

**UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ**

**FELIPE GUBAUA PORTES**

**GUSTAVO BOUFLEUR**

**VINICIUS SMAHA**

**ANÁLISE VIBROARTROGRÁFICA E MACHINE LEARNING NA  
IDENTIFICAÇÃO DE LESÕES ARTICULARES VIA SÉRIE DE  
FOURIER MULTIDIMENSIONAL**

**CURITIBA**

**2025**

**FELIPE GUBAUA PORTES**

**GUSTAVO BOUFLEUR**

**VINICIUS SMAHA**

**ANÁLISE VIBROARTROGRÁFICA E MACHINE LEARNING NA  
IDENTIFICAÇÃO DE LESÕES ARTICULARES VIA SÉRIE DE  
FOURIER MULTIDIMENSIONAL**

**Vibroarthrographic Analysis and Machine Learning in the  
Identification of Joint Injuries via Multidimensional Fourier Series**

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Controle e Automação do curso de Engenharia de Controle e Automação da Universidade Tecnológica Federal do Paraná (UTFPR).

Orientador: Prof. Dr. Elder Oroski

Coorientador: Prof. Dr. Narco Afonso Ravazzoli Maciejewski

**CURITIBA**

**2025**



[4.0 Internacional](https://creativecommons.org/licenses/by-nc-nd/4.0/)

Esta licença permite *download* e compartilhamento do trabalho desde que sejam atribuídos créditos ao(s) autor(es), sem a possibilidade de alterá-lo ou utilizá-lo para fins comerciais.

Conteúdos elaborados por terceiros, citados e referenciados nesta obra não são cobertos pela licença.

**FELIPE GUBAUA PORTES**

**GUSTAVO BOUFLEUR**

**VINICIUS SMAHA**

**ANÁLISE VIBROARTROGRÁFICA E MACHINE LEARNING NA  
IDENTIFICAÇÃO DE LESÕES ARTICULARES VIA SÉRIE DE  
FOURIER MULTIDIMENSIONAL**

Trabalho de Conclusão de Curso de Graduação apresentado como requisito para obtenção do título de Bacharel em Engenharia de Controle e Automação do curso de Engenharia de Controle e Automação da Universidade Tecnológica Federal do Paraná (UTFPR).

Data de aprovação: 01/dezembro/2025

---

Elder Oroski

Doutorado

Universidade Tecnológica Federal do Paraná

---

Narco Afonso Ravazzoli Maciejewski

Doutorado

Universidade Tecnológica Federal do Paraná

---

Mariana Antonia Aguiar Furucho

Doutorado

Universidade Tecnológica Federal do Paraná

---

Renata Coelho Borges

Doutorado

Universidade Tecnológica Federal do Paraná

**CURITIBA**

**2025**

## **AGRADECIMENTOS**

Gostaríamos de expressar nossa sincera gratidão a todos que, direta ou indiretamente, contribuíram para a realização deste trabalho. Agradecemos às nossas famílias, pelo apoio, incentivo contínuo e compreensão frente a todos os desafios enfrentados durante a graduação. Este suporte foi essencial para o sucesso de nossa trajetória acadêmica.

Estendemos nossos agradecimentos ao nosso orientador e ao nosso coorientador, além de todos os professores que, ao longo da formação, compartilharam não apenas seu conhecimento técnico, mas também valores éticos e profissionais que foram fundamentais para nossa evolução. Sua dedicação e compromisso com o ensino deixaram marcas profundas em nossa formação.

À banca avaliadora, agradecemos pela disponibilidade em contribuir com suas análises criteriosas, sugestões valiosas e observações construtivas, que enriqueceram significativamente este trabalho. Também registramos nossa gratidão à UTFPR - CT e a todos os colegas de curso, pela convivência, pelas trocas de experiências e pelo aprendizado coletivo que construímos ao longo desta jornada. A todos, o nosso muito obrigado.

## RESUMO

O aumento da longevidade da população e a maior incidência de problemas articulares têm ampliado a necessidade por métodos diagnósticos mais precisos, acessíveis e menos invasivos. Entre as diversas articulações do corpo humano, o joelho destaca-se por sua complexidade biomecânica e elevada suscetibilidade a lesões decorrentes de sobrecarga, envelhecimento ou patologias degenerativas. Essas condições representam uma das principais causas de limitação funcional e perda de qualidade de vida, especialmente entre idosos e atletas. Diante desse cenário, torna-se essencial o desenvolvimento de técnicas capazes de auxiliar na detecção de alterações articulares. A vibroartrografia (VAG) surge como uma alternativa promissora, fundamentada na análise das vibrações resultantes das interações mecânicas entre as superfícies articulares em movimento, que geram sinais vibroartrográficos, contendo informações sobre o estado funcional da articulação. Neste contexto, o presente trabalho propõe uma abordagem computacional para a detecção de anomalias articulares a partir de sinais vibroartrográficos adquiridos por sensores inerciais, como acelerômetros e giroscópios. A metodologia proposta emprega a Série de Fourier Multidimensional como principal ferramenta matemática para a decomposição e reconstrução dos sinais em múltiplas dimensões. Esse processo permite representar de forma mais abrangente as características dos sinais, possibilitando a extração de coeficientes relevantes. Então, esses coeficientes são utilizados como variáveis de entrada em classificadores baseados em modelos de aprendizado de máquina, como Máquinas de Vetores de Suporte (SVM), Árvores de Decisão e métodos *Ensemble*, responsáveis por realizar a detecção entre sinais normais e patológicos. Como diferencial, foi desenvolvido um *software* interativo que possibilita visualizar as curvas reconstruídas, comparar padrões e exibir, de forma intuitiva, o resultado dos classificadores. Os testes realizados apresentaram elevada acurácia e consistência entre os modelos, demonstrando que a integração entre vibroartrografia, análise espectral multidimensional e aprendizado de máquina constitui uma alternativa robusta e não invasiva para a detecção de disfunções articulares do joelho.

**Palavras-chave:** Vibroartrografia; Lesão no Joelho; Fourier Multidimensional; Aprendizado de Máquina.

## ABSTRACT

The increase in population longevity and the higher incidence of joint disorders have intensified the need for more precise, accessible, and less invasive diagnostic methods. Among the various joints of the human body, the knee stands out for its biomechanical complexity and high susceptibility to injuries caused by overload, aging, or degenerative pathologies. These conditions represent one of the main causes of functional limitation and loss of quality of life, especially among the elderly and athletes. In this scenario, it becomes essential to develop techniques capable of assisting, in an automated and objective manner, the detection of joint alterations. Vibroarthrography (VAG) emerges as a promising alternative, based on the analysis of vibrations resulting from mechanical interactions between joint surfaces in motion, which generate vibroarthrographic signals containing information about the functional state of the joint. In this context, the present work proposes a computational approach for detecting joint anomalies from vibroarthrographic signals acquired by inertial sensors, such as accelerometers and gyroscopes. The proposed methodology employs the Multidimensional Fourier Series as the main mathematical tool for decomposing and reconstructing the signals across multiple spectral and temporal dimensions. This process enables a more comprehensive representation of signal characteristics, allowing the extraction of relevant coefficients. These coefficients are then used as input variables for classifiers based on machine learning models, such as Support Vector Machines (SVM), Decision Trees, and Ensemble methods, responsible for distinguishing between normal and pathological signals. As a differential, an interactive software was developed to visualize the reconstructed curves, compare patterns, and intuitively display the results of the classifiers. The tests carried out demonstrated high accuracy and consistency among the models, showing that the integration of vibroarthrography, multidimensional spectral analysis, and machine learning constitutes a robust and non-invasive alternative for detection of knee joint dysfunctions.

**Keywords:** *Vibroarthrography; Knee lesion; Multidimensional Fourier; Machine Learning.*

## LISTA DE ILUSTRAÇÕES

<b>Figura 1</b> – Representação formada nos estudos de Lestrel (1989) para a vista lateral de um maxilar humano, incluindo o incisivo central superior, aproximado com um polígono consistindo de 42 pontos. . . . .	22
<b>Figura 2</b> – Estrutura fisiológica do joelho. . . . .	30
<b>Figura 3</b> – Métodos de aprendizado de máquina. . . . .	39
<b>Figura 4</b> – A função booleana $Y = X_1 \wedge X_2 \vee X_3$ e duas árvores de decisão que representam suas soluções. . . . .	42
<b>Figura 5</b> – Classificação de dados pelo modelo SVM em um espaço 2D. . . . .	43
<b>Figura 6</b> – Movimento realizado pelo paciente durante o processo de treinamento e aquisição dos dados. . . . .	50
<b>Figura 7</b> – Tela de coleta de informações e configuração da aquisição da IHM desenvolvida por Silva e Camargo (2025). . . . .	51
<b>Figura 8</b> – Tela de aquisição da IHM desenvolvida por Silva e Camargo (2025), com indicadores de status. . . . .	52
<b>Figura 9</b> – Exemplo de um arquivo “ <i>json</i> ” com valores genéricos. . . . .	53
<b>Figura 10</b> – Exemplo de parte de um arquivo “ <i>csv</i> ” com valores coletados. . . . .	53
<b>Figura 11</b> – Fluxograma que representa as etapas para o processamento dos dados. . . . .	55
<b>Figura 12</b> – Sinais de velocidade angular originais aquisitados pelos giroscópios. . . . .	56
<b>Figura 13</b> – Sinais de aceleração linear originais aquisitados pelos acelerômetros. . . . .	57
<b>Figura 14</b> – Sinal de aceleração do sensor lateral em x após remoção do tempo sem movimento (4s). . . . .	58
<b>Figura 15</b> – Sinal de aceleração do sensor lateral em x após a média dos 3 ciclos. . . . .	59
<b>Figura 16</b> – Sinal de aceleração do sensor lateral em x após a padronização. . . . .	60
<b>Figura 17</b> – Sinal de aceleração do sensor lateral em x após a uniformização. . . . .	61
<b>Figura 18</b> – Sinal de aceleração do sensor lateral em x após o filtro de outliers. . . . .	62
<b>Figura 19</b> – Sinal de aceleração do sensor lateral em x após o filtro de ruídos. . . . .	63
<b>Figura 20</b> – Sinal de aceleração do sensor lateral em x após ajuste e sincronização. . . . .	64
<b>Figura 21</b> – Resumo das 100 execuções para a Árvore de Decisão. . . . .	69
<b>Figura 22</b> – Matriz de confusão com os dados de teste para o modelo de Árvore de Decisão. . . . .	70
<b>Figura 23</b> – Resumo das 100 execuções para o SVM. . . . .	71
<b>Figura 24</b> – Matriz de confusão com os dados de teste para o modelo de SVM. . . . .	72
<b>Figura 25</b> – Resumo das 100 execuções para o Ensemble. . . . .	73
<b>Figura 26</b> – Matriz de confusão com os dados de teste para o modelo de Ensemble. . . . .	74
<b>Figura 27</b> – Fluxo de processamento de classificação. . . . .	75
<b>Figura 28</b> – Estrutura de diretórios do aplicativo. . . . .	77
<b>Figura 29</b> – Diagrama de Casos de Uso do Aplicativo. . . . .	79
<b>Figura 30</b> – Exemplo de reconstrução individual realizado para a componente aceleração do sensor lateral no eixo x, considerando 10 harmônicos. . . . .	81
<b>Figura 31</b> – Exemplo de reconstrução individual realizado para a componente aceleração do sensor lateral no eixo x, considerando 100 harmônicos. . . . .	81
<b>Figura 32</b> – Sinal reconstruído a partir dos coeficientes da Série de Fourier - 10 harmônicos. . . . .	83

