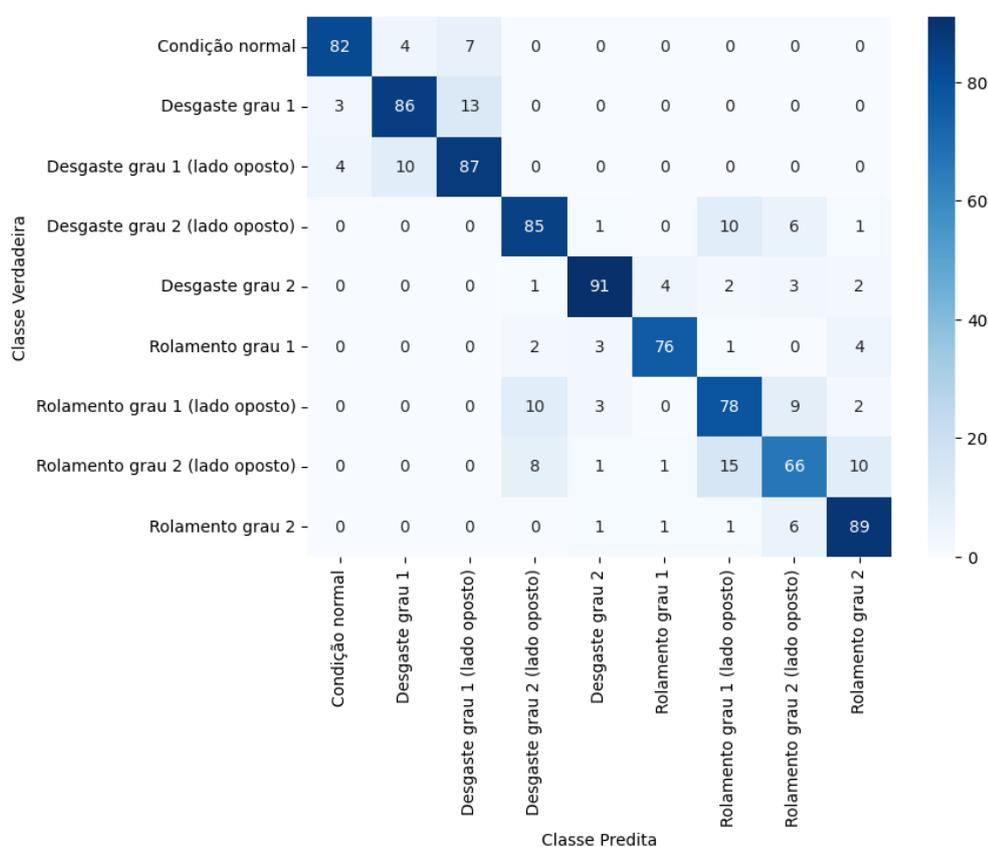
Esses casos são denominados de falso negativos, quando o modelo de aprendizado não conseguiu diagnosticar o defeito no rolo. Esse erro já havia sido discutido quando realizado a ANOVA em conjunto com os *boxplots* e, é indicativo que o modelo possui dificuldade em avaliar o rolo que apresenta desgaste de baixa severidade. Porém, não houve falso negativo para os demais estados defeituosos do rolo, resultando em mais de 92% dos casos com o diagnóstico correto.

Para o diagnóstico de desgaste de grau 1 do rolo, os erros mais frequentes concentraram-se na identificação equivocada da posição do rolo defeituoso, aos quais correspondem a, aproximadamente, 11,1% dos diagnósticos nesse estado. Os demais erros, foram falso positivos, ao qual, o rolo estava em estado saudável, porém, foi identificado com esse grau de desgaste.

Em relação ao rolo com desgaste grau 2, além dos defeitos no rolamento nos graus 1 e 2, não houve diagnósticos com falso positivos nem falsos diagnósticos de

desgaste grau 1. Apesar disso, houve erros de classificações entre esses estados, bem como, erros de identificação da posição do rolo. Dentre os erros, destaca-se a dificuldade de identificar qual o tipo de falha do rolo no lado oposto a posição do sensor, ao qual diagnosticou erroneamente, 29 de 107 casos classificados como defeito em 'rolamento grau 1 (lado oposto)', sendo 15 casos pertencentes a sinais do rolo com defeito em 'rolamento grau 2 (lado oposto)' e 10 casos a sinais do rolo com 'desgaste grau 2 (lado oposto)'.

Figura 5. 14 – Matriz de Confusão SVM (Sinal vertical)



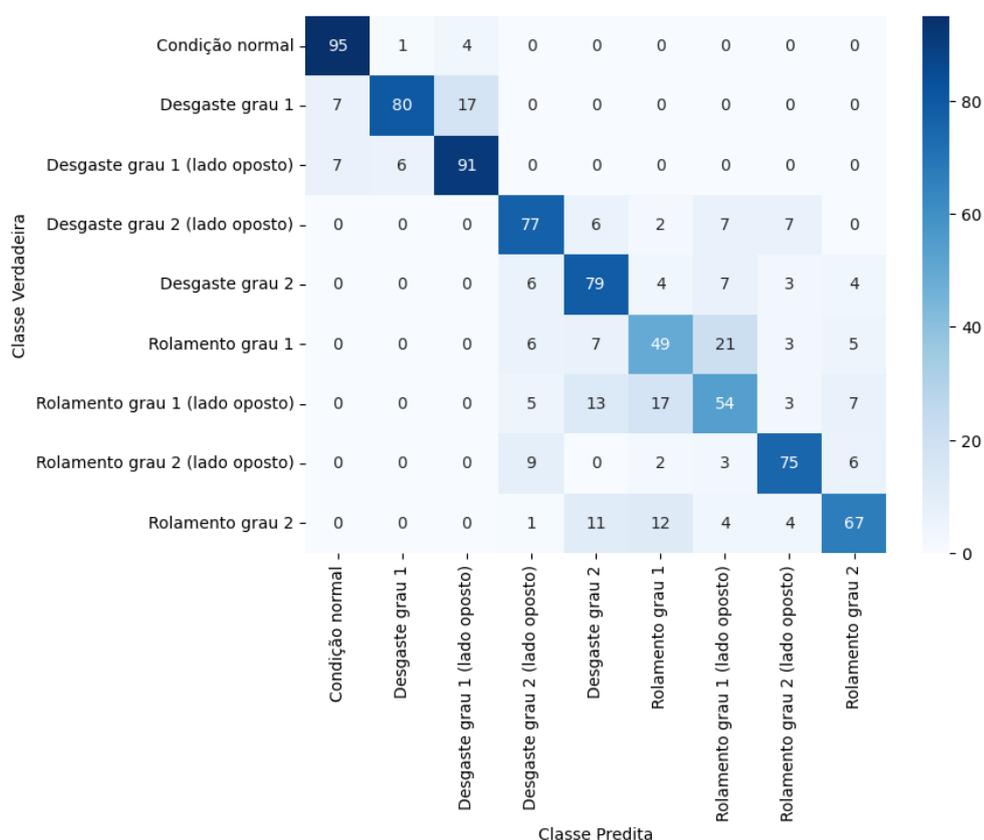
Fonte: O autor

A Figura 5.15 apresenta a matriz de confusão do modelo SVM com melhor acuracidade na direção horizontal. Na matriz de confusão é observável que houve, aproximadamente, 12,8% de diagnósticos com falso negativos, totalizando 14 em 109 casos. Assim como testado na direção vertical, na horizontal, os erros relacionados a falso negativo estão restritos a rolos que apresentam 'desgaste grau 1' em ambas extremidades do cavalete.

Em relação aos falsos positivos, foram diagnosticados erroneamente apenas 5 em 100 casos testados, também envolvendo diagnósticos equivocados de desgaste grau 1 para rolos em condições saudáveis. Portanto, de modo análogo aos diagnósticos na direção vertical, corrobora-se que rolos com desgastes nas fases mais prematuras estão mais suscetíveis a não serem detectadas, diferente dos demais modos de falha.

Também, é indubitável abordar os equívocos relacionados a posição do rolo defeituoso tendo o acelerômetro como referencial no cavalete. Os erros de diagnósticos mais evidentes, para este caso, foram em relação a defeitos em 'rolamento grau 1' que apresentaram 38 erros de posicionamento do rolo em 141 casos, apesar do modelo identificar a presença de um rolo defeituoso no cavalete verificado.

Figura 5.15 – Matriz de Confusão SVM (Sinal horizontal)

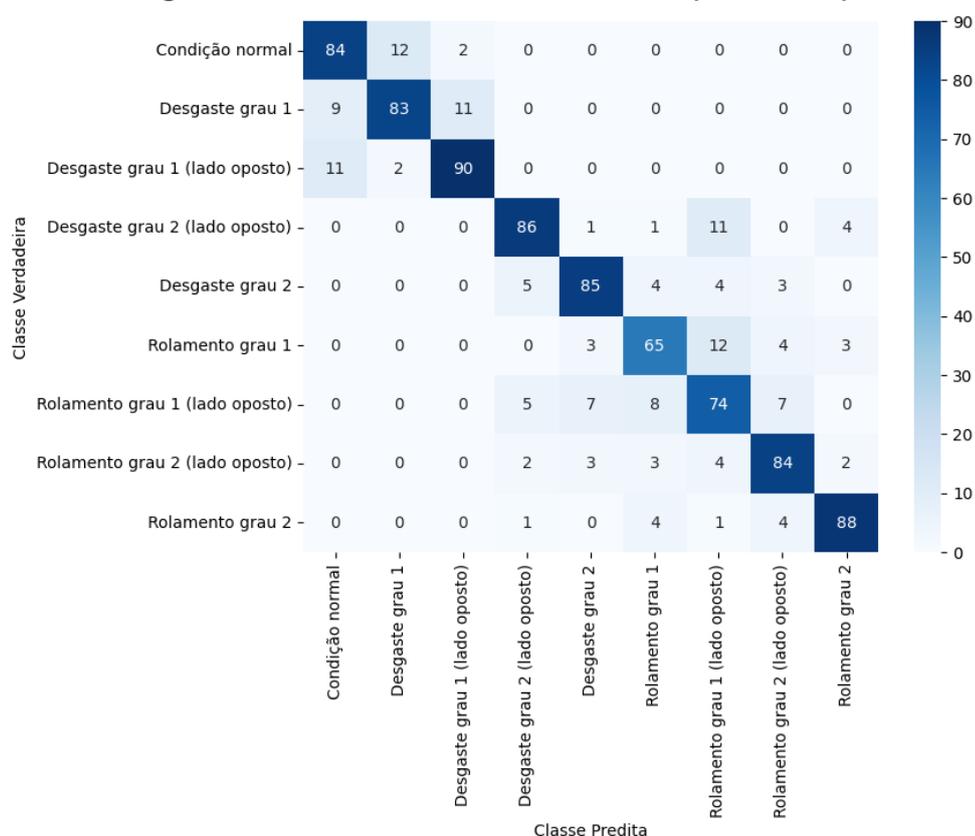


Fonte: O autor

A Figura 5.16 apresenta a matriz de confusão do melhor modelo SVM para o diagnóstico de falha em rolos a partir de sinais de vibração na direção axial. Sobre os

casos de diagnóstico de falso negativo, foram observados que em 19,2% dos testes foram diagnosticados, de modo equivocado, sinais em ‘condição normal’. Além disso, assim como nas demais direções os falsos negativos estavam relacionados, exclusivamente, com os rolos pertencentes as classes ‘desgaste grau 1’ e ‘desgaste grau 1 (lado oposto)’. Sobre os casos de falso positivo, houveram 14 diagnósticos incorretos, indicando sinais defeituosos em rolos com condições normais.

Figura 5. 16 – Matriz de Confusão SVM (Sinal axial)



Fonte: O autor

A Tabela 5.4 fornece uma avaliação geral dos melhores modelos de cada direção a partir da técnica SVM sem a seleção de *features*. Percebe-se que apesar da análise na direção vertical apresentar a melhor acurácia total, nas outras direções também se sobressaíram em alguns aspectos como na direção horizontal que apresentou menos diagnósticos equivocados de falso positivo, enquanto na direção axial houve ótima taxa de acertos do posicionamento do rolo danificado.

Tabela 5. 4 – Desempenho geral do algoritmo SVM (sem seleção de *features*)

Atributos	Vertical	Horizontal	Axial
Acurácia total (%)	83,2	74,6	82,8
Posicionamento do rolo (%)	94,0	87,3	93,5
Falso negativo (%)	7,86	12,8	19,2
Falso positivo (%)	11,8	5,00	14,3

Fonte: O autor

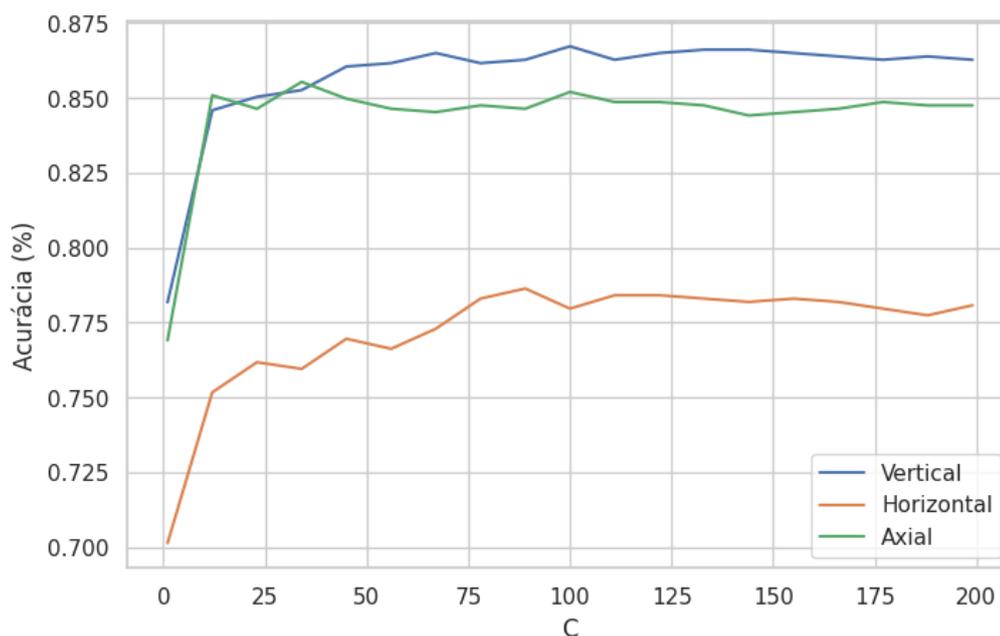
5.2.2. Com seleção de *features*

De modo similar a análise anterior, a abordagem com o algoritmo de aprendizado SVM com a seleção de *features* criou 19 diferentes modelos, para cada direção, por meio da técnica SVM, com variações do parâmetro C entre 1 a 199. A Figura 5.17 apresenta as acuracidades dos modelos SVM com seleção de *features* conforme a variação do parâmetro C. Assim como no caso anterior, as melhores direções para o diagnóstico de falhas foram a vertical e axial.