<b>Figura 33 – Sinal reconstruído a partir dos coeficientes da Série de Fourier - 100 harmônicos. . . . .</b>	<b>83</b>
<b>Figura 34 – Sinais cartesianos dos sensores e sua reconstrução com <math>h = 10</math>. . . . .</b>	<b>84</b>
<b>Figura 35 – Estrutura de telas do aplicativo. . . . .</b>	<b>88</b>
<b>Figura 36 – Tela inicial do aplicativo. . . . .</b>	<b>89</b>
<b>Figura 37 – Tela de Exames do aplicativo. . . . .</b>	<b>90</b>
<b>Figura 38 – Tela de Detalhes do Exame do aplicativo. . . . .</b>	<b>91</b>
<b>Figura 39 – Tela de Configurações do aplicativo. . . . .</b>	<b>92</b>
<b>Figura 40 – Tela de Resultados do aplicativo. . . . .</b>	<b>94</b>
<b>Figura 41 – Gráficos na Tela de Resultados do aplicativo. . . . .</b>	<b>95</b>
<b>Figura 42 – Tela de Classificação do aplicativo. . . . .</b>	<b>96</b>

## LISTA DE TABELAS

<b>Tabela 1 – Métodos de processamento e classificação para os diferentes artigos encontrados na literatura. . . . .</b>	<b>27</b>
<b>Tabela 2 – Matriz de confusão para classificação de <math>n</math> classes. . . . .</b>	<b>46</b>
<b>Tabela 3 – Comparação do NRMSE entre reconstrução com 100 e 1000 harmônicos - Sinal por posição do sensor e eixo cartesiano. . . . .</b>	<b>82</b>
<b>Tabela 4 – Métricas de Desempenho dos classificadores. . . . .</b>	<b>85</b>

## LISTA DE ABREVIATURAS E SIGLAS

CPU	Unidade Central de Processamento, <i>Central Processing Unit</i>
DAC	Conversor Analógico-Digital, <i>Digital-Analogic Converter</i>
DAF	Análise de Função Discriminante, <i>Discriminative Function Analysis</i>
EFA	Análise Elíptica de Fourier, <i>Elliptical Fourier Analysis</i>
FBSEWT	Transformada Wavelet Empírica Sincronizada Baseada em Banco de Filtros, <i>Filter-Bank-Based Synchrosqueezing Empirical Wavelet Transform</i>
FFT	Transformada Rápida de Fourier, <i>Fast Fourier Transform</i>
FN	Falso Negativo, <i>False Negative</i>
FP	Falso Positivo, <i>False Positive</i>
HFZCs	Cruzamentos por Zero em Alta Frequência, <i>High-Frequency Zero-Crossings</i>
I/O	Entradas e saídas, <i>Inputs/Outputs</i>
KNN	K-Vizinhos Mais Próximos, <i>k-Nearest Neighbors</i>
kbps	Quilo bits por segundo
LCA	Ligamento Cruzado Anterior
LCL	Ligamento Colateral Lateral
LCM	Ligamento Colateral Medial
LCP	Ligamento Cruzado Posterior
LED	Diodo emissor de luz, <i>Light-Emitting Diode</i>
MA	Média Móvel, <i>Moving Average</i>
Mbps	Mega bits por segundo
ML	<i>Machine Learning</i>
MP	<i>Matching Pursuit</i>
QDA	Análise Discriminante Quadrática, <i>Quadratic Discriminant Analysis</i>
RMS	Valor Quadrático Médio, <i>Root Mean Square</i>
RMSE	<i>Root Mean Square Error</i>
SCL	Relógio serial, <i>Serial Clock</i>
SDA	Dados seriais, <i>Serial Data</i>
SVM	Máquina de Vetores de Suporte, <i>Support Vector Machine</i>
TN	Verdadeiro Negativo, <i>True Negative</i>
TP	Verdadeiro Positivo, <i>True Positive</i>
VAG	Vibroartrografia, <i>Vibroarthrography</i>
VI	Curva de tensão por corrente, <i>Voltage-Current</i>
Z-score	Pontuação Padronizada, <i>Standard Score</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>12</b>
<b>1.1</b>	<b>Objetivos . . . . .</b>	<b>16</b>
1.1.1	Objetivo Geral . . . . .	16
1.1.2	Objetivos Específicos . . . . .	16
<b>1.2</b>	<b>Justificativa . . . . .</b>	<b>17</b>
<b>1.3</b>	<b>Estrutura do Documento . . . . .</b>	<b>17</b>
<b>2</b>	<b>REVISÃO DA LITERATURA . . . . .</b>	<b>19</b>
<b>2.1</b>	<b>AQUISIÇÃO DE SINAIS BIOMECÂNICOS . . . . .</b>	<b>19</b>
<b>2.2</b>	<b>TÉCNICAS DE PROCESSAMENTO . . . . .</b>	<b>20</b>
2.2.1	Série de Fourier . . . . .	21
2.2.2	Análise de Fourier Elíptica . . . . .	22
2.2.3	Série de Fourier Multidimensional . . . . .	24
<b>2.3</b>	<b>TÉCNICAS DE CLASSIFICAÇÃO . . . . .</b>	<b>25</b>
<b>3</b>	<b>FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>28</b>
<b>3.1</b>	<b>ESTRUTURA FISIOLÓGICA DO JOELHO . . . . .</b>	<b>28</b>
3.1.1	Ligamentos . . . . .	29
3.1.2	Ligamentos Colaterais . . . . .	29
3.1.3	Meniscos . . . . .	29
3.1.4	Músculos e Tendões . . . . .	29
3.1.5	Patela . . . . .	30
3.1.6	Lesões . . . . .	30
<b>3.2</b>	<b>Pré-Processamento . . . . .</b>	<b>31</b>
3.2.1	Z-score . . . . .	31
3.2.2	Remoção de outliers . . . . .	32
3.2.3	Filtragem e Suavização de Sinais . . . . .	32
3.2.4	Savitzky–Golay . . . . .	33
<b>3.3</b>	<b>PROCESSAMENTO DOS SINAIS . . . . .</b>	<b>33</b>
3.3.1	Série de Fourier Multidimensional . . . . .	34
<b>3.4</b>	<b>Métodos de Classificação . . . . .</b>	<b>38</b>
3.4.1	Árvore de Decisão . . . . .	40
3.4.2	Máquina de Vetores de Suporte (SVM) . . . . .	42
3.4.3	Ensemble Learning . . . . .	44
<b>3.5</b>	<b>Métricas de avaliação dos Classificadores . . . . .</b>	<b>45</b>
3.5.1	Acurácia . . . . .	46
3.5.2	Precisão . . . . .	47
3.5.3	Sensibilidade . . . . .	47
3.5.4	Especificidade . . . . .	48

<b>4</b>	<b>METODOLOGIA</b>	<b>49</b>
<b>4.1</b>	<b>Procedimento de coleta de dados</b>	49
<b>4.2</b>	<b>Características dos Dados Coletados</b>	52
<b>4.3</b>	<b>Organização e Processamento dos Dados</b>	54
4.3.1	Análise Estrutural dos Dados	55
4.3.2	Etapas de Pré-Processamento	57
4.3.2.1	<u>Movimento</u>	57
4.3.2.2	<u>Média dos ciclos</u>	58
4.3.2.3	<u>Padronização</u>	59
4.3.2.4	<u>Uniformização</u>	60
4.3.2.5	<u>Filtro de Outliers</u>	61
4.3.2.6	<u>Filtro de Ruídos</u>	62
4.3.2.7	<u>Ajuste e Sincronização dos Sinais</u>	63
<b>4.4</b>	<b>Aplicação da Série de Fourier e Reconstrução do Sinal</b>	64
4.4.1	Cálculo dos Erros Nas Reconstruções	66
<b>4.5</b>	<b>Desenvolvimento e Treinamento dos Classificadores</b>	66
4.5.1	Algoritmo da Árvore de Decisão	68
4.5.2	Algoritmo SVM	70
4.5.3	Algoritmo Ensemble	72
<b>4.6</b>	<b>Classificação e Geração de Métricas</b>	74
<b>4.7</b>	<b>Aplicativo de Apoio à Análise dos Sinais de Vibroartrografia</b>	75
4.7.1	Arquitetura	76
4.7.2	Integração	78
<b>5</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>80</b>
<b>5.1</b>	<b>Reconstrução das componentes cartesianas</b>	80
<b>5.2</b>	<b>Reconstrução do sinal em duas dimensões</b>	82
<b>5.3</b>	<b>Reconstrução dos sinais com 3 variáveis</b>	84
<b>5.4</b>	<b>Classificadores</b>	84
5.4.1	Limitações dos Dados	85
5.4.2	Análise dos Resultados de Classificação	85
<b>5.5</b>	<b>Aplicativo desenvolvido e suas funcionalidades</b>	87
5.5.1	Tela Inicial	88
5.5.2	Tela de Exames	89
5.5.3	Tela de Detalhes dos Exames	90
5.5.4	Tela de Configurações	91
5.5.5	Tela de Resultados	93
5.5.6	Tela de Classificação	96
<b>6</b>	<b>CONCLUSÃO</b>	<b>98</b>
	<b>REFERÊNCIAS</b>	<b>100</b>

## 1 INTRODUÇÃO

Em um contexto social de aumento da longevidade e maior preocupação com a qualidade de vida, surge a necessidade do desenvolvimento de diagnósticos médicos cada vez mais precisos e menos invasivos ao paciente. Tratando-se da avaliação de lesões articulares, especialmente as que afetam o joelho, há uma demanda bastante expressiva. Somente no ano de 2021, foram realizados 5,2 milhões de exames pela rede pública e mais de 16 mil cirurgias relacionadas a este tipo de lesão (PERCOPE, 2022).

A crescente demanda por diagnósticos contrasta com a limitação de métodos disponíveis, que majoritariamente envolvem procedimentos custosos ou invasivos, como a ressonância magnética e a artroscopia (BACHES *et al.*, 2024). Tendo essas limitações em vista, exames não invasivos ou que não afetam diretamente a região examinada possibilitam ganhos em diversas frentes relacionadas ao diagnóstico e ao prognóstico de lesões articulares. Nesse mesmo contexto, a vibroartrografia surge como uma técnica promissora, sendo menos onerosa e invasiva que os exames de imagem e outros tradicionalmente empregados na prática clínica. Essa metodologia baseia-se na captação e na análise de vibrações geradas pelas articulações durante movimentos (MCCOY *et al.*, 1987).