Ambas direções apresentaram taxas de acerto de diagnóstico acima 76,9% para todos os casos. Além disso, os modelos na direção vertical apresentaram as melhores acuracidades, com exceção nos casos de $C = 12$ e $C = 34$, ao qual a direção axial obteve melhor performance. O melhor modelo SVM de diagnóstico na posição vertical apresentou uma taxa de acerto de 86,7% para $C = 100$, enquanto na direção axial o melhor desempenho foi de 85,5% para $C = 34$.

Com desempenhos inferiores, os modelos de diagnóstico dos sinais na direção horizontal apresentaram taxas de acerto do diagnóstico de falha dos rolos que variaram entre 70,1% a 78,6%, sendo a acuracidade máxima alcançada para o modelo SVM com parâmetro $C = 89$. É notável o aumento substancial das performances dos modelos SVM após a seleção de *features*, indicando a eficácia de técnicas como ANOVA e árvore de decisão para avaliação das *features* mais significativas.

Figura 5. 17 – Acuracidade SVM (com seleção de *features*)



Fonte: O autor

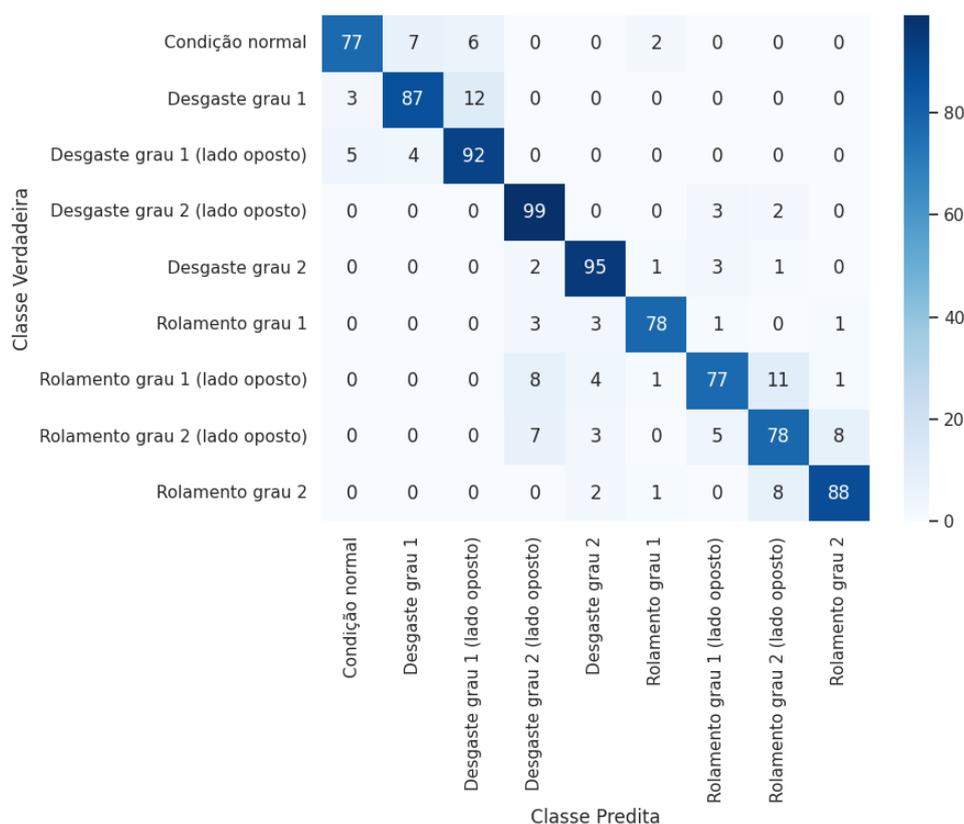
A Figura 5.18 apresenta a matriz de confusão do melhor modelo SVM com seleção de *features* na direção vertical. É possível observar que em 8 dos 85 diagnósticos do modelo foram classificados com o estado de ‘condição normal’ quando, na verdade, o rolo apresentava desgaste de grau 1 em ambas laterais do cavalete. Porém, assim como já discutido, não houve falso negativo para os demais estados defeituosos do rolo, resultando em mais de 90,6% dos casos com o diagnóstico correto.

Para o diagnóstico de desgaste de grau 1 do rolo, os erros mais frequentes concentraram-se na identificação equivocada da posição do rolo defeituoso, aos quais correspondem a, aproximadamente, 8,2% dos diagnósticos nesse estado. Os demais erros, foram falso positivos, ao qual, o rolo estava em estado saudável, todavia, foi identificado erroneamente com esse grau de desgaste.

Em relação aos defeitos no rolamento nos graus 1 e 2, não houve diagnósticos com falso positivos nem falsos diagnósticos de desgaste grau 1. Ainda assim, houve erros de classificações entre esses estados, bem como, erros de identificação da posição do rolo. Dentre os erros, destaca-se a dificuldade de identificar qual o tipo de falha do rolo no lado oposto a posição do sensor, ao qual diagnosticou erroneamente, 16 de 155 casos classificados, sendo 11 casos pertencentes a sinais do rolo com

defeito em ‘rolamento grau 1 (lado oposto)’ e 5 casos a sinais do rolo com ‘rolamento grau 2 (lado oposto)’.

Figura 5. 18 – Matriz de Confusão SVM com seleção de *features* (Sinal vertical)



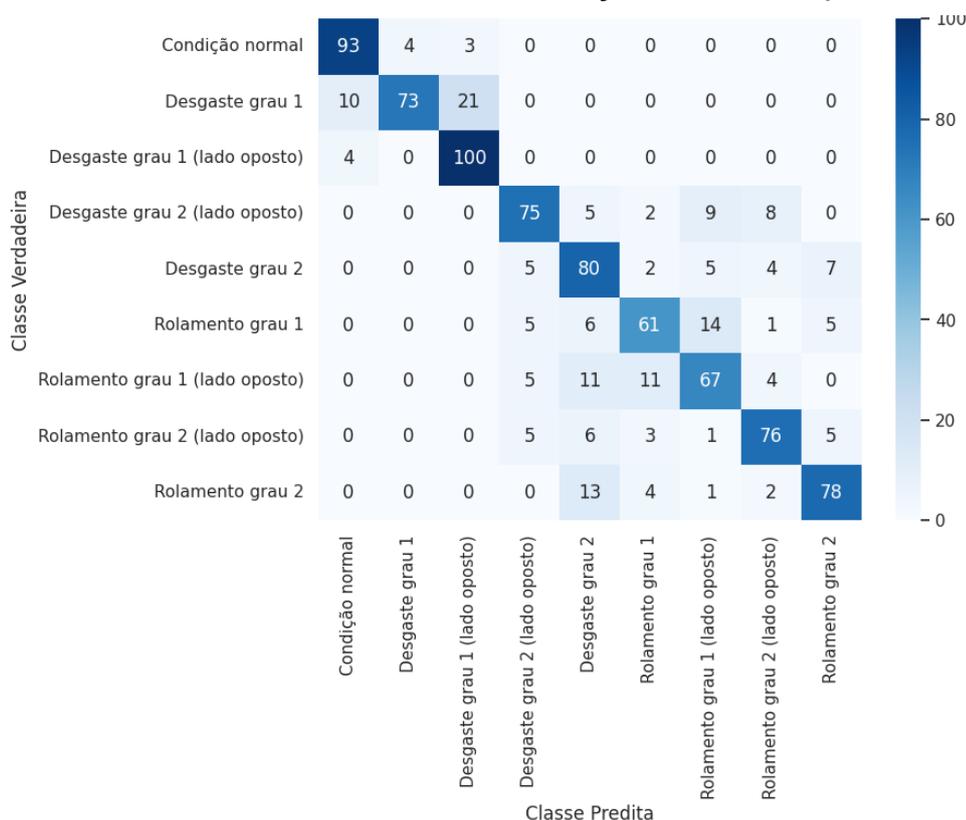
Fonte: O autor

Já a Figura 5.19 apresenta a matriz de confusão do modelo SVM com melhor acuracidade na direção horizontal. Na matriz de confusão é observável que houve, aproximadamente, 13,1% de diagnósticos com falso negativo, totalizando 14 em 107 casos. Assim como testado na direção vertical, na horizontal, os erros relacionados a falso negativo estão restritos a rolos que apresentam ‘desgaste grau 1’ em ambas extremidades do cavalete.

Em relação aos falsos positivos, foram diagnosticados erroneamente apenas 7 em 100 casos testados, também envolvendo diagnósticos equivocados de desgaste grau 1 para rolos em condições saudáveis. Sendo assim, similar aos demais casos, os rolos com desgastes nas fases mais prematuras estão mais suscetíveis a não serem detectadas, diferente dos demais modos de falha.

Por fim, vale também abordar os erros relacionados a posição do rolo defeituoso tendo o acelerômetro como referencial no cavalete. Os erros de diagnósticos mais evidentes, para este caso, foram relacionados aos defeitos em 'rolamento grau 1' que apresentaram 25 identificações equivocadas do posicionamento do rolo em 153 casos, apesar do modelo identificar corretamente a presença de um rolo defeituoso no cavalete verificado.

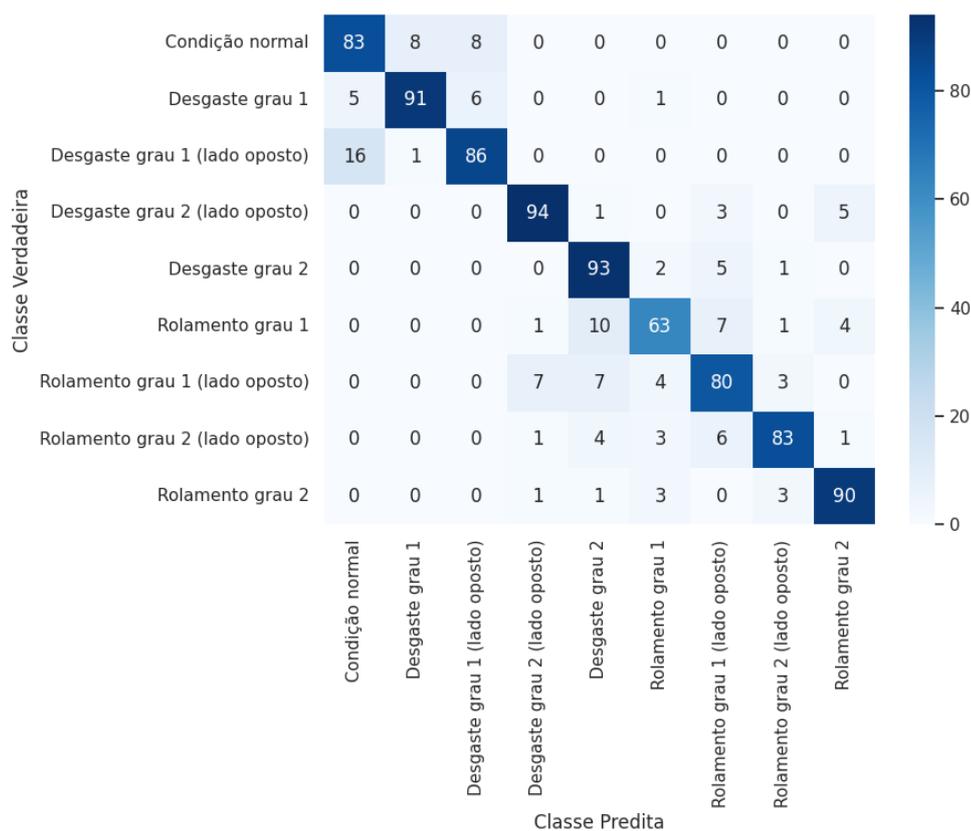
Figura 5. 19 – Matriz de Confusão SVM com seleção de *features* (Sinal horizontal)



Fonte: O autor

A Figura 5.20 apresenta a matriz de confusão do melhor modelo SVM para o diagnóstico de falha em rolos a partir de sinais de vibração na direção axial. Sobre os casos de diagnóstico de falso negativo, foram observados que em 20,2% dos testes foram diagnosticados, de modo equivocado, sinais em 'condição normal'. Igualmente nas demais direções, os falsos negativos estavam relacionados, exclusivamente, com os rolos pertencentes as classes 'desgaste grau 1' e 'desgaste grau 1 (lado oposto)'. Sobre os casos de falso positivo, houveram 16 diagnósticos incorretos, indicando sinais defeituosos em rolos com condições normais.

Figura 5. 20 – Matriz de Confusão SVM com seleção de *features* (Sinal axial)



Fonte: O autor

Os principais atributos dos melhores modelos SVM com seleção de *features* estão exibidos pela Tabela 5.5. Nota-se que de modo geral, a acurácia total dos modelos com seleção de *features* foram superiores aos modelos sem a seleção. Além disso, na direção axial houve maior taxa de acerto do posicionamento do rolo defeituoso, assim como na direção horizontal houve menor percentual de erro de diagnóstico por falso positivo, de modo similar a Tabela 5.4.