A grandeza a ser medida a partir das vibrações articulares depende do instrumento de medição utilizado pelo estudo, podendo ser a aceleração linear da vibração (acelerômetros) ou a velocidade angular da vibração (giroscópios), conforme proposto no estudo de Funaki e Lorentz (2022). Historicamente, já foram utilizados microfones e até mesmo estetoscópios como ferramentas para realizar as medições, com análise a partir de ondas sonoras (TOCQUEVILLE *et al.*, 2021).

Inicialmente, de acordo com os estudos de McCoy *et al.* (1987), nas primeiras abordagens da medicina diagnóstica voltadas às articulações, os exames eram realizados com o auxílio de estetoscópios, a fim de detectar variações na pressão sonora gerada pelas vibrações, como uma medida indireta do som gerado pelo contato articular. Contudo, avanços tecnológicos permitiram melhorias significativas na captação das vibrações geradas pelos movimentos articulares. Na década de 1990, microfones foram empregados para a aquisição, embora estes fossem altamente suscetíveis a ruídos externos (TOCQUEVILLE *et al.*, 2021). Mais recentemente, a partir dos anos 2000, o uso de acelerômetros contribuiu para mitigar o problema dos ruídos externos, com dispositivos capazes de captar vibrações de maneira mais precisa e limitada ao ponto de medição

(TOCQUEVILLE *et al.*, 2021).

No âmbito do processamento dos dados aqusitados, diversos estudos utilizam diferentes métodos de processamento e determinação da qualidade dos diagnósticos. McCoy *et al.* (1987) utilizaram a Transformada Rápida de Fourier (FFT) em suas análises. Rangayyan e Wu (2008), por sua vez, realizaram o processamento por meio de funções de densidade de probabilidade – janela de Parzen<sup>1</sup> e Gaussianas<sup>2</sup> – com pré-filtragem passa-faixa (10Hz a 1kHz) e normalização dos dados. Ainda, há estudos como o de Wu e Krishnan (2009) que utilizam o algoritmo de *Matching Pursuit* (MP) para realizar análises no tempo e na frequência, bem como técnicas baseadas em máquinas de vetores de suporte (SVM).

A análise de vibrações articulares mostra-se uma ferramenta promissora para a detecção de lesões no joelho. Para a aplicação do método da vibroartrografia, de forma geral, duas grandes etapas são necessárias: (i) a aquisição dos dados de vibração resultantes do movimento da articulação; e (ii) o processamento desses dados adquiridos, visando à identificação da lesão. Para a realização dessas duas etapas, diferentes métodos foram utilizados no decorrer do tempo com o intuito de otimizar os resultados.

Em um estudo mais recente, Funaki e Lorentz (2022) realizaram uma pesquisa utilizando um microcontrolador Arduino UNO R3, responsável pela comunicação com dois pares de acelerômetros e giroscópios agrupados em um módulo de sensores GY-521. Esse sistema foi empregado para a coleta dos sinais de vibração. As análises indicaram que as vibrações captadas durante a fase de aquisição podem ser representadas por sinais periódicos.

A partir disso, os autores realizaram o processamento dos sinais utilizando a Série de Fourier Bidimensional. Para facilitar o entendimento, essa abordagem pode ser comparada à metodologia empregada na análise de curvas VI na Engenharia Elétrica, em que a relação entre tensão e corrente é representada graficamente sem considerar o eixo temporal como referência direta. De forma semelhante, neste estudo, foram analisadas curvas que relacionam aceleração linear e velocidade angular no mesmo eixo de movimento. Essas curvas foram decompostas utilizando 50 harmônicos na Série de Fourier Bidimensional, a fim de representar as principais características da forma original do sinal, mantendo desempenho computacional otimizado.

Em união a esse processo, os autores empregaram algoritmos de *machine learning* na

<sup>1</sup> A Janela de Parzen é um método estatístico usado para estimar a densidade de probabilidade de um conjunto de dados sem assumir um modelo específico (PARZEN, 1962).

<sup>2</sup> A Janela Gaussiana é uma técnica do processamento de sinais utilizada para minimizar distorções em análises espectrais, aplicando uma função Gaussiana ao sinal antes de realizar transformadas de Fourier (BOASHASH, 1992).

etapa de detecção das lesões. Dessa maneira, por meio de classificadores como árvore de decisão, *Ensemble* e Máquinas de Vetores de Suporte (SVM), que serão conceituados posteriormente, foram classificadas as lesões articulares de maneira automatizada.

Os resultados obtidos demonstraram que essa abordagem permitiu a identificação de lesões, além de reforçar o potencial da vibroartrografia como uma ferramenta no diagnóstico médico, especialmente em cenários que demandam soluções menos invasivas e de menor custo. A melhora na identificação justifica-se pela obtenção de indicadores de performance para a detecção da presença das lesões – como acurácia, especificidade, precisão e sensibilidade –, mais elevados em relação aos obtidos em estudos anteriores como o de McCoy *et al.* (1987) ou o proposto por Tocqueville *et al.* (2021). Como os próprios autores sugerem na pesquisa, essa melhoria significativa percebida pode ser atribuída ao uso de acelerômetros e giroscópios em conjunto, permitindo a análise concomitante da aceleração linear e da velocidade angular, em cada eixo, por meio da Série de Fourier Bidimensional.

Partindo das premissas estabelecidas pelos resultados dos estudos apresentados, busca-se aprimorar a abordagem realizada por Funaki e Lorentz (2022), aplicando a Série de Fourier Multidimensional na análise dos sinais biomecânicos. Ao considerar um número maior de variáveis simultaneamente, objetiva-se obter um diagnóstico mais preciso e confiável das lesões articulares, bem como fundamentar as bases para melhor identificá-las.

Dada a ampla variedade de abordagens possíveis para a realização da aquisição e processamento dos sinais obtidos por meio da técnica de vibroartrografia, surge a necessidade do desenvolvimento de pesquisas que busquem identificar a metodologia com melhor performance na detecção de lesões. Neste estudo, propõe-se a aplicação de técnicas de processamento de sinais e análise espectral para a caracterização, parametrização e interpretação dos sinais de vibração gerados pelo movimento das articulações do joelho. Complementarmente, são utilizadas ferramentas de aprendizado de máquina com o objetivo de identificar e classificar diferentes tipos de lesões articulares.

A principal contribuição do trabalho reside no desenvolvimento de um método computacional que utiliza a Série de Fourier Multidimensional para decomposição e extração de características dos sinais, seguido de uma etapa de classificação por meio de técnicas de aprendizado de máquina. Com isso, espera-se oferecer uma ferramenta amigável ao usuário final, capaz de realizar a identificação das lesões com base nos padrões presentes nos dados obtidos de sensores inerciais.

O processo de identificação de lesões no joelho, por meio da técnica de vibroartrografia, conforme observado por Funaki e Lorentz (2022), pode ser dificultado pela semelhança entre os padrões de movimento gerados por diferentes tipos de lesões. Dessa forma, surge a necessidade de métodos cada vez mais precisos, que permitam distinguir esses sinais de forma adequada, a fim de reduzir a ocorrência de diagnósticos incorretos e evitar impactos negativos no tratamento dos pacientes.

Visando contribuir para essa demanda, este estudo propõe o desenvolvimento de um método baseado na aplicação da Série de Fourier Multidimensional para o processamento e análise dos sinais biomecânicos. A proposta parte da premissa de que os dados já terão sido previamente adquiridos por meio de sensores inerciais, fornecidos pelo Hospital do Trabalhador, utilizando pares de acelerômetros e giroscópios. Esses dispositivos fornecem sinais que podem ser representados em um espaço de estados<sup>3</sup> pertencente ao conjunto dos números reais com 18 dimensões ( $\mathbb{R}^{18}$ ).

Além disso, parte-se do pressuposto de que as amostras fornecidas contêm os parâmetros necessários para o adequado processamento dos sinais. Assim, faz-se possível a aplicação da decomposição por meio da Série de Fourier Multidimensional, visando à obtenção dos coeficientes de Fourier representativos de cada sinal. A partir desses coeficientes, realiza-se a extração de padrões e características relevantes que, posteriormente, são utilizadas como base para a etapa de classificação, empregando-se técnicas de aprendizado de máquina.

Dentre os desafios esperados durante o desenvolvimento do projeto, podem ser destacados:

1. Os dados adquiridos por sensores inerciais podem conter ruídos, interferências e variações devido à execução diferente dos movimentos pelos pacientes. Assim, podem ser necessárias etapas de pré-processamento para a correta análise.
2. Desbalanceamento no conjunto de dados devido a variedade do tipo de lesões analisadas, que pode contrastar com o número de amostras para cada uma delas;
3. Dificuldade na seleção de características (*features*), resultante da complexidade na interpretação das curvas geradas pela decomposição dos sinais utilizando a Série de Fourier Multidimensional.

---

<sup>3</sup> Dimensões dos eixos  $x$ ,  $y$  e  $z$  com as variáveis de aceleração linear (acelerômetro) e velocidade angular (giroscópio) para cada par de instrumentos utilizados.

4. Por conta da baixa quantidade de amostras, dificuldade em garantir que o modelo treinado com o conjunto limitado de dados disponíveis consiga generalizar bem para novos pacientes e diferentes condições clínicas, evitando sobreajuste (*overfitting*).

## 1.1 Objetivos

### 1.1.1 Objetivo Geral

Desenvolver um método para diagnóstico e análise de lesões articulares no joelho, por meio da aplicação de Série de Fourier Multidimensional e aprendizado de máquina, em sinais de vibroartrografia.

### 1.1.2 Objetivos Específicos

Para que o objetivo geral deste trabalho seja atingido, os seguintes objetivos específicos ficam estabelecidos:

1. Analisar, parametrizar e caracterizar os sinais vibroartrográficos utilizando a Série de Fourier Multidimensional;
2. Selecionar as técnicas de aprendizado de máquina que mais se adequam às condições dos dados obtidos;
3. Treinar o modelo de aprendizado de máquina;
4. Analisar a influência da dimensionalidade dos sinais no desempenho dos modelos de classificação;
5. Desenvolver uma interface que integre a seleção, o processamento e a classificação dos dados vibroartrográficos;
6. Identificar a presença de lesões articulares no joelho e classificar os diferentes tipos de lesões por meio das características específicas dos sinais vibroartrográficos;
7. Validar o desempenho dos classificadores desenvolvidos para identificar o modelo mais adequado ao diagnóstico das lesões articulares.

## 1.2 Justificativa

A crescente demanda por diagnósticos de lesões no joelho contrasta com a persistência de métodos de detecção que ainda se limitam devido a altos custos e características invasivas. A ressonância magnética, apesar de ser um procedimento com alto índice de assertividade na detecção da lesão, é um método que exige equipamentos de alta complexidade e mão de obra especializada para sua operação, tornando-se um diagnóstico de alto custo agregado (FILHO *et al.*, 2011). A artroscopia, por sua vez, é considerada o método “padrão-ouro” por apresentar resultados de elevada acurácia e precisão. No entanto, consiste em um procedimento cirúrgico, sendo um método diagnóstico de caráter invasivo. Por este motivo, é usualmente utilizada em casos nos quais os demais procedimentos não obtiveram resultados satisfatórios, como nas lesões condrais<sup>4</sup>, em que há incertezas recorrentes por falsos positivos gerados no diagnóstico por ressonância magnética (FILHO *et al.*, 2011).

A vibroartrografia, como apontado por estudos recentes, possui um grande potencial de diagnóstico de lesões articulares, e sua assertividade está diretamente relacionada ao procedimento metodológico adotado, tanto na etapa de aquisição dos dados quanto no processamento dos dados obtidos. Analisando o avanço das pesquisas realizadas, principalmente na última década, e tendo como base os resultados descritos por Funaki e Lorentz (2022), notam-se taxas de assertividade significativas. Como os próprios autores propõem, isso se deve à aquisição massiva de dados de aceleração angular e velocidade linear durante a execução dos movimentos biomecânicos. Esses fatores, combinados à extração de características, por meio da análise via Série Bidimensional de Fourier, favoreceram a performance dos classificadores baseados em Aprendizado de Máquina.

## 1.3 Estrutura do Documento

O presente trabalho está estruturado de forma a proporcionar uma compreensão dos componentes que permitem a aquisição, processamento e classificação de dados biomecânicos no contexto da vibroartrografia. Inicialmente, no Capítulo 2, é apresentada uma revisão da literatura, na qual são contextualizadas as principais contribuições acadêmicas relacionadas à técnica, incluindo métodos de aquisição validados e os resultados obtidos em estudos anteriores.

<sup>4</sup> Lesões condrais referem-se a danos na cartilagem articular do joelho, causados por fatores metabólicos, genéticos, vasculares ou traumáticos. Essas lesões podem variar desde um desgaste superficial até atingir o osso subcondral (FILHO *et al.*, 2012).

Após esta etapa, no Capítulo 3, é realizada a fundamentação teórica dos principais conceitos relacionados à vibroartrografia e a outros componentes que serão explorados durante a pesquisa, a fim de fornecer o embasamento científico necessário para a confiabilidade das conclusões.

A metodologia, por sua vez, é apresentada no Capítulo 4 e introduz os caminhos percorridos até a obtenção dos resultados, evidenciando todas as etapas realizadas, desde a aquisição dos sinais de vibração pelo movimento das articulações do joelho até o processamento dos dados brutos, os filtros e outras técnicas de manipulação empregadas. Dessa forma, os resultados surgem no Capítulo 5 como consequência de todos os passos descritos anteriormente, demonstrando o comportamento dos sinais após a reconstrução via Série de Fourier e o estágio de desenvolvimento da interface do aplicativo orientado ao usuário final. Por fim, no Capítulo 6, é realizada uma síntese dos conceitos abordados e do avanço obtido com a pesquisa.

## 2 REVISÃO DA LITERATURA

Este Capítulo tem como objetivo revisar os principais avanços e aplicações da vibro-artrografia na análise biomecânica e no diagnóstico de lesões articulares. São abordadas as metodologias empregadas em estudos anteriores, bem como os resultados obtidos, com ênfase na identificação de pontos de interesse e limitações. Essa análise permite compreender o estado da arte e fornecer subsídios teóricos para o desenvolvimento metodológico proposto neste trabalho.

### 2.1 AQUISIÇÃO DE SINAIS BIOMECÂNICOS

A fase inicial da análise do movimento humano compreende a aquisição dos sinais correspondentes às respostas geradas durante o deslocamento articular. Ferramentas de organização e processamento de grandes volumes de dados permitem que sinais provenientes de movimentos biomecânicos sejam coletados e analisados por algoritmos especializados, como evidenciado por Oliveira (2017), que analisou o sincronismo entre remadores através dos sinais gerados durante a execução do movimento. Devido à complexidade e presença de ruídos nas amostras, tal análise torna-se possível quando há confiabilidade na metodologia de aquisição, filtragem e pré-processamento dos dados. Neste sentido, a literatura apresenta diferentes abordagens a depender da natureza do sistema biológico alvo do estudo. Segundo Crenna *et al.* (2021), existem duas maneiras de realizar a extração de sinais de movimentos biomecânicos, que incluem vídeo ou sensores inerciais. No primeiro caso, o resultado consiste na obtenção, em 2D ou 3D, das posições no eixo do tempo. No segundo caso estão os acelerômetros, giroscópios, magnetômetros e outros elementos sensores capazes de medições minuciosas dos movimentos, cujo resultado são ângulos que variam no tempo. Todavia, em ambos há a presença de ruídos que interferem nos resultados das análises.

No estudo realizado por Borzucka *et al.* (2024), analisou-se a resposta das articulações do joelho a diferentes intensidades de carga durante a atividade de agachamento com barra. Os sinais de aceleração foram obtidos acima da região patelar a uma frequência de amostragem de  $10\text{ kHz}$ , passando por um amplificador de condicionamento multicanal antes do processamento digital. Foi então aplicado um filtro Butterworth passa-banda de quarta ordem, a fim de eliminar frequências abaixo de  $50\text{ Hz}$ , provenientes da contração muscular e movimentos do corpo, e acima de  $1000\text{ Hz}$ , uma vez que altas frequências não refletem eventos de caráter biomecânico.

No estudo de Basavaraju *et al.* (2025b), também no âmbito da VAG<sup>1</sup>, os sinais de aceleração foram obtidos na região médio patelar, amostrados em uma taxa de  $2\text{ kHz}$ , digitalizados com 12 bits de resolução e pré-filtrados utilizando um filtro passa-banda analógico nas frequências entre  $10\text{ Hz}$  e  $1\text{ kHz}$ . Após a digitalização, foi aplicado um filtro CMA<sup>2</sup>, buscando aumentar a acurácia dos classificadores. Os acelerômetros aparecem com frequência como mecanismo sensor nos estudos sobre o desenvolvimento da vibroartrografia, mas há também artigos que buscam outras alternativas, como o de Skoczylas *et al.* (2023) que utilizou um microfone de contato piezoelétrico para detecção de vibrações acústicas, posteriormente pré-processadas via EEMD<sup>3</sup>.

## 2.2 TÉCNICAS DE PROCESSAMENTO

Além das técnicas de aquisição, a etapa anterior à aplicação dos algoritmos classificadores é importante para o melhor aproveitamento dos dados obtidos. Uma série de procedimentos pode ser aplicada para que características relevantes possam ser extraídas dos sinais coletados.

Crenna *et al.* (2021) propõem em seu estudo uma análise comparativa entre diferentes técnicas de diferenciação e filtragem de sinais biomecânicos aplicados à análise de movimento humano, com foco na minimização da incerteza de medição. Neste sentido, é realizada a síntese das abordagens mais frequentemente usadas na literatura, sendo elas: (i) Diferenciação numérica seguida de filtragem passa-baixa, (ii) aproximação local polinomial e diferenciação direta e, por fim, (iii) filtragem de Fourier ótima (CRENNA *et al.*, 2021). Dentre os métodos mais explorados no estudo, refere-se ao Filtro de Butterworth com Fase Zero como uma boa alternativa para a filtragem de sinais biomecânicos. Este recurso foi utilizado no campo de estudo da vibroartrografia por Borzucka *et al.* (2024), para selecionar somente as faixas de frequência com influência de movimentos articulares, desprezando frequências relacionadas aos movimentos corporais e à interferência eletrônica. Além disso, o estudo explora também filtros de média móvel (MA<sup>4</sup>) e filtros polinomiais de Savitsky–Golay (CRENNA *et al.*, 2021).

Os trabalhos de Soudan e Dierckx (1979) e Hatze (1981), por outro lado, apresentam abordagens voltadas ao tratamento de dados experimentais obtidos a partir de sinais biomecânicos,

<sup>1</sup> VAG - Abreviação comum na literatura para Vibroartrografia

<sup>2</sup> CMA (*Cascaded Moving Average*) - técnica de filtragem digital utilizada para suavizar sinais e remover componentes de baixa frequência indesejadas.

<sup>3</sup> EEMD (*Ensemble Empirical Mode Decomposition*) - técnica de processamento de sinais não lineares e não estacionários

<sup>4</sup> Sigla para *Moving Average*

utilizando a decomposição por séries de Fourier. No primeiro, o uso de funções *spline* em conjunto com a série permitiu a suavização dos dados e o cálculo dos coeficientes de Fourier mesmo em condições de ruído. No segundo, Hatze (1981) introduziu o uso de séries de Fourier regularizadas para estimar derivadas de ordens superiores, evitando os problemas de amplificação de ruído associados à diferenciação numérica. Partindo destes resultados, serão exploradas as vantagens da utilização desta abordagem no contexto da vibroartrografia. Para isso, é necessário resgatar os alicerces matemáticos que tornam possível a aplicação desta metodologia ao estudo, iniciando pelo próprio conceito de Séries de Fourier.

### 2.2.1 Série de Fourier

A série de Fourier é uma das ferramentas mais clássicas e adequadas para representar sinais periódicos e convergentes no tempo, ou seja, sinais definidos para todo o domínio  $t$  real e que possuem um período de repetição  $p$ , tal que  $f(t + p) = f(t)$ . Sua principal utilidade reside na capacidade de decompor funções complexas no domínio do tempo em uma soma de funções harmônicas — senos, cossenos e uma componente constante, sendo tão precisa quanto o limite do somatório proposto na Equação 1. Essa transformação permite uma análise mais simplificada no domínio da frequência, facilitando a interpretação e o tratamento de sinais que, de outra forma, seriam de difícil manipulação (KREYSZIG, 2009).

Em sua versão mais simplificada, a Série de Fourier pode ser representada por meio da Equação 1, cuja constante  $A_0$  e seus coeficientes  $A_h$  e  $B_h$  são definidos pelas Equações 2, 3 e 4, respectivamente.

$$f(x) = A_0 + \sum_{h=1}^{\infty} \left[ A_h \cos \left( \frac{2\pi h}{T} x \right) + B_h \sin \left( \frac{2\pi h}{T} x \right) \right], \quad (1)$$

$$A_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx, \quad (2)$$

$$A_h = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(h x) dx, \quad (3)$$

$$B_h = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(h x) dx. \quad (4)$$

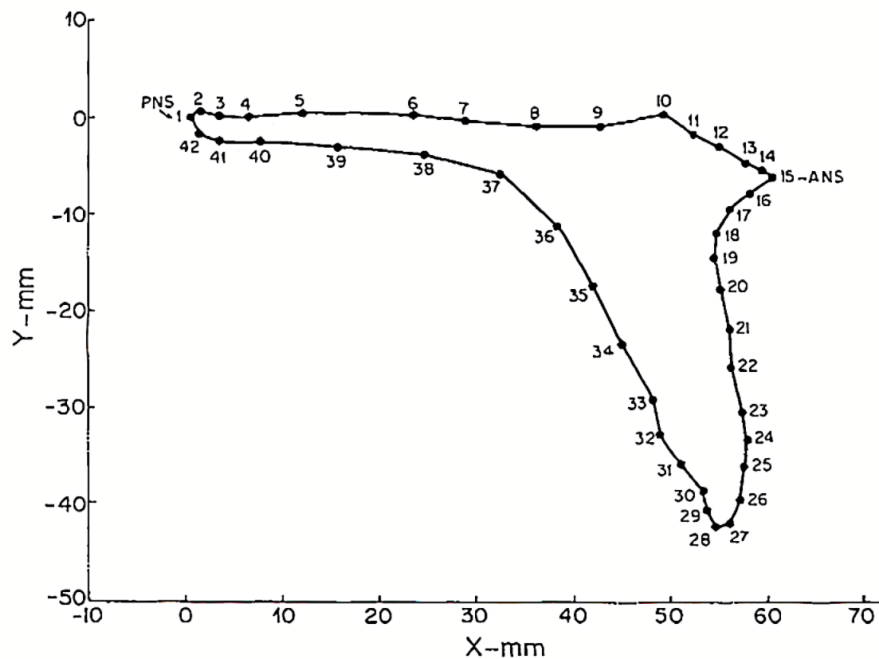
Na Equação 1, o índice  $h \in \mathbb{Z}$  do somatório representa a ordem dos harmônicos utilizados na decomposição, influenciando diretamente a precisão da aproximação da Série de Fourier em relação ao sinal original (KREYSZIG, 2009). Teoricamente, a série pode conter infinitos termos ( $h \rightarrow \infty$ ), o que garante a reconstrução exata da função original sob determinadas

condições. No entanto, como discutido por Funaki e Lorentz (2022), na prática, essa série é truncada em um valor finito de  $h$ . Esse valor é definido de modo que o erro de aproximação resultante seja suficientemente pequeno para não comprometer os resultados da decomposição e da análise subsequente.