Tabela 5. 5 – Desempenho geral do algoritmo SVM (com seleção de *features*)

Atributos	Vertical	Horizontal	Axial
Acurácia total (%)	86,7	78,6	85,5
Posicionamento do rolo (%)	95,1	90,6	96,7
Falso negativo (%)	9,4	13,1	20,2
Falso positivo (%)	14,4	7,00	16,7

Fonte: O autor

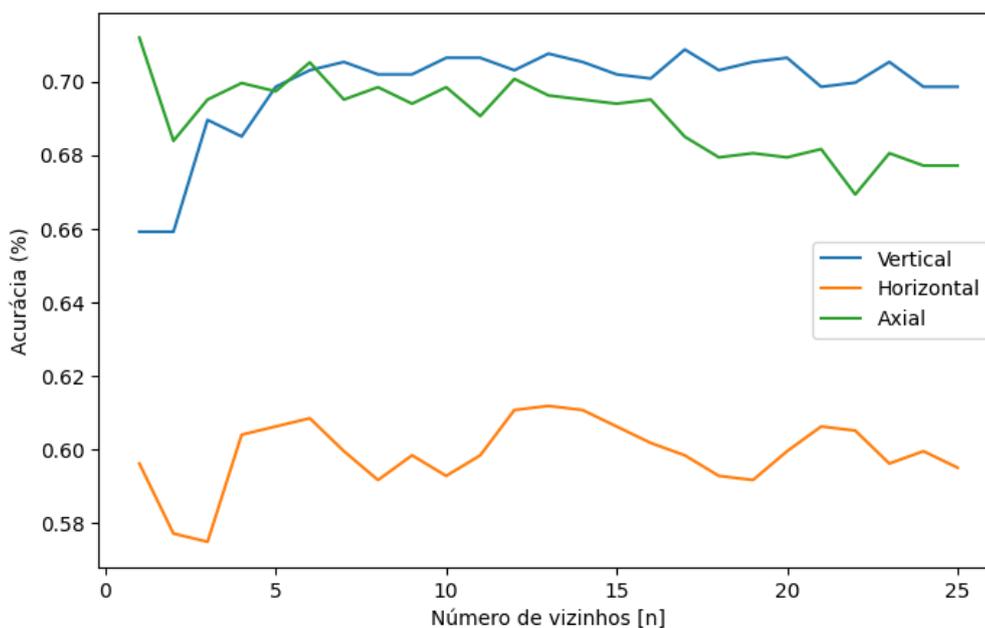
5.3. K-ésimo Vizinho mais Próximo

5.3.1. Sem seleção de *features*

Para avaliar o algoritmo de aprendizado kNN. sem a seleção de *features*, para o diagnóstico de falhas dos rolos, foram criados 25 modelos para cada direção do sinal de vibração, variando o número de vizinhos (n) de 1 a 25. A Figura 5.21 exibe a acurácia dos modelos para cada direção em relação ao número de vizinhos. Em linhas gerais, nota-se que as direções vertical e axial apresentaram melhores resultados, ao qual conforme a variação de n . foi observado que os diagnósticos na posição vertical obtiveram melhor performance, com exceção de alguns valores de n . ao qual a direção axial obteve a maior acuracidade dentre todos os modelos abordados com taxa de acerto de, aproximadamente, 71,2% dos diagnósticos quando $n = 1$.

Já o melhor modelo na posição vertical obteve acurácia de 70,9% para $n = 17$. Por último, com desempenhos inferiores o diagnóstico dos sinais na direção horizontal com maior acuracidade foi o modelo com precisão de 61,2% quando $n = 13$. Também é importante notar que em todas as direções, os piores modelos foram obtidos quando $n = 2$, indicando não ser aconselhável a escolha dessa configuração para esse parâmetro.

Figura 5. 21 – Acuracidade kNN (sem seleção de *features*)



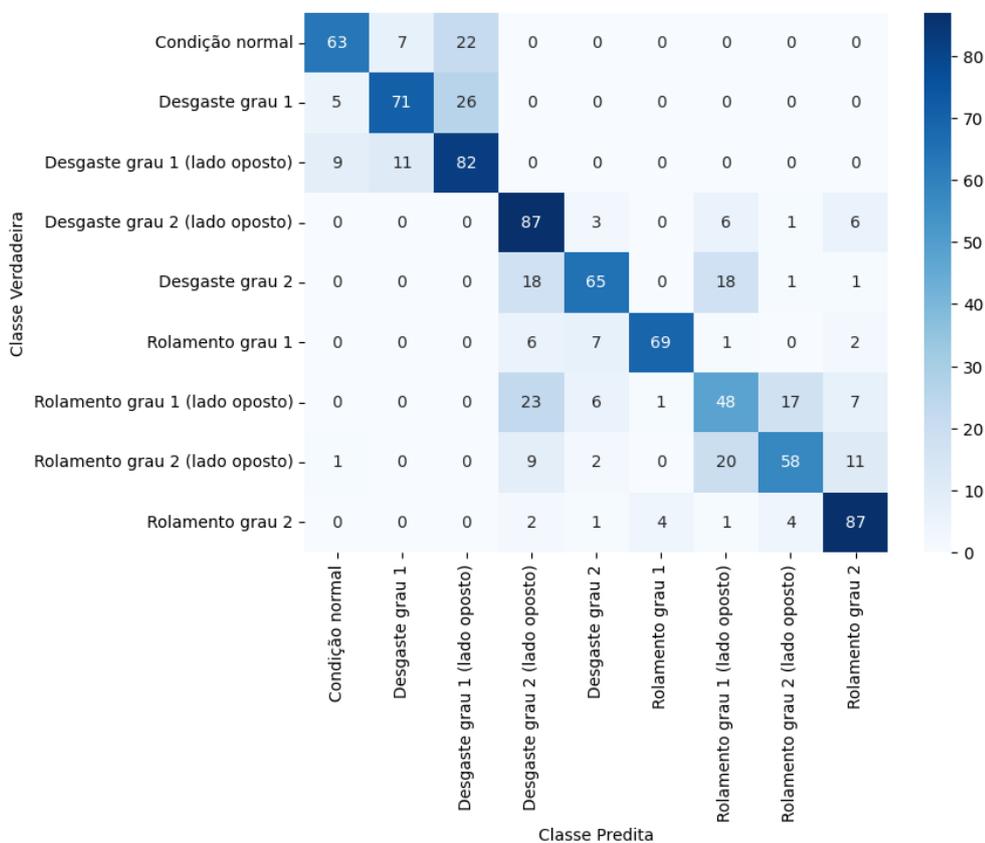
Fonte: O autor

A Figura 5.22 apresenta a matriz de confusão do melhor modelo kNN na direção vertical, sem a seleção de *features*. Foi constatado que em 14 dos 77 diagnósticos do modelo foram classificados como falso negativo. Como já discutido em outras abordagens, não houve falso negativo para os demais estados defeituosos do rolo, resultando em mais de 81,8% dos casos com o diagnóstico correto.

Os erros mais frequentes, para o diagnóstico de desgaste de grau 1 do rolo, concentraram-se, majoritariamente, na identificação equivocada da posição do rolo defeituoso, aos quais correspondem a, aproximadamente, 27,6% dos diagnósticos nesse estado. Sobre os defeitos no rolamento nos graus 1 e 2, houve erros de identificação da posição do rolo.

Por fim, destaca-se a dificuldade de identificar qual o tipo de falha do rolo no lado oposto a posição do sensor, ao qual diagnosticou erroneamente, 37 de 143 casos classificados, sendo 20 casos pertencentes a sinais do rolo com defeito em ‘rolamento grau 1 (lado oposto)’ e 17 casos a sinais do rolo com ‘rolamento grau 2 (lado oposto)’.

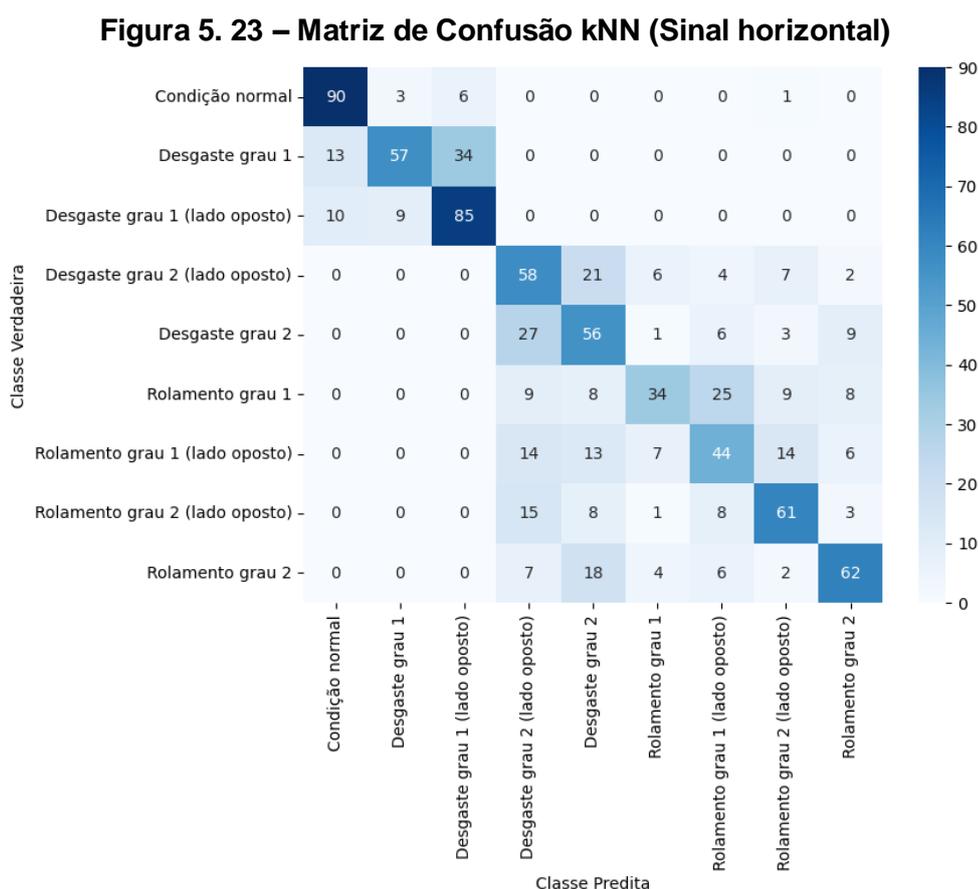
Figura 5. 22 – Matriz de Confusão kNN (Sinal vertical)



Fonte: O autor

A Figura 5.23 apresenta a matriz de confusão do modelo kNN com melhor acuracidade na direção horizontal. Na matriz de confusão é notável que houve, aproximadamente, 20,4% de diagnósticos com falso negativos, totalizando 23 em 113 casos. Não diferente dos demais casos, os erros relacionados a falso negativo estão restritos a rolos que apresentam ‘desgaste grau 1’ em ambas extremidades do cavalete.

Em relação aos falsos positivos, foram diagnosticados erroneamente apenas 9 em 100 casos testados, também envolvendo diagnósticos equivocados de ‘desgaste grau 1’, ‘desgaste grau 1 (lado oposto)’, além de um erro relacionado a defeito em ‘rolamento grau 2 (lado oposto)’ para rolos em condições saudáveis. Logo, corrobora-se mais uma vez que os rolos com desgastes nas fases mais prematuras estão mais suscetíveis a não serem detectadas, diferente dos demais modos de falha.

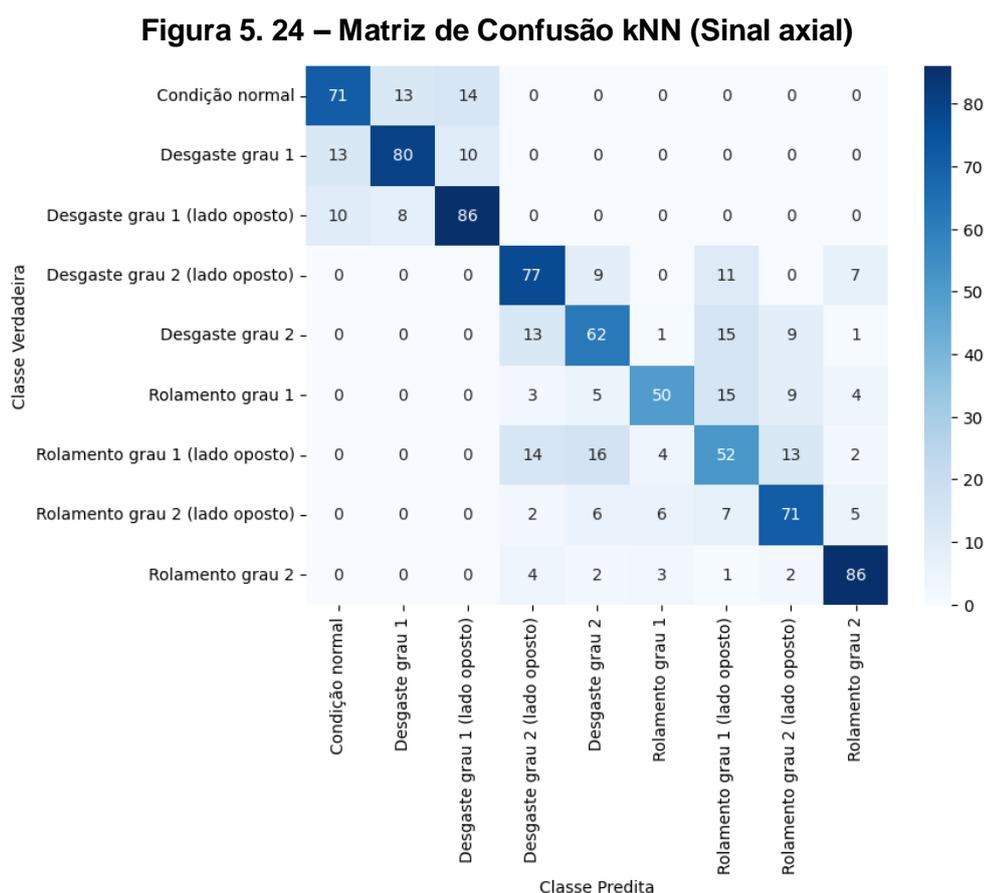


Fonte: O autor

É válido ressaltar também os erros relacionados a posição do rolo defeituoso tendo o acelerômetro como referencial no cavalete. Os erros de diagnósticos mais

evidentes, para este caso, foram relacionados aos defeitos em rolamento grau 1 que apresentaram 32 identificações equivocadas do posicionamento do rolo em 110 casos, apesar do modelo identificar corretamente a presença de um rolo defeituoso no cavalete verificado.

A Figura 5.24 apresenta a matriz de confusão do melhor modelo kNN para o diagnóstico de falha em rolos a partir de sinais de vibração na direção axial. Sobre os casos de diagnóstico de falso negativo, foram observados que em 24,5% dos testes foram diagnosticados, de modo equivocado, sinais em ‘condição normal’. Sobre os casos de falso positivo, houveram 27 diagnósticos incorretos, indicando sinais defeituosos em rolos com condições normais. Apesar de haver outros erros, devido os testes na direção axial apresentarem a maior acuracidade dentre todas as direções, nota-se que houve erros menos significativos.



Fonte: O autor

A Tabela 5.6 apresenta o desempenho geral dos principais modelos ANN-MLP nas três direções. Apesar do modelo de aprendizado na direção axial apresenta

a maior acurácia total, nota-se que nos demais atributos os demais modelos se sobressaíram. No modelo da direção vertical, houve maior performance relacionada a taxa de posicionamento do rolo danificado, bem como, nos casos de falso negativo. Enquanto no modelo da direção horizontal, houve a menor taxa de erro de diagnóstico por falso positivo

Tabela 5. 6 – Desempenho geral do algoritmo kNN (sem seleção de *features*)

Atributos	Vertical	Horizontal	Axial
Acurácia total (%)	70,9	61,2	71,2
Posicionamento do rolo (%)	88,3	78,1	78,1
Falso negativo (%)	18,2	20,4	24,5
Falso positivo (%)	27,6	9,1	27,6

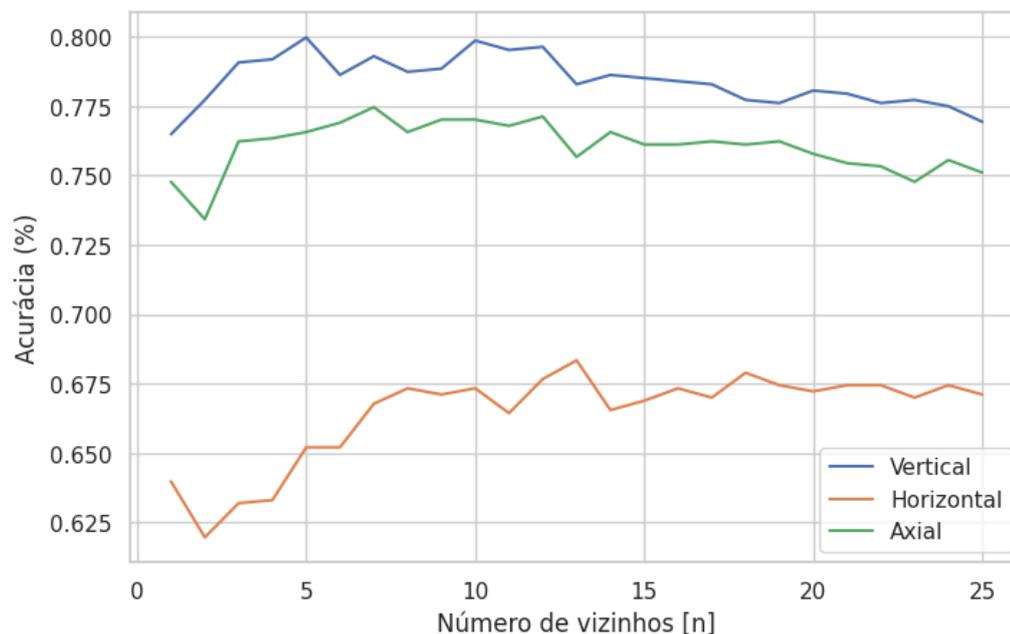
Fonte: O autor

5.3.2. Com seleção de *features*

Para a análise de diagnóstico de falha da técnica de aprendizado kNN com a seleção de *features*, também foi feito um comparativo entre acuracidades de diferentes modelos nas três direções conforme a variação do número de vizinhos (n), conforme exibido na Figura 5.25. Para esta nova abordagem, é perceptível que o aprimoramento da performance da técnica kNN com a seleção de *features* nas três direções.

Também se nota maior disparidade entre as acuracidades da direção vertical, em relação a direção axial, na qual nos modelos na direção vertical foram superiores para todo n analisado. Em relação aos melhores modelos para cada direção, obteve-se uma acurácia de, aproximadamente, 80% na direção vertical para $n = 5$, 68,3% na direção horizontal para $n = 13$ e, por último, 77,5% de acuracidade na direção axial para $n = 7$.

Figura 5. 25 – Acuracidade kNN (com seleção de *features*)

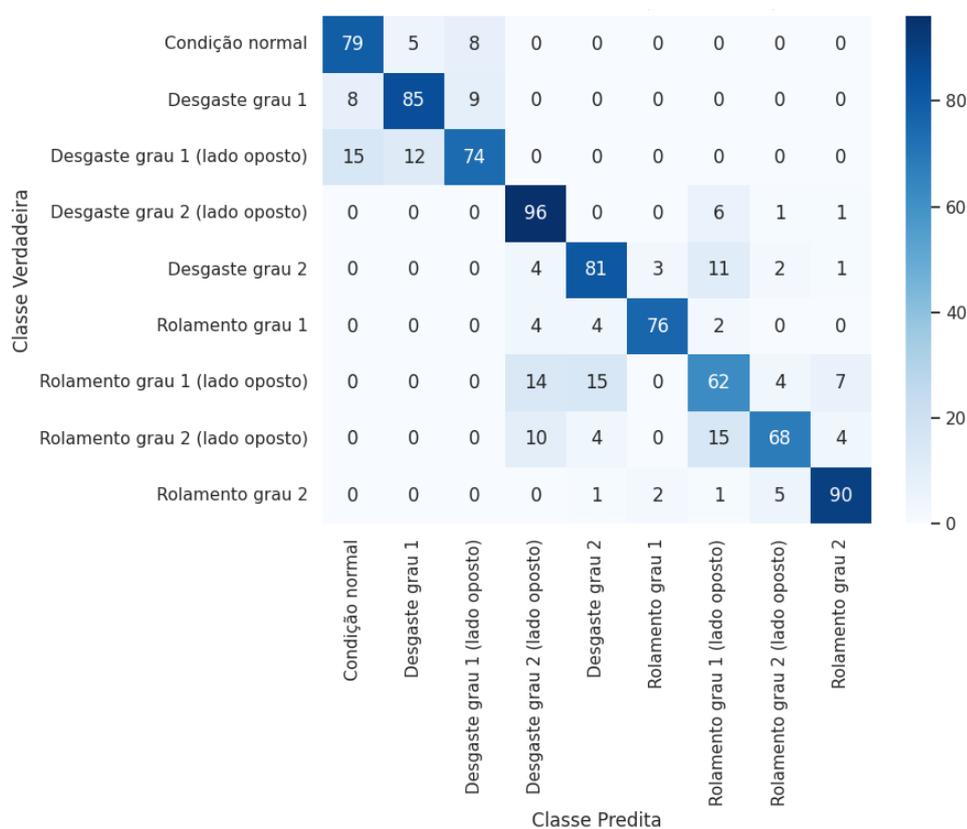


Fonte: O autor

Na matriz de confusão do modelo kNN na direção vertical com a seleção de *features*, conforme ilustrado pela Figura 5.26, foi constatado que em 23 dos 102 diagnósticos do modelo foram classificados como falso negativo. Como já discutido em outras abordagens, não houve falso negativo para os demais estados defeituosos do rolo, performando em mais de 77,4% dos casos com o diagnóstico correto.

Para o diagnóstico de desgaste de grau 1 do rolo, os principais erros se concentraram, principalmente, na identificação equivocada da posição do rolo defeituoso, aos quais correspondem a, aproximadamente, 11,7% dos diagnósticos nesse estado. Sobre os falsos positivos houve falha em 14,1% dos diagnósticos de ‘desgaste grau 1’ e ‘desgaste grau 1 (lado oposto)’. Por último, é válido observar que as maiores dificuldade de identificar qual o estado defeituoso do sinal foi na predição de classes cuja falha estava presente nos rolos do lado oposto do sensor de vibração.

Figura 5. 26 – Matriz de Confusão kNN com seleção de *features* (Sinal vertical)



Fonte: O autor

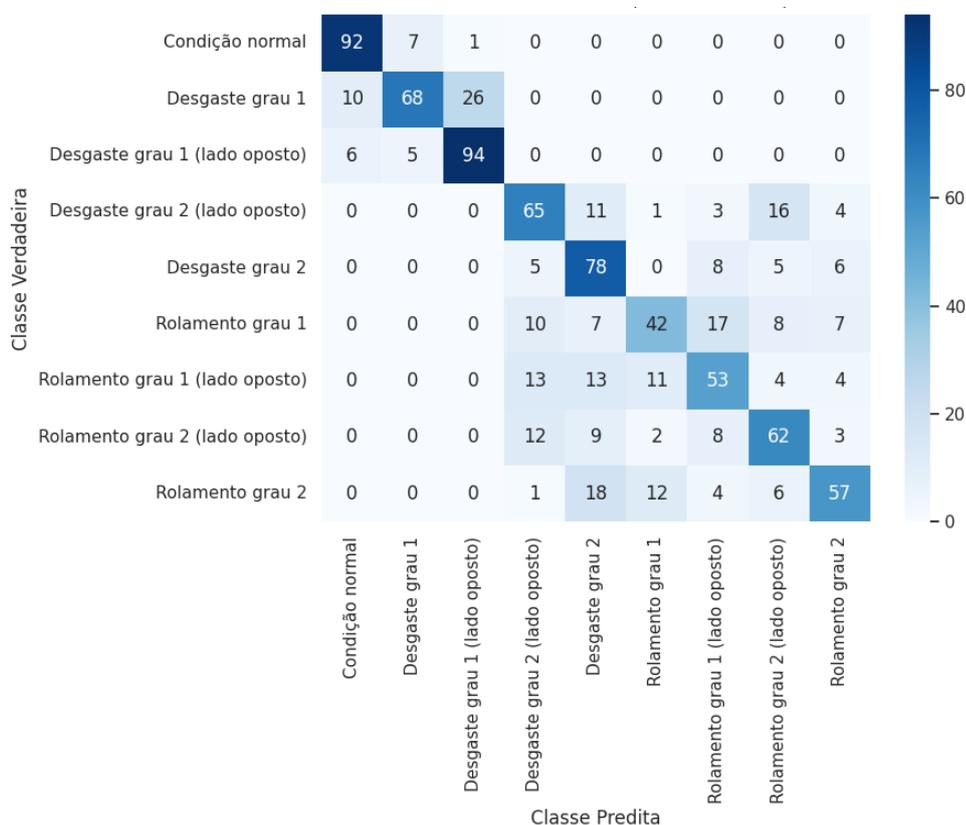
A Figura 5.27 apresenta a matriz de confusão do melhor modelo kNN na direção horizontal com a seleção de *features*. Na matriz de confusão é notável que houve, aproximadamente, 14,8% de diagnósticos com falso negativos, totalizando 16 em 108 casos. Reitera-se o que já foi abordado, ao qual os erros relacionados a falso negativo estão restritos a rolos que apresentam ‘desgaste grau 1’ em ambas extremidades do cavalete.

Em relação aos falsos positivos, foram diagnosticados erroneamente apenas 8 em 100 casos testados, também envolvendo diagnósticos equivocados de ‘desgaste grau 1’ e ‘desgaste grau 1 (lado oposto)’ para rolos em condições saudáveis. Portanto, mais uma vez os rolos com desgastes nas fases mais prematuras são menos detectáveis, diferente dos demais modos de falha.

É válido ressaltar também os erros relacionados a posição do rolo defeituoso tendo o acelerômetro como referencial no cavalete. Os erros de diagnósticos mais evidentes, para este caso, foram relacionados aos defeitos em rolamento grau 1 que apresentaram 28 identificações equivocadas do posicionamento do rolo em 123

casos, apesar do modelo identificar corretamente a presença do rolo defeituoso no cavalete avaliado.

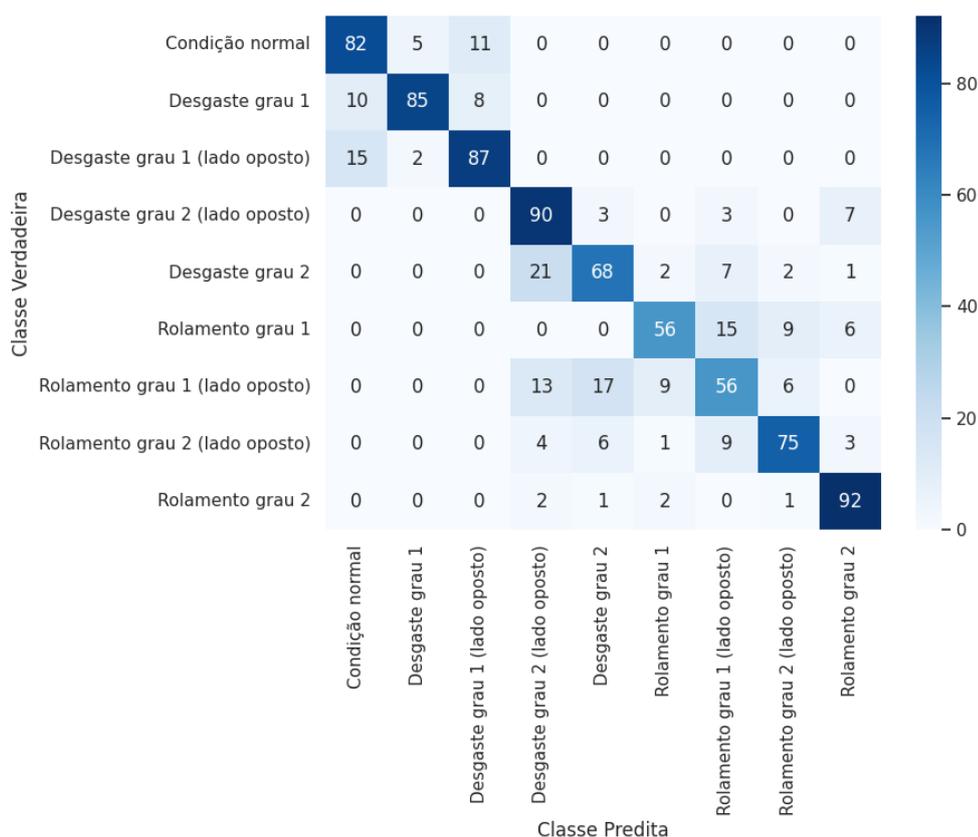
Figura 5. 27 – Matriz de Confusão kNN com seleção de *features* (Sinal horizontal)



Fonte: O autor

Por último, a Figura 5.28 apresenta a matriz de confusão do modelo kNN de melhor acuracidade na direção axial com a seleção de *features*. Sobre os casos de diagnóstico de falso negativo, foram observados que em 23,4% dos testes foram diagnosticados, de modo equivocado, sinais em ‘condição normal’. Sobre os casos de falso positivo, houveram 16 diagnósticos incorretos, indicando sinais defeituosos em rolos com condições normais. Consoante, apesar de haver outros erros, destaca-se os erros relacionados a predição de ‘desgaste grau 2’ e ‘desgaste grau 2 (lado oposto)’, ao qual o rolo apresentava o estado de defeito em ‘rolamento grau 1 (lado oposto)’.