### 2.2.2 Análise de Fourier Elíptica

Diante da complexidade inerente à análise de sinais e à necessidade de considerar múltiplas variáveis simultaneamente, Kuhl e Giardina (1982) propuseram em seus estudos a Análise de Fourier Elíptica (EFA)<sup>5</sup>, uma abordagem que ampliou significativamente a capacidade de descrever e analisar formas bidimensionais complexas (LESTREL, 1989). A EFA representa uma solução paramétrica voltada ao estudo de contornos fechados, fundamentada na derivação de um par de equações em função de uma terceira variável. Essa representação pode ser observada por meio de um polígono formado em um gráfico de coordenadas cartesianas  $(x,y)$ , como ilustrado na Figura 1. Assim, a EFA permite uma representação matemática tão precisa quanto à proximidade dos pontos do sinal original (KUHL; GIARDINA, 1982).

**Figura 1 – Representação formada nos estudos de Lestrel (1989) para a vista lateral de um maxilar humano, incluindo o incisivo central superior, aproximado com um polígono consistindo de 42 pontos.**



Fonte: (LESTREL, 1989).

<sup>5</sup> Do inglês *Elliptical Fourier Analysis*, também chamada de Série de Fourier bidimensional.

Os pontos das coordenadas  $x$  e  $y$  do contorno presente na Figura 1 podem ser analisados separadamente ao serem projetados em função de uma nova variável, o tempo  $t$ , desde que a velocidade de deslocamento deles no contorno seja considerada constante (LESTREL, 1989). Ou seja, isso significa representar os valores de  $x$  e  $y$  como variações univariadas ao longo de um "eixo do tempo", permitindo que ambos sejam tratados de forma independente. Com essa abordagem, cada coordenada passa a descrever um sinal convergente e periódico, que pode ser decomposto e analisado por meio do uso da Série de Fourier (KUHL; GIARDINA, 1982).

Dessa forma, ao aplicar a Série de Fourier, conforme expressa na Equação 1, para analisar separadamente as funções das coordenadas  $x(t)$  e  $y(t)$ , Kuhl e Giardina (1982) observaram que os harmônicos resultantes de cada componente assumem o formato de elipses — característica que deu origem ao nome EFA. Quando essas elipses, correspondentes aos diferentes harmônicos, são combinadas ponto a ponto para os mesmos valores de tempo  $t$ , o contorno original tende a ser reconstruído. Esse processo é possível devido à ortogonalidade dos harmônicos, ou seja, a independência entre funções seno e cosseno de diferentes frequências, cujo produto entre elas, quando integrado ao longo de um intervalo, é igual a zero (KUHL; GIARDINA, 1982). Essa propriedade garante que os harmônicos utilizados na Série de Fourier não interfiram entre si e permite a decomposição e reconstrução precisa da forma. Assim, mesmo para o caso da EFA, a precisão da reconstrução do contorno depende diretamente do número de harmônicos utilizados, representado pelo termo  $h$  no somatório da Equação 1, que indica o grau de truncamento da série.

Estudos mais recentes têm utilizado o conceito da EFA, originalmente definido por Kuhl e Giardina (1982) para o processamento de imagens, adaptando-o para diferentes áreas de interesse. Por exemplo, Caple *et al.* (2017) aplicou a EFA em estudos antropológicos de crânios. Já Funaki e Lorentz (2022) e Silva e Camargo (2025), motivados pela complexidade dos sinais de vibração obtidos pela técnica da vibroartrografia e pela necessidade de uma análise conjunta, adaptaram essa técnica do processamento de imagens para o processamento de sinais biomecânicos.

A Série de Fourier elíptica para a análise de sinais é uma extensão natural da versão unidimensional apresentada na Equação 1, sendo aplicada a funções de duas variáveis — por exemplo,  $f(x, y)$  — e representando a decomposição do sinal em termos de funções senoidais e cossenoidais em ambas as direções. Em sua forma geral, a função é expressa como um somatório duplo envolvendo coeficientes harmônicos que capturam a variação conjunta entre os dois

domínios analisados (FUNAKI; LORENTZ, 2022).

### 2.2.3 Série de Fourier Multidimensional

Em muitos casos, um fenômeno pode ser influenciado simultaneamente não apenas por duas, mas sim por três, quatro ou, de modo geral,  $n$  variáveis. Nessas situações, a análise por meio da Série de Fourier elíptica não é suficiente para capturar todas as características das diferentes variáveis influenciando, no mesmo instante de tempo, o sinal original. Objetivando contemplar essa complexidade, existe uma abordagem matemática denominada de Série de Fourier Multidimensional, também conhecida como Somas Parciais Multidimensionais (WEISZ, 2002), a qual permite uma representação mais completa das influências múltiplas e interações entre variáveis.

Essa abordagem, assim como a EFA, também se baseia na extensão da Série de Fourier simples, representada na Equação 1, mas para um espaço de estados de ordem  $n$ . Weisz (2002) apresenta em seus estudos um formato base para a representação trigonométrica  $n$  dimensional do sistema, introduzida pelo que o autor chama de "Produto de Kronecker", apresentado na Equação 5 a seguir:

$$e^{ik \cdot x} = \prod_{j=1}^n e^{ik_j x_j}. \quad (5)$$

Nessa Equação 5,  $k$  é um vetor de inteiros, tal que  $k = (k_1, k_2, \dots, k_n)$ , representando os índices de frequência nas  $n$  dimensões de análise. De forma análoga,  $x$  corresponde a um vetor cujas componentes  $x = (x_1, x_2, \dots, x_n)$  representam o valor da variável analisada para cada uma das  $n$  dimensões. Assim, na Equação 5 o produto escalar desses termos aparece como argumento da exponencial complexa  $(e^{ik_j \cdot x_j})$ , expressando a contribuição de cada frequência  $k_j$  para a fase total da oscilação da função base. Essa fase determina como cada componente de frequência se combina na reconstrução da função original (WEISZ, 2002).

A partir dessa base, a série de Fourier multidimensional é construída pelo autor como uma soma infinita ponderada pelas frequências presentes na função original  $f(x)$ . Os coeficientes da Série de Fourier Multidimensional são calculados utilizando a ortogonalidade das exponenciais complexas no toro<sup>6</sup>  $n$ -dimensional, dada pela Equação 6:

<sup>6</sup> Em geral, o toro  $n$ -dimensional  $(\mathbb{T}^n)$  é um espaço topológico que generaliza a ideia de um círculo para múltiplas dimensões. Ele é representado pelo produto cartesiano de  $n$  círculos  $(S^1)$ , formando o domínio natural para funções periódicas em análise de Fourier.

$$\hat{f}(k) = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x) e^{-ik \cdot x} dx \quad (k \in \mathbb{N}^n). \quad (6)$$

Na Equação 6 acima,  $\hat{f}(k)$  representa os coeficientes de Fourier associados ao vetor de frequências  $k$ .  $\mathbb{T}^n$  representa o toro  $n$ -dimensional e, o termo multiplicador  $\frac{1}{(2\pi)^n}$  é um fator de normalização necessário para manter a ortonormalidade da base (WEISZ, 2002). Este fator surge da integral de ortogonalidade no espaço de funções periódicas em cada dimensão, garantindo que a base formada pelas funções exponenciais seja ortonormal. Isso significa que essas funções são mutuamente ortogonais e têm norma bem definida, o que é essencial para a correta decomposição da função original na Série de Fourier Multidimensional (WEISZ, 2002).

Finalmente, o autor chega à série trigonométrica formal que define a representação da Série de Fourier Multidimensional para uma função, conforme a Equação 7 a seguir:

$$f(x) = \sum_{k \in \mathbb{Z}^n} \hat{f}(k) e^{ik \cdot x} \quad (x \in \mathbb{T}^n). \quad (7)$$

Na Equação 7 acima, que expressa a soma formal da série de Fourier multidimensional, o termo  $\mathbb{Z}^n$  representa o conjunto de todos os vetores de inteiros de  $n$  dimensões, assim garantindo que todas as frequências possíveis sejam consideradas na reconstrução da função original. Portanto, por meio dessa representação da Série de Fourier Multidimensional, é possível reconstruir uma função  $f(x)$  original de maneira completa apenas com base em suas componentes harmônicas (WEISZ, 2002).

## 2.3 TÉCNICAS DE CLASSIFICAÇÃO

Com o desenvolvimento das aplicações de inteligência artificial na engenharia biomédica, diversos estudos dedicam-se ao desenvolvimento e análise de eficácia de técnicas de processamento de sinais e aprendizado de máquina para análise de sinais biomecânicos. Tais técnicas visam melhorar a detecção da presença ou ausência de lesões articulares e, eventualmente, a caracterização do tipo de lesão.

Tratando-se dos estudos recentes, um exemplo são os trabalhos de Funaki e Lorentz (2022) e Silva e Camargo (2025), que empregaram abordagens de pré-processamento como filtragem e série de Fourier elíptica para a decomposição espectral, seguidas por métodos de classificação como *Ensemble*, árvore de decisão e máquinas de vetores de suporte (SVM). Tais

abordagens demonstram a diversidade de metodologias que podem ser aplicadas, variando desde o domínio do tempo até o domínio da frequência e tempo-frequência.

No entanto, fatores como a similaridade entre padrões vibroartrográficos de diferentes patologias, bem como a limitação no número de amostras para cada categoria de lesão, ainda representam desafios consideráveis a serem explorados. Conforme destacado por Krecisz e Baczkowicz (2018), tais desafios, como o desbalanceamento do conjunto de dados e o sobreajuste (overfitting) dos modelos de aprendizado, impactam não apenas na identificação binária (presença ou ausência de lesão), mas principalmente na classificação multiclasse.

Apesar dos desafios, os estudos na área têm obtido resultados promissores. Embora a acurácia na classificação multiclasse seja geralmente inferior à binária, ainda se observa um desempenho satisfatório (KRECISZ; BACZKOWICZ, 2018), o que reforça a importância de aprofundar a pesquisa e aprimorar as técnicas empregadas.

Com o objetivo de consolidar o conhecimento atual e agregar as principais tendências e lacunas de pesquisa, realizou-se um levantamento sistemático dos principais estudos publicados na literatura científica. A Tabela 1, que foi elaborada a partir de uma revisão sistemática da literatura, sintetiza esses trabalhos, organizando-os cronologicamente e apresentando, para cada estudo, as principais características extraídas dos sinais vibroartrográficos, os métodos de classificação empregados, as acurácias de desempenho obtidas e o tipo de classificação realizado (presença ou identificação da lesão).

**Tabela 1 – Métodos de processamento e classificação para os diferentes artigos encontrados na literatura.**

<b>Autores (Ano)</b>	<b>Características utilizadas</b>	<b>Método de classificação</b>	<b>Acurácia obtida</b>	<b>Identificação</b>
Rangayyan e Wu (2008)	Domínio do tempo com: Fator de Fórmula; Assimetria; Kurtosis e Entropia	Análise discriminante linear de Fisher; Máquina de vetores de suporte (SVM); Rede de funções de base radial.	Insatisfatórias de acordo com os autores.	Binária: presença ou não de lesão
Wu e Krishnan (2009)	Domínio do tempo e tempo-frequência com: Transformada Rápida de Fourier	Análise discriminante linear de Fisher; Máquina de vetores de suporte de mínimos quadrados	65,17%; 73,03%	Binária: presença ou não de lesão
Moreira (2015)	Domínio do tempo e tempo-frequência com: Transformada Wavelet	K-Vizinhos Mais Próximos; Máquina de vetores de suporte (SVM); Árvore de decisão	89,7%; 84,3%; 82,0%	Binária: presença ou não de lesão
Nalband <i>et al.</i> (2016)	Transformada Wavelet com: Análise de quantificação de recorrência, entropia de amostra, entropia aproximada e energia wavelet	Máquina de vetores de suporte de mínimos quadrados	94,31%	Binária: presença ou não de lesão
Funaki e Lorentz (2022)	Domínio da Frequência baseando-se na curva VI e com decomposição por série de Fourier Bidimensional	Ensemble; Máquina de vetores de suporte (SVM); Árvore de decisão	66,1%; 97,9%; 100%-(máximo)	Binária: presença ou não de lesão
Silva e Camargo (2025)	Domínio da Frequência baseando-se na curva VI e com decomposição por série de Fourier Bidimensional	Máquina de vetores de suporte (SVM); Árvore de decisão	100%-(máximo); 100%-(máximo)	Binária: presença ou não de lesão e identificação se é lesão no ligamento cruzado anterior
Basavaraju <i>et al.</i> (2025a)	Descritores espectrais e suas medidas de variabilidade para HFZCs <sup>7</sup> obtidos de FBSEEW <sup>8</sup>	Classificador Ensemble usando votação suave (ENSC)	94,38%	Binária: presença ou não de lesão

**Fonte: Adaptado de Silva e Camargo (2025).**

<sup>7</sup> HFZCs (*High-Frequency Zero-Crossings*) - Referem-se ao número de vezes que a amplitude de um sinal, após ser filtrado para reter apenas suas componentes de alta frequência, passa pelo valor zero dentro de um intervalo de tempo específico.

<sup>8</sup> FBSEEW (*Filter-Bank-Based Synchrosqueezing Empirical Wavelet Transform*) - É um método de processamento de sinais para análise tempo-frequência que decompõe um sinal em seus modos oscilatórios intrínsecos de forma adaptativa, utilizando um banco de filtros construído a partir do espectro do próprio sinal.

### 3 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os principais fundamentos teóricos que sustentam o desenvolvimento deste trabalho. Inicialmente, são abordados os aspectos anatômicos essenciais da articulação do joelho, com ênfase nas implicações biomecânicas relevantes para a captação de sinais. Em seguida, discutem-se os princípios físicos de funcionamento dos sensores utilizados no processo de aquisição — acelerômetros e giroscópios — bem como os protocolos adotados para a transmissão de dados. Na etapa subsequente, detalha-se o pré-processamento e o processamento dos sinais obtidos, com destaque para a aplicação da Série de Fourier Multidimensional, aqui adaptada e deduzida para representar as assinaturas dos movimentos articulares em múltiplas dimensões. Por fim, são apresentadas as técnicas de extração de características relevantes a partir dos coeficientes de Fourier e os métodos de classificação de lesões articulares, por meio de algoritmos de aprendizado de máquina.

#### 3.1 ESTRUTURA FISIOLÓGICA DO JOELHO

O joelho é a maior e mais complexa articulação sinovial do corpo humano, sendo classificada como uma articulação do tipo dobradiça modificada, capaz de realizar movimentos de flexão, extensão e uma limitada rotação axial (KUMAR *et al.*, 2020). Sua relevância biomecânica é evidenciada pela função de suporte do peso corporal, especialmente na postura ereta, além de participar ativamente na locomoção, absorvendo impactos e permitindo a progressão segura durante a marcha (KUMAR *et al.*, 2020). A articulação do joelho é formada pelas extremidades inferiores do fêmur, superiores da tíbia e pela face posterior da patela, que se articulam entre si para possibilitar a mobilidade e a sustentação de carga. O entalhe intercondilar do fêmur e a eminência intercondilar da tíbia contribuem para a estabilidade óssea da articulação. Devido à sua complexidade anatômica e às forças biomecânicas a que está submetido, o joelho é suscetível a lesões que podem comprometer sua estabilidade e função locomotora. Assim, compreender sua anatomia e fisiologia é imprescindível para prevenir lesões, otimizar a performance biomecânica e entender patologias articulares, como a osteoartrite (KUMAR *et al.*, 2020).

### 3.1.1 Ligamentos

Os ligamentos cruzados anterior (LCA) e posterior (LCP) estão localizados no centro da articulação do joelho e exercem papéis essenciais para a estabilidade mecânica da articulação. O LCA é responsável por limitar a translação anterior da tíbia em relação ao fêmur, sendo mais tenso durante a extensão e mais relaxado na flexão. Por sua vez, o LCP atua estabilizando a tíbia posteriormente e controlando movimentos rotacionais, contribuindo de maneira decisiva para a integridade articular.

### 3.1.2 Ligamentos Colaterais

O ligamento colateral medial (LCM) e o ligamento colateral lateral (LCL) fornecem estabilidade ao joelho nos planos medial e lateral, respectivamente. O LCM resiste a forças em valgo<sup>1</sup>, enquanto o LCL limita forças em varo<sup>2</sup>, sendo auxiliado pela cápsula articular e pelo complexo arcuato na região pósterio-lateral.

### 3.1.3 Meniscos

Os meniscos medial e lateral funcionam como amortecedores e distribuidores de carga entre o fêmur e a tíbia, auxiliando na estabilidade articular, absorvendo choques, contribuindo para a lubrificação e prevenindo lesões degenerativas. Anatomicamente, o menisco medial cobre cerca de 50% da superfície tibial medial, enquanto o lateral cobre 70% da superfície tibial lateral, o que influencia sua mobilidade e vulnerabilidade a lesões (Kumar et al., 2020).

### 3.1.4 Músculos e Tendões

Os músculos e tendões exercem funções essenciais no joelho, sendo o quadríceps femoral responsável pela extensão do joelho, enquanto o semimembranoso atua como estabilizador medial, retraindo o menisco medial durante a flexão para evitar aprisionamentos (Kumar et al., 2020).

---

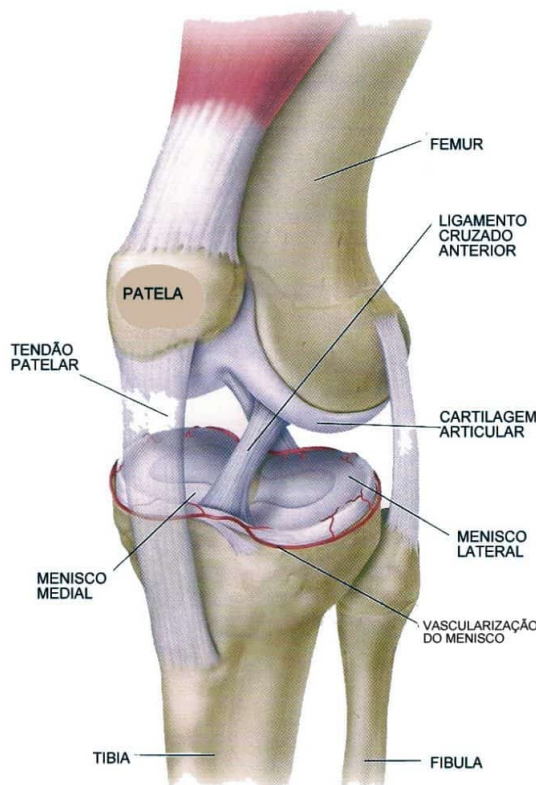
<sup>1</sup> Força em valgo é a força que empurra o joelho para dentro (deslocamento medial da articulação)

<sup>2</sup> Força em varo é a força que empurra o joelho para fora (deslocamento lateral da articulação).

### 3.1.5 Patela

A patela é um osso sesamoide que protege a frente da articulação do joelho e atua como uma polia, responsável por aumentar a eficiência do músculo quadríceps durante a extensão do joelho. Além disso, distribui as forças de compressão na articulação femoropatelar e ajuda a estabilizar o mecanismo extensor.

**Figura 2 – Estrutura fisiológica do joelho.**



**Fonte: Paschoa (2022).**

### 3.1.6 Lesões

Todas as estruturas apresentadas acima desempenham papéis fundamentais no funcionamento e na estabilidade do joelho. Por conta desse uso constante, são nessas mesmas estruturas que ocorrem a maior parte das lesões que afetam a articulação. Lesões ligamentares (como do LCA e LCM), lesões meniscais, danos à cartilagem e problemas envolvendo a patela estão entre as mais comuns (SOUZA; MORAIS, 2016). Estas lesões podem ocorrer de forma aguda, decorrentes de movimentos de torção ou impacto direto, ou podem ter origem em mecanismos degenerativos associados a sobrecarga, treinos excessivos e instabilidades articulares, culminando

em patologias como a osteoartrite.

Dessa maneira, diversas lesões no joelho — como as ligamentares, meniscais, condrais e patelares — podem gerar vibrações mecânicas que são captadas por sensores estrategicamente posicionados na articulação. Esses sinais, quando processados, fornecem subsídios importantes para análise clínica e identificação de padrões associados às lesões. Estudos demonstram que sensores de vibração, como acelerômetros e microfones de contato, têm sido aplicados com sucesso para detectar vibrações articulares em lesões de cartilagem e menisco (FUNAKI; LORENTZ, 2022).