Figura 5. 28 – Matriz de Confusão kNN com seleção de *features* (Sinal axial)



Fonte: O autor

A Tabela 5.7 apresenta os principais atributos dos melhores modelos kNN com a seleção de *features*. Vale destacar que apesar de apresentar a pior performance dentre as direções, o modelo na direção horizontal apresentou as menores taxa de erros por falso negativo e falso positivo.

Tabela 5. 7 – Desempenho geral do algoritmo kNN (com seleção de *features*)

Atributos	Vertical	Horizontal	Axial
Acurácia total (%)	80,0	68,3	77,5
Posicionamento do rolo (%)	94,6	86,1	90,8
Falso negativo (%)	22,5	14,8	23,4
Falso positivo (%)	14,1	8,00	16,3

Fonte: O autor

5.4. Perceptron Multicamadas

5.4.1. Sem seleção de *features*

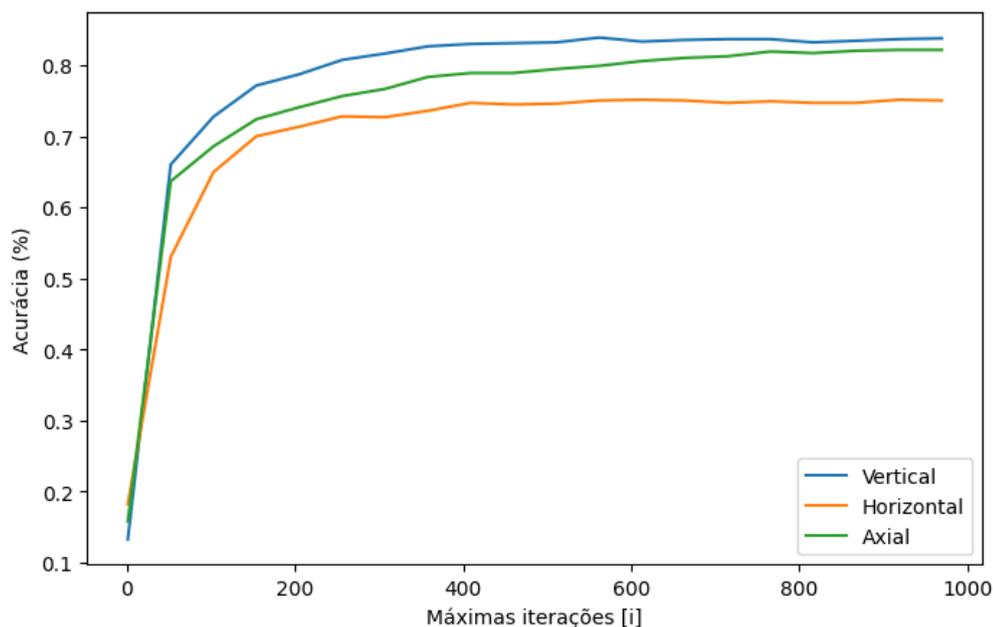
O algoritmo de aprendizado ANN-MLP, inicialmente, analisou todo o conjunto de dados sem a eliminação de nenhuma *feature*, criou 20 diferentes modelos, para cada direção, com variações do parâmetro de máximas iterações (i) entre 1 a 970. A Figura 5.29 apresenta a acuracidade de cada direção dos modelos ANN-MLP para diferentes valores de i .

Primeiramente, nota-se que as performances dos modelos nas três direções apresentaram um comportamento mais progressivo e gradual conforme aumentava o número de máximas iterações. Isso pode ser justificado, devido ao fato que redes neurais tendem a ter maior ajuste conforme o tamanho do conjunto de dados, bem como na otimização do modelo a partir do gradiente de erro da função preditora, o que é compensado pelo aumento do custo computacional.

Além disso, apesar das direções vertical e axial continuarem predominantes em termos de acuracidade de classificação, os modelos na direção horizontal apresentaram melhoras significativas relacionadas a taxa de acerto, podendo ser útil na combinação com as outras direções para um diagnóstico de falhas mais preciso. Outro ponto válido para ressaltar, é que nas três curvas da Figura 7.29, houve um ponto de inflexão em $i = 52$ ao qual percebe-se, que para $i > 52$, a taxa de progressão de acurácia em relação às máximas iterações começa a evoluir mais lentamente, até apresentaram um comportamento mais constante, ao qual pode-se encontrar um ponto ótimo de acurácia com o menor valor de i possível.

Por conseguinte, o modelo de maior desempenho na direção vertical obteve uma acuracidade de 83,9% para $i = 562$. Já na direção horizontal, houve máxima acuracidade de, aproximadamente, 75,2% para $i = 613$. Por fim, na direção axial, o modelo de maior acurácia foi de 82,2% para $i = 919$.

Figura 5. 29 – Acuracidade ANN-MLP (sem seleção de *features*)

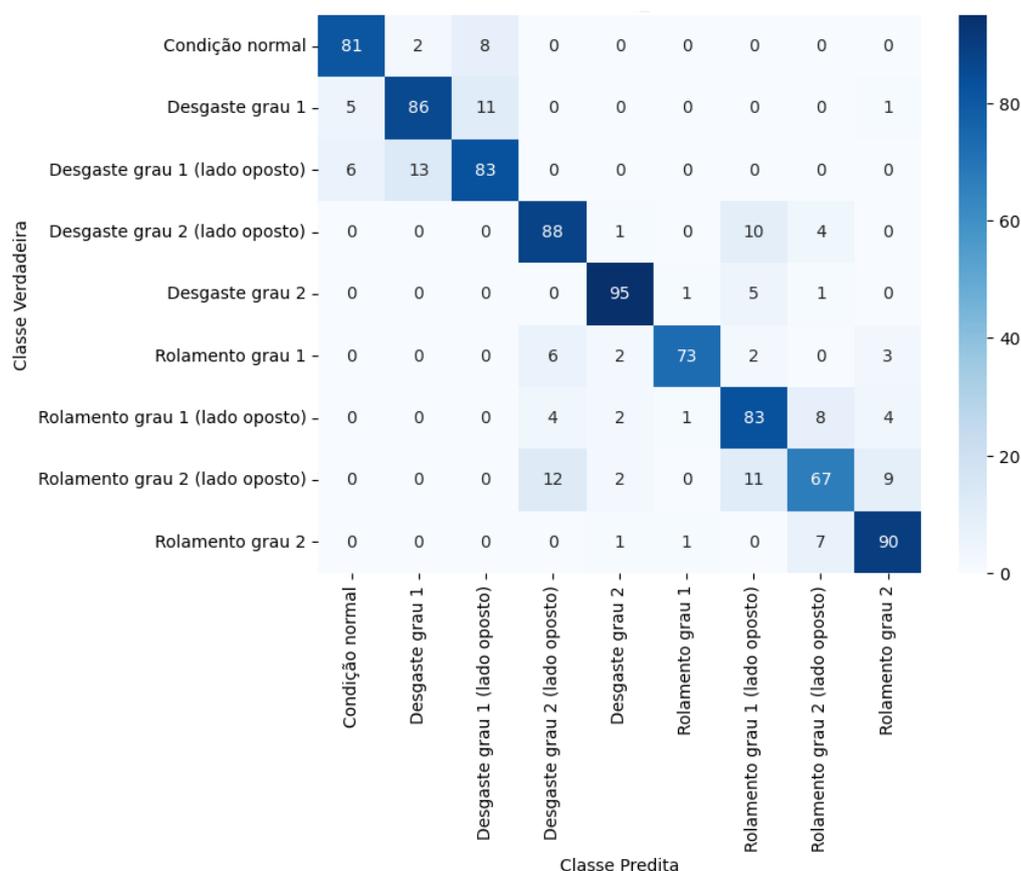


Fonte: O autor

A Figura 5.30 apresenta a matriz de confusão do melhor modelo de aprendizado ANN-MLP na direção vertical, sem a seleção de *features*. Em relação aos casos com falso negativo, observou-se que houve 11 em 92 diagnósticos errados, representando, aproximadamente, 12% dos casos ao qual o modelo classificou o sinal como um sinal saudável, apesar de apresentar as anomalias de desgaste grau 1, como já observado nas demais técnicas de aprendizado.

Acerca dos falsos positivos, foi detectado 10 erros dentro de 101 casos diagnosticados, representando um nível percentual de 9,9%, também relacionados, exclusivamente, a sinais de rolos que apresentam desgaste grau 1. Além disso, observou-se erros na indicação do posicionamento do rolo lateral com desgaste grau 1, ao qual, houve erro de 12,4%.

Figura 5. 30 – Matriz de Confusão ANN-MLP (Sinal vertical)

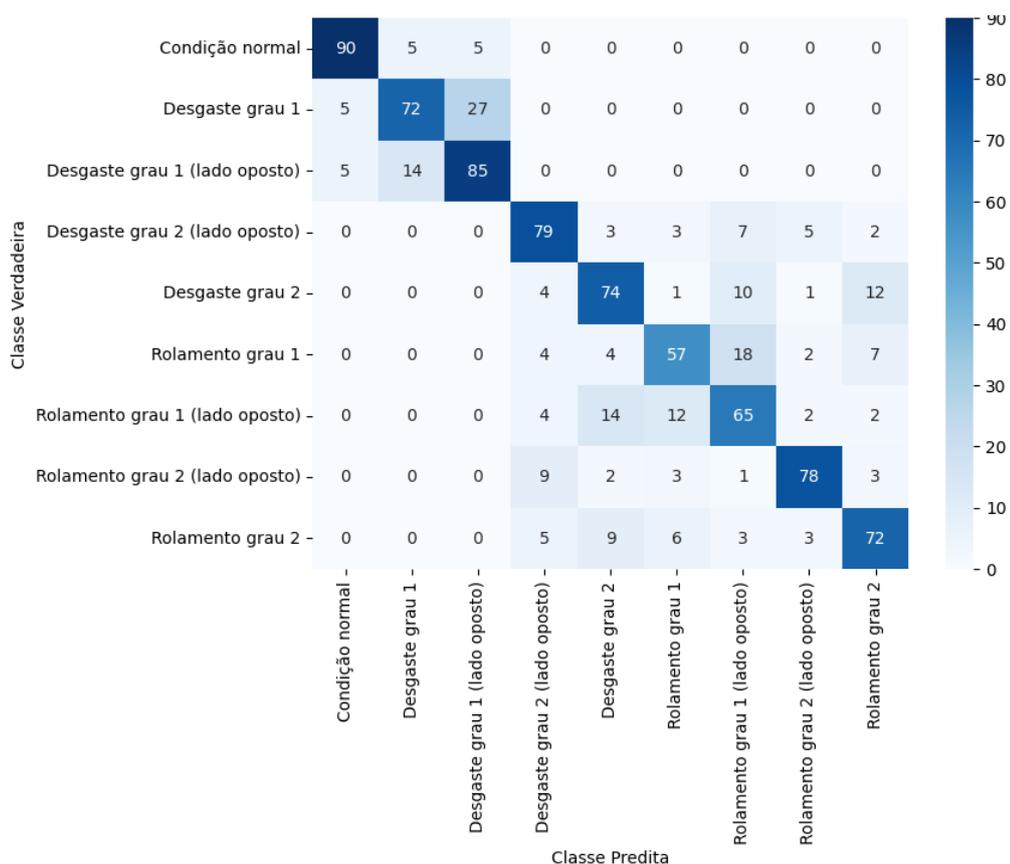


Fonte: O autor

A Figura 5.31 apresenta a matriz de confusão do melhor modelo ANN-MLP na direção horizontal, sem seleção de *features*. Para os casos de falso negativo, foi detectado 10 dentre 100 casos com diagnósticos errôneos acerca da condição do rolo indicando, em 90% das vezes, um sinal com boa saúde apesar da sua condição defeituosa. Para esses casos, também só foram para casos na qual o sinal pertencia a classe de sinais com desgaste grau 1.

Em relação aos falsos positivos, notou-se também a presença de 10% de erro de diagnóstico dentre 100 casos, também relacionados a rolos que apresentavam desgaste grau 1. Já sobre o erro de posicionamento do rolo que apresenta desgaste grau 1 no cavalete, notou-se que houve 20,7% de diagnósticos equivocados acerca do posicionamento, apesar da identificação correta do tipo de falha. De modo análogo, observa-se o mesmo caso para defeitos em rolamento grau 1, obtendo uma taxa de erro de 19,7%

Figura 5.31 – Matriz de Confusão ANN-MLP (Sinal horizontal)

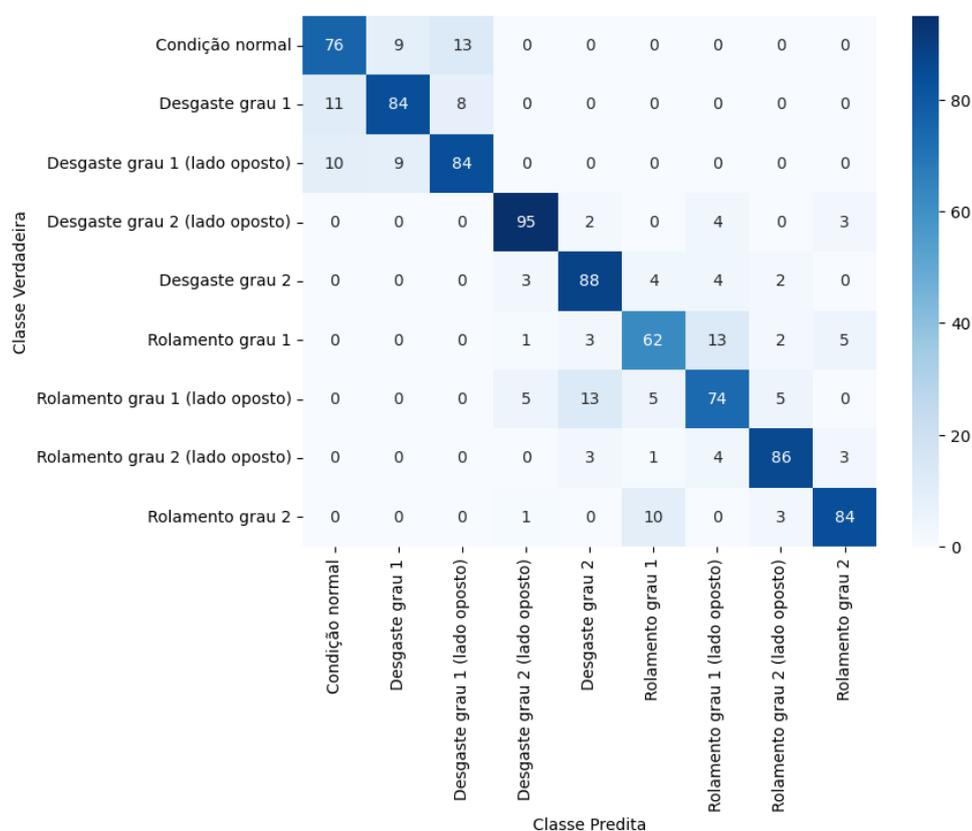


Fonte: O autor

Por fim, a Figura 5.32 exibe a matriz de confusão do melhor modelo ANN-MLP na direção axial, sem seleção de *features*. Sobre os casos com falsos negativos, identificou-se que em 21 de 97 casos (taxa de erro de 21,6%), houve a classificação indevida do rolo defeituoso (também relacionado a defeitos de desgaste grau 1), ao qual foi predito como um rolo em condição normal.