## 3.2 Pré-Processamento

### 3.2.1 Z-score

A técnica de Z-score é um método de padronização estatística amplamente utilizado em tarefas de pré-processamento de dados, especialmente em algoritmos de classificação sensíveis à escala, como o k-Nearest Neighbor (KNN)<sup>3</sup>, Redes Neurais e SVM. Seu objetivo é transformar os valores de uma variável de forma que eles passem a ter uma distribuição com média igual a zero e desvio padrão igual a um, padronizando os dados em uma mesma escala (HENDERI *et al.*, 2021).

O funcionamento do Z-score baseia-se no cálculo do número de desvios padrão que um determinado valor está distante da média da variável. Isso é feito por meio da fórmula:

$$Z = \frac{x - \mu}{\sigma}. \quad (8)$$

em que  $x$  representa o valor original,  $\mu$  a média dos valores da variável, e  $\sigma$  o seu desvio padrão. Essa transformação permite que variáveis com escalas diferentes sejam comparadas de forma justa, tornando o modelo de aprendizado menos suscetível ao domínio de atributos com maiores magnitudes.

Embora o Z-score não altere a forma da distribuição dos dados, sua aplicação é mais apropriada em variáveis com distribuição aproximadamente normal. Em distribuições assimétricas ou com presença significativa de outliers, a padronização pode apresentar menor representatividade. Diferentemente da técnica de Min-Max — que comprime os dados em um

<sup>3</sup> KNN classifica um dado com base nas classes dos K vizinhos mais próximos.

intervalo fixo —, o Z-score preserva a distribuição original, sendo, por isso, mais robusto frente a valores extremos (HENDERI *et al.*, 2021).

Conforme Henderi *et al.* (2021), a aplicação do Z-score contribuiu para uma acurácia estável (95% a 97%) em tarefas de classificação com kNN, destacando-se como uma técnica confiável para cenários em que os dados apresentam variabilidade natural e distribuição próxima à gaussiana. Por essa razão, sua utilização é recomendada em etapas de pré-processamento onde se busca equilibrar variáveis com escalas distintas sem distorcer sua distribuição estatística.

### 3.2.2 Remoção de outliers

Durante a aquisição de dados experimentais ou observacionais, é comum que certos valores se desviem significativamente do comportamento predominante do conjunto. Esses valores atípicos, conhecidos como *outliers*, podem surgir por diversos fatores, como erros de medição, falhas em sensores, ruídos ambientais ou até mesmo eventos raros que não representam o padrão geral dos dados. A presença desses elementos pode comprometer análises estatísticas, enviesar algoritmos de aprendizado de máquina e reduzir a confiabilidade dos resultados obtidos (PIMENTEL *et al.*, 2014).

O impacto dos *outliers* torna-se ainda mais relevante em contextos onde os dados são naturalmente mais sensíveis a ruídos e flutuações, como em sinais fisiológicos ou biomecânicos. Nesses casos, a variação inesperada pode mascarar padrões reais ou induzir interpretações equivocadas. Consoante a Pimentel *et al.* (2014), a identificação e remoção de *outliers* constituem uma etapa essencial do pré-processamento dos dados, contribuindo para maior robustez analítica e fidelidade na reconstrução dos sinais avaliados.

Dentre os métodos empregados para esse fim, destaca-se o filtro por limiar (*threshold*), no qual define-se um intervalo aceitável de valores baseado em parâmetros estatísticos (como a média e o desvio padrão, ou os quartis da distribuição). Valores que excedem esse intervalo são considerados *outliers* e, dependendo da abordagem adotada, podem ser eliminados ou substituídos por estimativas mais coerentes com o conjunto (GOIC *et al.*, 2012).

### 3.2.3 Filtragem e Suavização de Sinais

A aquisição de sinais biomecânicos, como os vibroartrográficos (VAG), está frequentemente sujeita a interferências de natureza instrumental, fisiológica e ambiental. Para assegurar a

qualidade dos dados e mitigar distorções que possam comprometer análises posteriores, é essencial a aplicação de métodos de filtragem e suavização de ruído como parte do pré-processamento. Dentre os métodos mais eficazes e amplamente adotados, destacam-se o filtro de *Savitzky–Golay* e os filtros FIR (*Finite Impulse Response*).

### 3.2.4 Savitzky–Golay

O filtro *Savitzky–Golay* aplica um polinômio de grau definido a uma janela deslizante sobre o sinal, suavizando-o por meio de regressão por mínimos quadrados. Ao contrário de filtros convencionais passa-baixa, este método tem a vantagem de preservar as características morfológicas importantes, como picos e inflexões, sendo ideal para sinais que contêm transições rápidas, como os oriundos da articulação do joelho (HAIDER, 2018). Ele é especialmente útil na remoção de ruídos de alta frequência sem causar distorções significativas na fase do sinal (MARTINEK *et al.*, 2021).

## 3.3 PROCESSAMENTO DOS SINAIS

Como evidenciado por Funaki e Lorentz (2022) e aprofundado por Silva e Camargo (2025), os sinais obtidos por meio da técnica de vibroartrografia, utilizando giroscópios e acelerômetros, apresentam natureza periódica, o que viabiliza sua análise por meio da Série de Fourier. Considerando que os dados são coletados em três dimensões (eixos  $x$ ,  $y$  e  $z$ ), cada conjunto de sensores — composto por um acelerômetro e um giroscópio — gera seis sinais distintos, refletindo diferentes aspectos do movimento do joelho durante a avaliação.

Dada a complexidade desses sinais e a necessidade de analisá-los de forma conjunta, Funaki e Lorentz (2022) propuseram em seu estudo não apenas o uso da Série de Fourier tradicional, mas sim a aplicação da Série de Fourier bidimensional. Essa abordagem permite a análise simultânea da influência mútua entre dois sinais de naturezas diferentes, fornecendo uma representação mais abrangente e precisa das dinâmicas envolvidas nos movimentos articulares capturados, conforme explicado na Seção 2.2.2 do Capítulo anterior.

No entanto, embora a decomposição pela Série de Fourier bidimensional tenha se mostrado eficaz e tenha trazido avanços nos estudos de Funaki e Lorentz (2022), os sinais adquiridos ainda apresentam uma complexidade adicional. Como apontado pelos próprios autores, a cada instante de tempo são registradas múltiplas variáveis simultaneamente — mais

precisamente, seis sinais distintos por par de sensores (acelerômetro e giroscópio). Essa alta dimensionalidade por amostra abre espaço para o uso de abordagens mais robustas, capazes de representar, de forma conjunta e integrada, os diferentes aspectos do movimento articular. Assim, com o objetivo de extrair o máximo de características possíveis, preservando a fidelidade do sinal original, propõe-se o uso da Série de Fourier Multidimensional, abordada na Seção 2.2.3.

### 3.3.1 Série de Fourier Multidimensional

Portanto, baseando-se na Equação 1, que descreve a série de Fourier unidimensional (simples), e seguindo a abordagem proposta por Funaki e Lorentz (2022) para a aplicação da Série de Fourier Bidimensional na análise de sinais, é possível estender esse conceito para a Equação 7, com o mesmo objetivo da análise de sinais. Essa última equação representa a Série de Fourier Multidimensional, permitindo a análise simultânea de múltiplas variáveis e mantendo o mesmo propósito: decompor e interpretar sinais complexos em domínios mais compreensíveis.

Sendo assim, partindo da representação da exponencial complexa  $e^{ik \cdot x}$ , ela pode ser decomposta em funções trigonométricas como a soma de cosseno e um seno imaginário (identidade de Euler), tal que  $e^{-i k \cdot x} = \cos(k \cdot x) - i \sin(k \cdot x)$ . Substituindo isso então na Equação 6, obtém-se a expressão do coeficiente de Fourier Multidimensional  $\hat{f}(k)$  em termos de funções cosseno e seno, conforme apresentado na equação a seguir:

$$\hat{f}(k) = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x) [\cos(k \cdot x) - i \sin(k \cdot x)] dx. \quad (9)$$

Define-se a parte real do coeficiente  $\hat{f}(k)$  como  $A_k$ , e a parte imaginária como  $B_k$ , de modo que  $\hat{f}(k) = A_k - i B_k$ . Substituindo a forma trigonométrica da exponencial complexa na expressão  $\hat{f}(k)e^{ik \cdot x}$ , e utilizando a decomposição de  $\hat{f}(k)$  em termos de  $A_k$  e  $B_k$ , obtém-se a seguinte equação:

$$\hat{f}(k)e^{ik \cdot x} = [A_k \cos(k \cdot x) + B_k \sin(k \cdot x)] + i[A_k \sin(k \cdot x) - B_k \cos(k \cdot x)]. \quad (10)$$

Observa-se, contudo, que na Equação 10, ainda persiste a presença de uma parte imaginária, mesmo que a função  $f(x)$  (o sinal original) seja puramente real. Essa questão é discutida por Stein e Shakarchi (2003), mais especificamente no Capítulo 2 ("Basic Properties of Fourier Series"). Nesse capítulo, os autores demonstram que  $\hat{f}(-k) = \overline{\hat{f}(k)}$ . Isso implica que a

parte imaginária dos coeficientes de Fourier é uma função ímpar de  $k$ , enquanto a parte real é uma função par. Consequentemente, ao somar os termos correspondentes a  $k$  e  $-k$  na Série de Fourier, as partes imaginárias se cancelam, enquanto as partes reais se reforçam, garantindo que a soma final represente uma função real.

Dessa forma, estendendo esse conceito para a abordagem multidimensional na Equação 7, como o somatório é realizado para todo  $k \in \mathbb{Z}^n$ , ou seja, considerando frequências de coeficientes  $k$  e  $-k$ , cada termo e seu oposto aparecem na série. Assim, analisando uma única dimensão  $n$  do somatório, obtém-se então o termo combinado no somatório  $\hat{f}(k)e^{ik \cdot x} + \overline{\hat{f}(k)}e^{-ik \cdot x}$ . Nessa soma, as partes imaginárias se anulam e as partes reais se reforçam, de acordo com a identidade de que  $z + \bar{z} = 2 \operatorname{Re}(z)$ . Assim, é obtida a seguinte representação geral para a soma de termos de frequências  $k$  opostas:

$$\hat{f}(k)e^{ik \cdot x} + \hat{f}(-k)e^{-ik \cdot x} = 2 \operatorname{Re}[\hat{f}(k)e^{ik \cdot x}]. \quad (11)$$

No entanto, percebe-se que o termo correspondente à frequência  $k = 0$  não é contemplado nessa generalização, pois não há uma frequência oposta a ela. Então, para esse termo específico, considerando que  $\sin(0) = 0$  e que o termo imaginário presente na Equação 9 é proporcional a  $\sin(k \cdot x)$ , a parte imaginária para a frequência  $k = 0$  é nula, restando apenas a parte real. Assim, chamando esse termo para a frequência  $k = 0$  de  $A_0$  e considerando que  $\cos(0) = 1$ , ele pode ser escrito como:

$$A_0 = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x) dx. \quad (12)$$

Portanto, a parte imaginária do somatório se anula para as frequências opostas e é inexistente para o termo independente  $A_0$ , restando apenas a parte real da série. Dessa forma, a soma resultante converge para a representação esperada da função original, que é puramente real. Assim, ao somar os termos correspondentes às frequências opostas, juntamente com o termo  $A_0$ , generaliza-se o somatório da Equação 7, reconstruindo a função original.