Acerca dos diagnósticos equivocados por falso positivo, foi identificado que houve 22 casos com diagnósticos errados dentre 98 casos (taxa de erro de 22,4%), também com erros de confusão apenas para rolos com desgaste grau 1. Outro ponto a ser observado foi erro de posicionamento do rolo com desgaste grau 1, ao qual, houve confusão em 9,2% dos casos.

Figura 5. 32 – Matriz de Confusão ANN-MLP (Sinal axial)



Fonte: O autor

A Tabela 5.8 apresenta o desempenho geral do algoritmo ANN-MLP nas três direções, sem seleção de *features*. Apesar da superioridade do modelo na direção vertical, principalmente, acerca da acurácia total e taxa de acerto do posicionamento do rolo defeituoso, o modelo na direção horizontal apresentou as menores taxas de erros relacionados a falso negativo e positivo.

Tabela 5. 8 – Desempenho geral do algoritmo ANN-MLP (sem seleção de *features*)

Atributos	Vertical	Horizontal	Axial
Acurácia total (%)	83,9	75,2	82,2
Posicionamento do rolo (%)	93,8	87,4	93,5
Falso negativo (%)	11,3	10,0	21,6
Falso positivo (%)	11,0	10,0	21,6

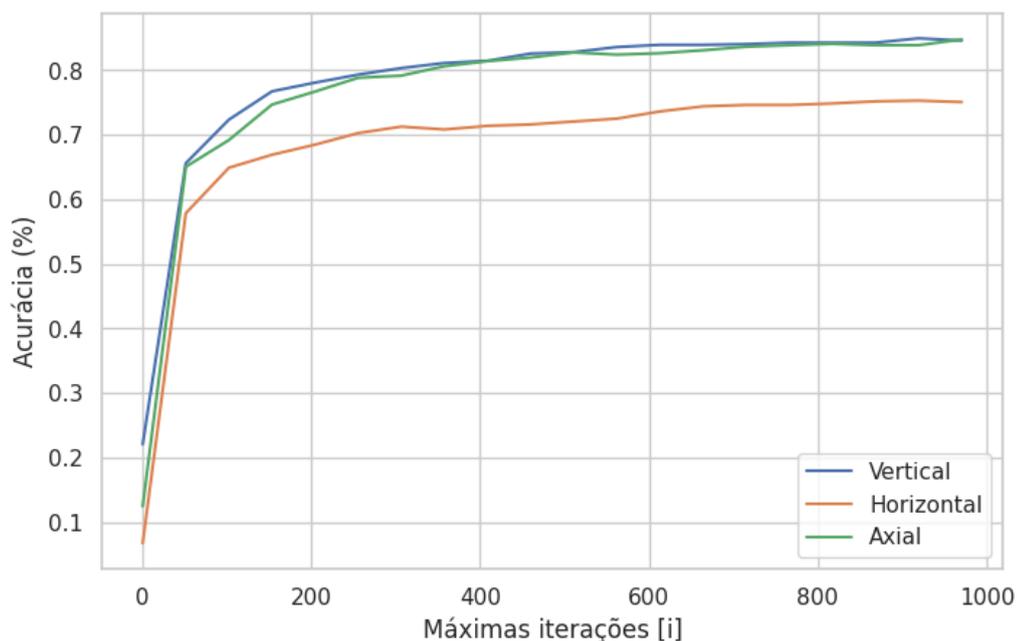
Fonte: O autor

5.4.2. Com seleção de features

Em relação aos algoritmos de aprendizado ANN-MLP nas três direções, com a seleção de *features*, foi exibido na Figura 5.33 as acuracidades dos modelos a partir da variação de i . Apesar de apresentar comportamento similar com a abordagem do ANN-MLP sem a seleção de *features*, é evidente que os modelos na direção axial obtiveram um desempenho mais próximos aos modelos na direção vertical.

Outro ponto a ser observado é que não houve uma evolução do desempenho entre modelos com e sem seleção de *features*, haja vista que as redes neurais possuem comportamentos mais estáveis. Os modelos na direção horizontal apresentaram as piores performances em relação as demais, assim como em todas as demais abordagens já realizadas. Sobre os melhores modelos, na direção vertical a maior acuracidade obtida foi de 84,9% para $i = 919$, na direção horizontal a maior acuracidade foi de 75,3% também para $i = 919$ e, por último, na direção axial o modelo de maior acurácia foi de 84,8% para $i = 970$.

Figura 5. 33 – Acuracidade ANN-MLP (com seleção de *features*)



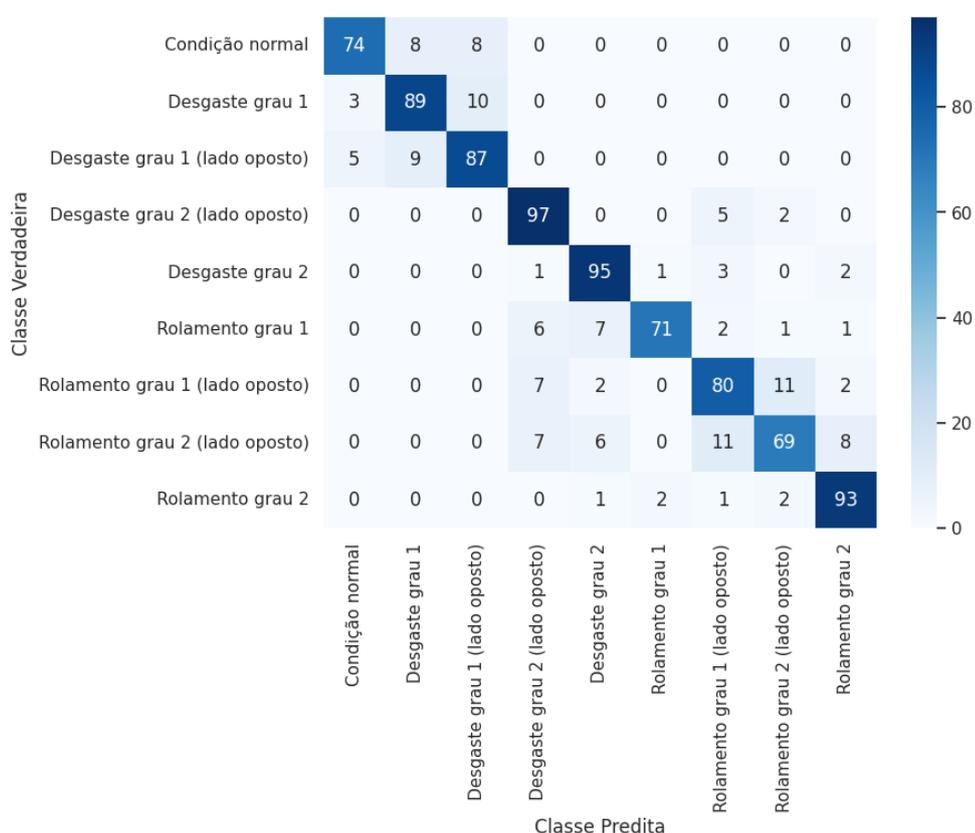
Fonte: O autor

A Figura 5.34 apresenta a matriz de confusão do modelo ANN-MLP na direção vertical, com seleção de *features*. Nos casos de falso negativos, foram detectados 8 casos de 82 que classificaram equivocadamente o sinal do rolo como condição

normal, resultando em 9,75% dos diagnósticos. Já em relação aos falsos positivos, foram um total de 16 em 90 casos, aos quais foram preditos como sinais de rolos com ‘desgaste grau 1’ e ‘desgaste grau 1 (lado oposto)’.

Em relação aos erros de posicionamento, é notável que no diagnóstico assertivo de desgaste grau 1, houve 9,74% de erros de identificação da posição do rolo defeituoso no cavalete. Outra dificuldade foi a identificação do grau de severidade dos defeitos em rolamento dos rolos no lado oposto ao sensor, cuja taxa de erro foi de 12,9%.

Figura 5. 34 – Matriz de Confusão ANN-MLP com seleção de *features* (Sinal vertical)

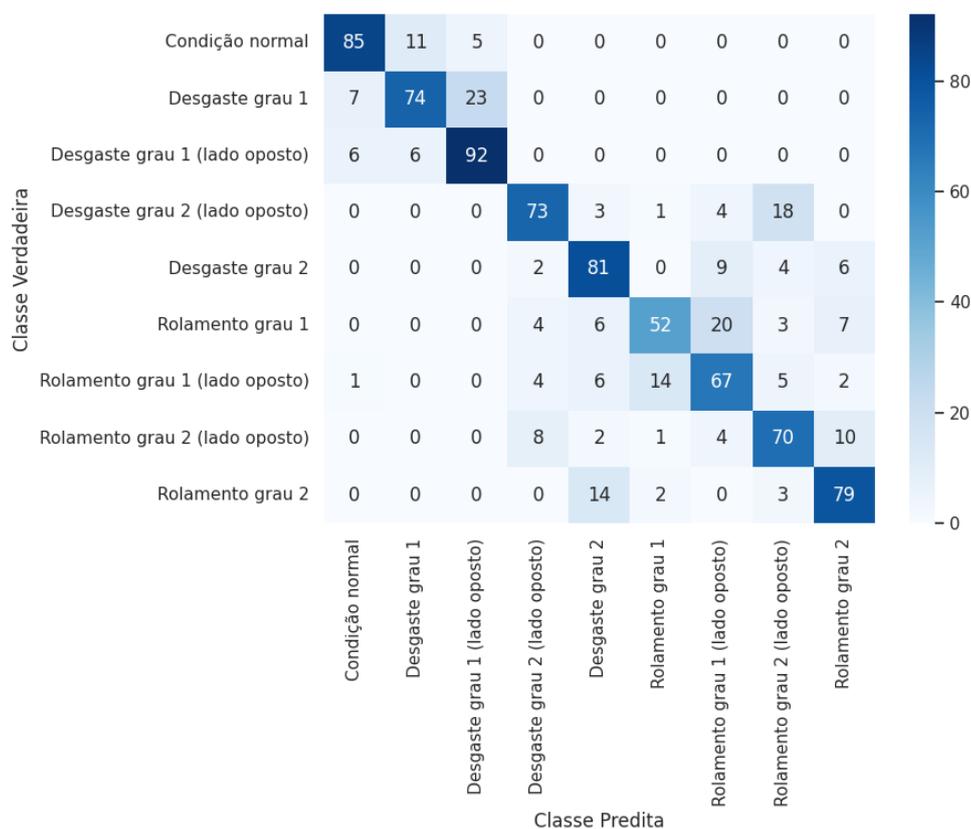


Fonte: O autor

A Figura 5.35 exibe a matriz de confusão do melhor modelo ANN-MLP na direção horizontal, com seleção de *features*. Sobre os casos de falso negativo, foi identificado 13 casos em 98 diagnósticos, totalizando uma taxa de erro de 15,3%. Enquanto para os casos de falso positivo, houve 16 falsas predições de sinais defeituosos dentre 101 casos, resultando uma taxa de 15,8% de erro. Sobre erros de posicionamento de rolos defeituosos já detectados, destacam-se as classes de

desgaste grau 1 e defeitos em rolamentos grau 1, que apresentaram percentual de erro de 14,9% e 22,2%, respectivamente.

Figura 5. 35 – Matriz de Confusão ANN-MLP com seleção de *features* (Sinal horizontal)

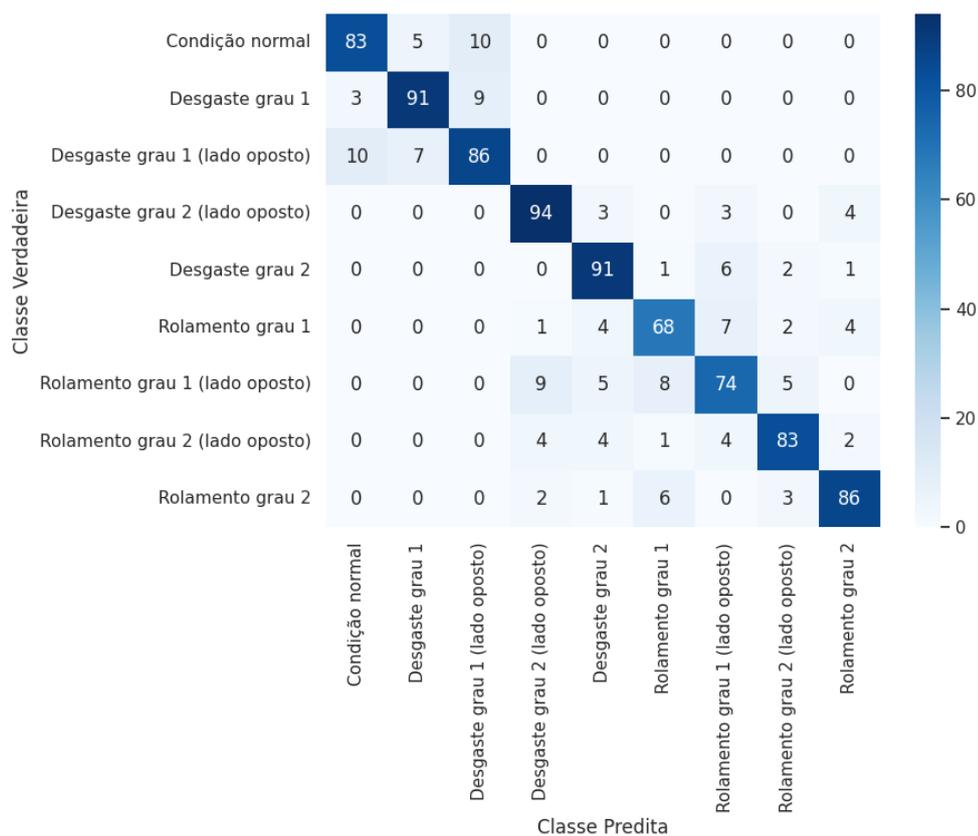


Fonte: O autor

A Figura 5.36 apresenta a matriz de confusão do melhor modelo ANN-MLP, com seleção de *features*, na direção axial. Para os casos de falso negativo, 13 em 96 casos apresentaram diagnósticos errôneos que classificaram sinais defeituosos como 'condição normal'. Todavia, em todas as técnicas e condições de seleção de *features*, os casos de falso negativos se restringiram a defeitos de desgaste grau 1.

Para casos de falso positivo, foram identificados 15 em 98 casos, totalizando em 15,3% de taxa de erro, também relacionados a rolos com desgaste grau 1. Em relação a erros de posicionamento dos rolos defeituosos como desgaste grau 1 e defeitos em rolamento grau 1, verificou-se que houve taxas de erro de 8,3% e 9,6%, respectivamente.

Figura 5. 36 – Matriz de Confusão ANN-MLP com seleção de *features* (Sinal axial)



Fonte: O autor

A Tabela 5.9 apresenta os principais atributos dos melhores modelos ANN-MLP para cada direção. De modo geral, o modelo na direção vertical apresentou os melhores *status*, com exceção dos casos de falso positivo, ao qual o modelo na direção axial apresentou a menor taxa de erro.

Tabela 5. 9 – Desempenho geral do algoritmo ANN-MLP (com seleção de *features*)

Atributos	Vertical	Horizontal	Axial
Acurácia total (%)	85,9	75,3	84,8
Posicionamento do rolo (%)	95,5	87,9	94,5
Falso negativo (%)	9,75	15,3	13,5
Falso positivo (%)	17,8	15,8	15,3

Fonte: O autor

6. CONSIDERAÇÕES FINAIS

De modo geral foi desenvolvido um sistema de diagnóstico de falhas em roletes de transportador de correia com diferentes modos de falha e graus de severidade, a partir de diferentes técnicas de seleção de *features* e aprendizado de máquina combinado a manutenção preditiva de vibração. Além de diagnosticar o estado, o modo de falha e o grau de severidade, também foi identificado a localização do rolo defeituoso no cavalete.

Aplicou-se ANOVA e árvores de decisão como técnica para seleção de *features*, para melhor entendimento das bandas de energia *wavelet* mais significativas para cada direção do sinal de vibração. É notável que houve similaridade em ambas técnicas para a identificação das *features* que possibilitavam um maior aprendizado dos modelos de classificação. No entanto, apesar da árvore de decisão ser um método prático de seleção de *features*, há limitações devido ao ajuste de hiperparâmetros da técnica de aprendizado, sendo necessário uma análise de otimização dos modelos gerados. Além disso, o modelo não exibe a significância de todas as *features*, limitando-se apenas ao ranqueamento das *features* aplicadas nos nós da árvore. Por outro lado, a ANOVA avalia todas as *features*, por meio do F estatístico e o valor-*p*.

Também é válido ressaltar a análise comparativa das três direções, sobre os modelos de aprendizado. Foi possível observar que apesar dos modelos na direção vertical apresentarem um desempenho geral superior, as demais direções também tiveram destaque em outras características como a taxa de acerto do posicionamento do rolo defeituoso, ou em casos de falso negativo e falso positivo. A identificação do posicionamento dos rolos pode reduzir os custos de instalação de mais de um acelerômetro por rolete. Enquanto os casos de falso positivo e negativo é imprescindível para a otimizar o tempo de manutenção do rolete, em inspeções manuais e preventivas.

Sobre o comparativo entre as técnicas de aprendizado de máquinas com e sem a seleção de *features*, corroborou-se que a eliminação das *features* mais ruidosas otimiza o processo de aprendizado e aumenta a acurácia dos modelos de classificação. Também é válido ressaltar a importância do ajuste de hiperparâmetros em cada técnica de aprendizado para a seleção dos modelos de maior acuracidade, além de evitar o *overfitting*. Em relação a SVM, houve melhoria na acuracidade dos

modelos com a seleção de *features*. Já na técnica kNN, foi o que apresentou maior aumento de acuracidade com a seleção de *features*. O ANN-MLP manteve o padrão das demais técnicas de classificação e apresentou melhorias na acurácia com a seleção de *features*.

De modo geral, os modelos SVM apresentaram diagnósticos mais assertivos, principalmente, com a seleção de *features*, ao qual o melhor modelo foi na direção vertical com uma taxa de acerto de 86,7%. Já o melhor modelo criado pela técnica kNN, apresentou acurácia de 80,0% com a seleção de *features*, também na direção vertical. Por fim, o ANN-MLP apresentou o melhor modelo de diagnóstico de falhas com acurácia de 85,9%, na posição vertical com a seleção de *features*. Conseqüentemente, o melhor modelo de diagnóstico de falhas em roletes de transportador de correia, abordado no trabalho, foi o sistema de extração das bandas de energia *wavelet* normalizada dos sinais de vibração na direção vertical, na qual seleciona-se as melhores bandas por ANOVA para criação de um conjunto de dados de treino para um modelo de classificação SVM kernel RBF e parâmetro $C = 100$.

Dentre os fatores que geram oportunidades para melhorias futuras, pode-se destacar a inclusão do diagnóstico em rolos posicionados na região central do cavalete, proporcionando um diagnóstico mais completo do rolete. Além disso, a implementação de carga na bancada do transportador, resultará na coleta de sinais de vibração mais próximas dos sinais coletados em campo. Também se recomenda um estudo mais detalhado acerca da criação de falhas como desgastes superficiais na concha e, principalmente, de defeitos nos rolamentos do rolo, a partir de técnicas não-tradicionais de usinagem para lidar com rolamentos de dimensões menores.

Em relação a extração de *features*, é válido o teste de outras técnicas de análise de vibração no domínio de tempo-frequência, além de técnicas de processamento de imagens para a formação de uma base de dados aplicadas como variáveis de entrada para os modelos de aprendizado. Sugere-se também a aplicação da curva ROC para maior qualidade na análise do diagnóstico. Por fim, a aplicação de técnicas de redes neurais mais complexas, a partir do *deep learning*, afim de avaliar a eficácia dessas técnicas em sinais de vibração em transportador de correia.

REFERÊNCIAS BIBLIOGRÁFICAS

ABDI, Hervé, WILLIAMS, Lynne J, Principal component analysis, Wiley interdisciplinary reviews: computational statistics, v. 2, n. 4, p. 433-459, 2010.

AL-BADOUR, Fadi, SUNAR, Mehmet, CHEDED, Lahouari, Vibration analysis of rotating machinery using time–frequency analysis and wavelet techniques, **Mechanical Systems and Signal Processing**, v. 25, n. 6, p. 2083-2101, 2011.

ALHARBI, Fahad et al, A brief review of acoustic and vibration signal-based fault detection for belt conveyor idlers using machine learning models, **Sensors**, v. 23, n. 4, p. 1902, 2023.

ALSYOUF, Imad, The role of maintenance in improving companies' productivity and profitability, **International Journal of production economics**, v. 105, n. 1, p. 70-78, 2007.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, ABNT, NBR 5462: 1994, **Confiabilidade e Mantabilidade**, Rio de Janeiro: ABNT, 1994.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, ABNT, NBR 6591: 2008, **Tubos de aço-carbono com solda longitudinal de seção circular, quadrada, retangular e especial para fins industriais - Especificação**, Rio de Janeiro: ABNT, 2008.

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, ABNT, NBR 6678: 2017, **Transportadores contínuos - Transportadores de correias - Roletes - Projeto, seleção e padronização**, Rio de Janeiro: ABNT, 2017.

BASU, Jayanta Kumar, BHATTACHARYYA, Debnath, KIM, Tai-hoon. **Use of artificial neural network in pattern recognition**. International journal of software engineering and its applications, v. 4, n. 2, 2010.

BATISTA, G.E.A.P.A., SILVA, D.F., How k-nearest neighbor parameters affect its performance, **In: Argentine symposium on artificial intelligence**, 2009, p. 1-12.

BILLINGS, Steve A., LEE, Kian L, Nonlinear Fisher discriminant analysis using a minimum squared error cost function and the orthogonal least squares algorithm, **Neural networks**, v. 15, n. 2, p. 263-270, 2002.

BISHOP. Christopher M., NASRABADI, Nasser M, **Pattern recognition and machine learning**, New York: Springer, 2006.

BORTNOWSKI, Piotr, KRÓL, Robert, OZDOBA, Maksymilian. Roller damage detection method based on the measurement of transverse vibrations of the conveyor belt, **Eksplatacja i Niezawodność**, v. 24, n. 3, 2022.

BOSWELL, Dustin. Introduction to support vector machines, **Department of Computer Science and Engineering, University of California San Diego**, v. 11, 2002.

BRUNTON. Steven L., KUTZ, J, Nathan. **Data-driven science and engineering: Machine learning, dynamical systems, and control**, Cambridge University Press, 2019.

CASTILHO J, E, et al, **Introdução ao mundo das wavelets**, Sociedade Brasileira de Matemática Aplicada e Computacional, vol, 62, pp. 144, São Carlos, 2012.

CHEN. Chien-Hsing. FEATURE SELECTION BASED ON COMPACTNESS AND SEPARABILITY: COMPARISON WITH FILTER-BASED METHODS. **Computational Intelligence**, v. 30, n. 3, p. 636-656, 2014.

CORTES, Corinna, VAPNIK, Vladimir, **Support-vector networks**, Machine learning, v. 20, p. 273-297, 1995.

CUNNINGHAM, Pdraig, DELANY, Sarah Jane, k-Nearest neighbour classifiers-A Tutorial, **ACM computing surveys (CSUR)**, v. 54, n. 6, p. 1-25, 2021.

CYRINO, L, Análise de vibração - método de Preditiva, **Manutenção em foco**, 08 abr, 2015, Disponível em: <<https://www.manutencaoemfoco.com.br/analise-de-vibracao/>>, Acesso em: 18 maio 2021.

DA SILVA, Paulo César Linhares, **Aplicação das wavelets de Daubechies em conjunto com o método de propagação vetorial de feixes na análise de estruturas fotônicas**, Editora Dialética, 2022.

DA SILVA, Roger R, et al, Fault diagnosis in rotating machine using full spectrum of vibration and fuzzy logic, **J. Eng. Sci. Technol.**, v. 12, n. 11, p. 2952-2964, 2017.

DAUBECHIES, I., **Orthonormal Bases of Compactly Supported Wavelets**, Communications on Pure and Applied Mathematics vol, 41, pp. 909–996, New Jersey, 1988.

DAUBECHIES, Ingrid, Orthonormal bases of compactly supported wavelets, **Communications on pure and applied mathematics**, v. 41, n. 7, p. 909-996, 1988.

DAUBECHIES, Ingrid, **Ten lectures on wavelets**, Society for industrial and applied mathematics, 1992.

DE JONGE, Bram, SCARF, Philip A, A review on maintenance optimization. **European journal of operational research**, v. 285, n. 3, p. 805-824, 2020.

DE SILVA, Clarence W, **Vibration: fundamentals and practice**, CRC press, 2006.

DE VILLE, Barry, **Decision trees**, Wiley Interdisciplinary Reviews: Computational Statistics, v. 5, n. 6, p. 448-455, 2013.

DINIZ, Hélio Augusto Goulart et al, Análise dos modos e efeitos das falhas: FMEA em um transportador de correia, **CONEM**, 2016.

DOMINGUES, Margarete Oliveira, KAIBARA, Magda Kimico, Wavelets biortogonais, **Revista Brasileira de Ensino de Física**, v. 34, p. 1-16, 2012.

DOMNICK, B, Belt conveyor idler life, factors and conditions, **Technical Paper, Superior Industries**, p. 5-8, 2002.

EFEOĞLU, Ebru, TUNA, Gurkan. Machine Learning for Predictive Maintenance: Support Vector Machines and Different Kernel Functions, **Journal of Machinery Manufacture and Reliability**, v. 51, n. 5, p. 447-456, 2022.

EMERSON. Process Management, **Análise de Vibrações Nível I**, Apostila do Curso de Análise de Vibrações da CSI means Reliability, 2006.

FENG, Jianli, LU, Shengnan. Performance analysis of various activation functions in artificial neural networks, **In: Journal of physics: conference series, IOP Publishing**, p. 022030, 2019.

FENG, Zhipeng, LIANG, Ming, CHU, Fulei, Recent advances in time–frequency analysis methods for machinery fault diagnosis: A review with application examples, **Mechanical systems and signal Processing**, v. 38, n. 1, p. 165-205, 2013.

FERNANDEZ, Marta et al, Early detection and fighting of fires in belt conveyor, **Research Fund for Coal and Steel**, 2013.

GÉRON. A, **Hands-On Machine Learning with Scikit-Learn and TensorFlow Concepts, Tools, and Techniques to Build Intelligent Systems**, 2019.

GOURIVEAU, Rafael, MEDJAHAR, Kamal, ZERHOUNI, Noureddine, **From prognostics and health systems management to predictive maintenance 1: Monitoring and prognostics**, John Wiley & Sons, 2016.

GURJAR, Raghvendra Singh, Failure analysis of belt conveyor system, **Int. J. Eng. Soc. Sci.** v. 2, p. 11-23, 2012.

GURJAR, Raghvendra Singh, Failure analysis of belt conveyor system by pareto chart, **Int. J. Eng. Soc. Sci.**, v. 2, p. 60-71, 2012.

IZBICKI, R, & SANTOS, T, M, **Aprendizado de Máquina: Uma abordagem estatística**, Machine Learning, São Carlos, 2020.

JAIN. Anil K., MAO, Jianchang, MOHIUDDIN. K, Moidin. Artificial neural networks: A tutorial, **Computer**, v. 29, n. 3, p. 31-44, 1996.

JANIESCH, Christian. ZSCHECH, Patrick, HEINRICH, Kai, Machine learning and deep learning, **Electronic Markets**, v. 31, n. 3, p. 685-695, 2021.

JANIKOW, Cezary Z, **Fuzzy decision trees: issues and methods**, IEEE Transactions on Systems, Man. and Cybernetics, Part B (Cybernetics), v. 28, n. 1, p. 1-14, 1998.

JJO, Bahzad, ABDULAZEEZ, Adnan. **Classification based on decision tree algorithm for machine learning**, Journal of Applied Science and Technology Trends, v. 2, n. 01, p. 20-28, 2021.

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

KARAMIZADEH, Sasan et al, **An overview of principal component analysis**, Journal of Signal and Information Processing, v. 4, n. 3B, p. 173, 2013.

KARDEC, Alan. NASCIF, Júlio, **Manutenção: função estratégica**, 3ª ed., Rio de Janeiro: Qualitymark: Petrobras, 2009.

KUMAR, Devendra, MANDLOI, R. K., Analysis & prospects of modification in belt conveyors, **International Journal of Engineering Research and Applications (IJERA)**, v. 3, n. 1, p. 581-587, 2013.

LACEY, Steve, The role of vibration monitoring in predictive maintenance. **Asset Management & Maintenance Journal**, v. 24, n. 1, p. 42-51, 2011.

LI, Wei et al, Design of online monitoring and fault diagnosis system for belt conveyors based on wavelet packet decomposition and support vector machine, **Advances in Mechanical Engineering**, v. 5, p. 797183, 2013.

LIPSCHUTZ, S, & LIPSON. M, **Álgebra linear**, Porto Alegre, Bookman. 2011.

LIU, Xiangwei et al, Acoustic signal based fault detection on belt conveyor idlers using machine learning, **Advanced Powder Technology**, v. 31, n. 7, p. 2689-2698, 2020.

LIU, Xiangwei et al, Integrated decision making for predictive maintenance of belt conveyor systems, **Reliability Engineering & System Safety**, v. 188, p. 347-351, 2019.

LIU, Xiangwei, **Prediction of belt conveyor idler performance**, TRAIL Research School, 2016.

LIU, Yi et al, Research on the fault analysis method of belt conveyor idlers based on sound and thermal infrared image features, **Measurement**, v. 186, p. 110177, 2021.

LOBATO, Thiago HG et al, An integrated approach to rotating machinery fault diagnosis using, EEMD, SVM, and augmented data, **Journal of Vibration Engineering & Technologies**, v. 8, p. 403-408, 2020.

LODEWIJKS, Gabriel, **Two decades dynamics of belt conveyor systems**, Bulk Solids Handling, v. 22, n. 2, p. 124-132, 2002.

LUXHØJ, James T., RIIS, Jens O., THORSTEINSSON. Uffe, Trends and perspectives in industrial maintenance management, **Journal of manufacturing systems**, v. 16, n. 6, p. 437-453, 1997.

MAHESH, Batta, Machine learning algorithms-a review, **International Journal of Science and Research (IJSR)**, [Internet], v. 9, n. 1, p. 381-386, 2020.

MALLAT, Stephane G, A theory for multiresolution signal decomposition: the wavelet representation. **IEEE transactions on pattern analysis and machine intelligence**, v. 11, n. 7, p. 674-693, 1989.

MALLAT, Stephane G, Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$, **Transactions of the American mathematical society**, v. 315, n. 1, p. 69-87, 1989.

MORALES, Anibal S, et al, A technology review of idler condition based monitoring systems for critical overland conveyors in open-pit mining applications, In: **2017 IEEE industry applications society annual meeting**, IEEE, 2017.

MUÑIZ, Rubén et al, Real-time monitoring solution with vibration analysis for industry 4.0 ventilation systems, **The Journal of Supercomputing**, v. 79, n. 6, p. 6203-6227, 2023.

MURALIDHARAN. V., RAVIKUMAR, S., KANGASABAPATHY, H, Condition monitoring of Self aligning carrying idler (SAI) in belt-conveyor system using statistical features and decision tree algorithm, **Measurement**, v. 58, p. 274-279, 2014.

NEPOMUCENO, Lauro Xavier, **Técnicas de manutenção preditiva**, Volume 2, São Paulo: Edgard Blucher, 1989.

OBAID, Hadeel S., DHEYAB, Saad Ahmed, SABRY, Sana Sabah, The impact of data pre-processing techniques and dimensionality reduction on the accuracy of machine learning, In: **2019 9th Annual Information Technology, Electromechanical Engineering and Microelectronics Conference (IEMECON)**, IEEE, p. 279-283, 2019.

OCAK, Hasan. LOPARO, Kenneth A., DISCENZO, Fred M, Online tracking of bearing wear using wavelet packet decomposition and probabilistic modeling: A method for bearing prognostics, **Journal of sound and vibration**. v. 302, n. 4-5, p. 951-961, 2007.

ÖZFIRAT, Pınar Mızrak et al, Risk evaluation of belt conveyor accidents using failure modes and effects analysis and event tree analysis, **ITEGAM-JETIA**, v. 8, n. 36, p. 24-31, 2022.

PARK, Cheong Hee, PARK, Haesun. A comparison of generalized linear discriminant analysis algorithms, **Pattern Recognition**. v. 41, n. 3, p. 1083-1097, 2008.

PARK, Cheong Hee, PARK, Haesun. Nonlinear discriminant analysis using kernel functions and the generalized singular value decomposition. **SIAM journal on matrix analysis and applications**, v. 27, n. 1, p. 87-102, 2005.

PEÑA, Mario et al. ANOVA and cluster distance based contributions for feature empirical analysis to fault diagnosis in rotating machinery. In: **2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC)**. IEEE, 2017. p. 69-74.

PEÑA, Mario et al. Feature engineering based on ANOVA, cluster validity assessment and KNN for fault diagnosis in bearings. **Journal of Intelligent & Fuzzy Systems**, v. 34, n. 6, p. 3451-3462, 2018.

PENG, Chen et al, An audio-based intelligent fault diagnosis method for belt conveyor rollers in sand carrier, **Control Engineering Practice**, v. 105, p. 104650, 2020.

PENG, Z, K, et al, Time–frequency data fusion technique with application to vibration signal analysis, **Mechanical Systems and Signal Processing**, v. 29, p. 164-173, 2012.

POPESCU, Theodor D., AIORDACHIOAIE, Dorel, CULEA-FLORESCU, Anisia, Basic tools for vibration analysis with applications to predictive maintenance of rotating machines: an overview, **The International Journal of Advanced Manufacturing Technology**, p. 1-17, 2022.

PRÜFTECHNIK, **Vibration Handbook: Fundamentals – Transducers – Measurement methods**, 1992.

RAO, Singiresu, **Vibrações Mecânicas**, 4ed., Pearson Prentice Hall, São Paulo, 2009.

RAUBER, Thomas W., DE ASSIS BOLDT, Francisco, VAREJAO, Flavio Miguel, Heterogeneous feature models and feature selection applied to bearing fault diagnosis, **IEEE Transactions on Industrial Electronics**, v. 62, n. 1, p. 637-646, 2014.

RAVIKUMAR, S., KANAGASABAPATHY, H., MURALIDHARAN. V. **Fault diagnosis of self-aligning troughing rollers in belt conveyor system using k-star algorithm**, Measurement, v. 133, p. 341-349, 2019.

RAVIKUMAR, S., KANAGASABAPATHY, H., MURALIDHARAN. V. Multi-component fault diagnosis of Self Aligning Troughing Roller (SATR) in belt conveyor system using decision tree: A statistical approach, **FME Transactions**, v. 48, n. 2, p. 364-371, 2020.

REDDY, G, Thippa et al, Analysis of dimensionality reduction techniques on big data, **IEEE Access**, v. 8, p. 54776-54788, 2020.

REICKS, Allen V. Belt conveyor idler roll behaviors, **Bulk material handling by conveyor belt**, v. 7, p. 35-40, 2008.

ROCHA, Filipe et al, Rosi: A robotic system for harsh outdoor industrial inspection-system design and applications, **Journal of Intelligent & Robotic Systems**, v. 103, p. 1-22, 2021.

ROOS, Willem Abraham, HEYNS, Philippus Stephan. In-belt vibration monitoring of conveyor belt idler bearings by using wavelet package decomposition and artificial intelligence, **International Journal of Mining and Mineral Engineering**, v. 12, n. 1, p. 48-66, 2021.

ROUILLARD, Vincent. Quantifying the non-stationarity of vehicle vibrations with the run test. **Packaging Technology and Science**, v. 27, n. 3, p. 203-219, 2014.

ROWE, Alistair CH, ABBOTT, Paul C, Daubechies wavelets and mathematica, **Computers in Physics**, v. 9, n. 6, p. 635-648, 1995.

RUBACK, Livia; AVILA, Sandra; CANTERO, Lucia. Vieses no aprendizado de máquina e suas implicações sociais: Um estudo de caso no reconhecimento facial. In: **Anais do II Workshop sobre as Implicações da Computação na Sociedade**. SBC, 2021. p. 90-101.

RUSCHEL, Edson. SANTOS, Eduardo Alves Portela, LOURES, Eduardo de Freitas Rocha, Industrial maintenance decision-making: A systematic literature review, **Journal of Manufacturing Systems**, v. 45, p. 180-194, 2017.

SAHA, Dip Kumar, AHMED, Sultan. SHAUROV. Md Shariare, Different Machine Maintenance Techniques of Rotary Machine and Their Future Scopes: A Review, In: **2019 4th International Conference on Electrical Information and Communication Technology (EICT)**, IEEE, p. 1-6, 2019.

SHARMA, Sagar, SHARMA, Simone, ATHAIYA, Anidhya, Activation functions in neural networks, **Towards Data Sci.**, v. 6, n. 12, p. 310-316, 2017.

SWINDERMAN. R, Todd, MARTI, Andrew D., MARSHALL, Daniel, **Foundations for Conveyor Safety: The Global Best Practices Resource for Safer Bulk Material Handling**, Martin Engineering Company, 2016.

SYGOUNI, Varvara, TSAKIROGLOU, Christos D., PAYATAKES, Alkiviades C, Capillary pressure spectrometry: Toward a new method for the measurement of the fractional wettability of porous media, **Physics of Fluids**, v. 18, n. 5, 2006.

TANDON. Naresh, CHOUDHURY, Achintya, A review of vibration and acoustic measurement methods for the detection of defects in rolling element bearings, **Tribology international**, v. 32, n. 8, p. 469-480, 1999.

THARWAT, Alaa et al, Linear discriminant analysis: A detailed tutorial, **AI communications**, v. 30, n. 2, p. 169-190, 2017.

TORRENCE, Christopher, COMPO, Gilbert P. A practical guide to wavelet analysis, **Bulletin of the American Meteorological society**, v. 79, n. 1, p. 61-78, 1998.

VAPNIK, Vladimir, The nature of statistical learning theory, **Springer science & business media**, 1999.

VASIĆ, Milan. STOJANOVIĆ, Blaža, BLAGOJEVIĆ, Mirko, **Failure analysis of idler roller bearings in belt conveyors**, Engineering Failure Analysis, v. 117, p. 104898, 2020.

WANG, Danshi et al, **Nonlinearity mitigation using a machine learning detector based on k-nearest neighbors**, IEEE Photonics Technology Letters, v. 28, n. 19, p. 2102-2105, 2016.

WANG, Xia et al, Fault diagnosis of diesel engine based on adaptive wavelet packets and EEMD-fractal dimension. **Mechanical Systems and Signal Processing**, v. 41, n. 1-2, p. 581-597, 2013.

WEBER, Abilio José, *et al*, **Telecurso: profissionalizante de Mecânica: manutenção**, Rio de Janeiro: Fundação Roberto Marinho, 2009.

WENSKOVITCH, John et al, Towards a systematic combination of dimension reduction and clustering in visual analytics, **IEEE transactions on visualization and computer graphics**, v. 24, n. 1, p. 131-141, 2017.

WIJAYA, Hendrik et al, Automatic fault detection system for mining conveyor using distributed acoustic sensor, **Measurement**, v. 187, p. 110330, 2022.

WIJAYA, Hendrik et al, Distributed optical fibre sensor for condition monitoring of mining conveyor using wavelet transform and artificial neural network, **Structural Control and Health Monitoring**, v. 28, n. 11, p. e2827, 2021.

XANTHOPOULOS, Petros et al, Linear discriminant analysis, **Robust data mining**, p. 27-33, 2013.

XIA, Liqiao et al, Toward cognitive predictive maintenance: A survey of graph-based approaches, **Journal of Manufacturing Systems**, v. 64, p. 107-120, 2022.

XIANG, Ling, HU, Aijun. Comparison of Methods for Different Time-frequency Analysis of Vibration Signal, **J, Softw.**, v. 7, n. 1, p. 68-74, 2012.

YANG, Mingjin. ZHOU, Wenju, SONG, Tianxiang, Audio-based fault diagnosis for belt conveyor rollers, **Neurocomputing**, v. 397, p. 447-456, 2020.

YEN. Gary G., LIN, K. C., Wavelet packet feature extraction for vibration monitoring, **IEEE transactions on industrial electronics**, v. 47, n. 3, p. 650-667, 2000.

ZAMAN, Qamruz et al, **Fundamentals of classification and regression trees**, Bio Technology, v. 40, p. 5407-5414, 2011.

ZHANG, Mengchao et al, Deep learning-based damage detection of mining conveyor belt, **Measurement**, v. 175, p. 109130, 2021.

ZHANG, Weiting, YANG, Dong, WANG, Hongchao, Data-driven methods for predictive maintenance of industrial equipment: A survey, **IEEE systems journal**, v. 13, n. 3, p. 2213-2227, 2019.

ZIMROZ, Radosław, KRÓL, Robert, Failure analysis of belt conveyor systems for condition monitoring purposes, **Mining Science**, v. 128, n. 36, p. 255, 2009.