Com isso, considerando a soma sobre as frequências opostas e o termo  $A_0$ , obtém-se a seguinte equação geral para a Série de Fourier Multidimensional (com coeficientes  $A_k$  e  $B_k$  definidos a partir da Equação 9):

$$f(x) = A_0 + \sum_{k \in \mathbb{Z}_+^n} \left[ A_k \cos(k \cdot x) + B_k \sin(k \cdot x) \right], \quad (13)$$

$$A_k = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x) \cos(k \cdot x) dx, \quad (14)$$

$$B_k = \frac{1}{(2\pi)^n} \int_{\mathbb{T}^n} f(x) \sin(k \cdot x) dx. \quad (15)$$

Portanto, a Equação 13 apresentada acima é equivalente à Equação 6 discutida nos estudos de Weisz (2002). Nessa formulação,  $n$  representa o número de dimensões consideradas na análise. Os vetores  $k$  e  $x$  são definidos da seguinte maneira:  $k$  é um vetor de inteiros, expresso como  $k = (k_1, k_2, \dots, k_n)$ , que representa os índices de frequência em cada uma das  $n$  dimensões de análise; já  $x$  é um vetor cujas componentes  $x = (x_1, x_2, \dots, x_n)$  correspondem aos valores da variável analisada para cada dimensão.

Analisando essa equação, percebe-se que ela necessita de vetores  $k$  e  $x$ , diferentemente das equações propostas por Funaki e Lorentz (2022) para a análise bidimensional em seus estudos. Sendo assim, de forma a deixar as equações para a análise multidimensional em uma estrutura semelhante às apresentadas por Funaki e Lorentz (2022), algumas manipulações algébricas são necessárias.

Inicialmente, é possível reescrever o termo  $k \cdot x$  presente na Equação 13 em termos de somatórios escalares, considerando a frequência angular fundamental de cada dimensão. Como cada dimensão  $n$  pode ter um período  $T_n$ , a frequência angular fundamental associada a essa dimensão é definida por  $\frac{2\pi}{T_n}$ . Portanto, de forma a incorporar explicitamente os períodos reais de cada dimensão, pode-se fazer a substituição<sup>4</sup>:  $k_n \mapsto h_n \frac{2\pi}{T_n}$ . Assim, o produto escalar  $k \cdot x$  pode ser expandido como:

$$k \cdot x = \sum_{j=1}^n \frac{2\pi h_j}{T_j} x_j. \quad (16)$$

Dessa forma, elimina-se a notação vetorial e passa-se a trabalhar apenas com índices escalares e frequências reais, deixando a série multidimensional em uma estrutura mais próxima daquela apresentada por Funaki e Lorentz (2022) na análise bidimensional. Além disso, o período fundamental de cada dimensão é incorporado explicitamente, tornando a fórmula mais

<sup>4</sup> Assim,  $h_n$  representa o número de harmônicos associado à dimensão  $n$ , ou seja, a quantidade de ciclos completos da frequência fundamental que cabem dentro do período  $T_n$ . Essa substituição torna explícita a relação entre o índice de frequência e a frequência angular fundamental da função periódica analisada, conforme apresentado nas equações para a análise bidimensional utilizadas por Funaki e Lorentz (2022).

alinhada com aplicações práticas, como análise de sinais periódicos em domínios físicos - o que é necessário para trabalhar com os sinais de vibração adquiridos.

Com o mesmo objetivo, é possível transformar o somatório vetorial  $\sum_{(k \in \mathbb{Z}_+^n)}$ , que aparece na Equação 13 da série de Fourier multidimensional. Então, apresentando-o em uma forma mais expandida, usando somatórios escalares aninhados — um para cada dimensão, o somatório vetorial único é substituído por:

$$\sum_{h \in \mathbb{Z}_+^n} \longrightarrow \sum_{h_1=1}^{\infty} \sum_{h_2=1}^{\infty} \cdots \sum_{h_n=1}^{\infty}. \quad (17)$$

De forma análoga, a integral vetorial única que aparece nos coeficientes de Fourier pode ser substituída por integrais aninhadas em cada dimensão. Então, realizando isso e centrando o intervalo de integração em torno da origem<sup>5</sup>, como utilizado por Funaki e Lorentz (2022), a integral vetorial é expressa como:

$$\int_{-\frac{T_1}{2}}^{\frac{T_1}{2}} \int_{-\frac{T_2}{2}}^{\frac{T_2}{2}} \cdots \int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} f(x) dx_1 \cdots dx_n. \quad (18)$$

Essa mudança facilita a visualização da contribuição de cada harmônico ao sinal, além de permitir uma implementação computacional mais intuitiva, pois dispensa o uso da notação vetorial e da álgebra de produtos vetoriais. Ainda, aproxima as equações da Série Multidimensional ao formato apresentado por Funaki e Lorentz (2022) em seus estudos bidimensionais.

Finalmente, após as manipulações algébricas realizadas e bem definidas sobre a forma inicial da Série de Fourier Multidimensional apresentada por Weisz (2002), adapta-se a equação para um formato mais prático e viável de ser aplicado na análise de sinais vibroartrográficos. Essa adaptação visa também alinhar a estrutura com a apresentada por Funaki e Lorentz (2022) para o caso bidimensional, facilitando a comparação entre diferentes dimensões. Assim, obtém-se a equação final que descreve a Série de Fourier Multidimensional, juntamente com seus coeficientes associados, apresentados a seguir:

<sup>5</sup> A escolha de integrar de  $-\frac{T_n}{2}$  a  $\frac{T_n}{2}$  em cada dimensão é uma alternativa comum à notação de 0 a  $T_n$ . Ambas representam uma integral em um período completo da função periódica. Entretanto, o uso de limites centrados na origem é especialmente vantajoso em análises que exploram as simetrias das funções seno e cosseno, além de facilitar a separação de termos pares e ímpares em algumas aplicações. Essa escolha é didática e não altera o valor final dos coeficientes de Fourier (STEIN; SHAKARCHI, 2003).

$$f(x) = A_0 + \sum_{h_1=1}^{\infty} \sum_{h_2=1}^{\infty} \cdots \sum_{h_n=1}^{\infty} \left[ A_h \cos \left( \sum_{j=1}^n h_j \frac{2\pi}{T_j} x_j \right) + B_h \sin \left( \sum_{j=1}^n h_j \frac{2\pi}{T_j} x_j \right) \right], \quad (19)$$

$$A_0 = \frac{1}{T_1 T_2 \cdots T_n} \int_{-\frac{T_1}{2}}^{\frac{T_1}{2}} \int_{-\frac{T_2}{2}}^{\frac{T_2}{2}} \cdots \int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} f(x) dx_1 dx_2 \cdots dx_n, \quad (20)$$

$$A_h = \frac{2}{T_1 T_2 \cdots T_n} \int_{-\frac{T_1}{2}}^{\frac{T_1}{2}} \int_{-\frac{T_2}{2}}^{\frac{T_2}{2}} \cdots \int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} f(x) \cos \left( \sum_{j=1}^n h_j \frac{2\pi}{T_j} x_j \right) dx_1 dx_2 \cdots dx_n, \quad (21)$$

$$B_h = \frac{2}{T_1 T_2 \cdots T_n} \int_{-\frac{T_1}{2}}^{\frac{T_1}{2}} \int_{-\frac{T_2}{2}}^{\frac{T_2}{2}} \cdots \int_{-\frac{T_n}{2}}^{\frac{T_n}{2}} f(x) \sin \left( \sum_{j=1}^n h_j \frac{2\pi}{T_j} x_j \right) dx_1 dx_2 \cdots dx_n. \quad (22)$$

### 3.4 Métodos de Classificação

Após o processamento dos sinais, torna-se essencial identificar padrões e características de destaque, ou seja, assinaturas específicas que permitam diferenciar os tipos de lesão articular. Essa etapa é fundamental para viabilizar a correta classificação das lesões e, conseqüentemente, uma avaliação clínica mais precisa.

Então, uma vez que as características relevantes tenham sido extraídas a partir dos coeficientes da Série de Fourier Multidimensional, essas informações passam a compor um vetor de atributos representativo de cada sinal analisado. Esse vetor sintetiza os aspectos mais discriminativos dos movimentos articulares, servindo como base para a etapa seguinte: a classificação, com o objetivo de apoiar o diagnóstico das lesões. Nessa fase, algoritmos de aprendizado de máquina são aplicados para identificar padrões nos atributos a eles apresentados, e assim atribuir, a cada amostra, uma classe correspondente ao tipo de lesão articular.

O Aprendizado de Máquina<sup>6</sup> constitui uma subárea da Inteligência Artificial (IA), voltada ao desenvolvimento de algoritmos capazes de identificar padrões e relações em conjuntos de dados. A partir dessas correlações, os modelos de aprendizado de máquina podem realizar classificações ou predições de maneira automática, aprimorando seu desempenho com base na experiência adquirida ao longo do tempo (MISHRA, 2024).

<sup>6</sup> Do inglês *Machine Learning*.

Figura 3 – Métodos de aprendizado de máquina.



Fonte: Adaptado de (MISHRA, 2024).

A Figura 3 ilustra as diferentes abordagens que podem ser utilizadas no contexto do Aprendizado de Máquina. Na sequência, são detalhadas cada uma dessas abordagens, de acordo com o proposto por Mishra (2024):

- **Aprendizado Supervisionado:** consiste no treinamento de um modelo a partir de um conjunto de dados rotulados, ou seja, cujas saídas corretas são previamente conhecidas. O objetivo é que o modelo aprenda a generalizar os padrões observados nos dados de treinamento para que consiga realizar previsões sobre um conjunto de novos dados não utilizados durante o processo de treinamento. Essa abordagem exige, portanto, conhecimento prévio e representativo sobre o problema em questão. Entre os algoritmos comumente utilizados nesta categoria estão as Árvores de Decisão, Máquinas de Vetores de Suporte (*Support Vector Machines* – SVM), Redes Neurais Artificiais e métodos de *Ensemble*.
- **Aprendizado Não Supervisionado:** nessa abordagem, os dados utilizados não são rotulados, ou seja, não se conhece previamente a saída correta. O foco está na descoberta de padrões, agrupamentos ou estruturas subjacentes diretamente nos dados de entrada, sem a necessidade de um treinamento anterior supervisionado. Algoritmos de clusterização, como o *K-Means* e métodos de agrupamento hierárquico, são exemplos dessa abordagem.

- **Aprendizado por Reforço:** baseia-se na interação de um agente com o ambiente no qual está inserido. O agente toma decisões sequenciais com o objetivo de maximizar uma recompensa cumulativa. A cada ação realizada, ele recebe uma resposta (*feedback*) do ambiente, positiva ou negativa, permitindo ajustar suas decisões futuras. Essa abordagem é útil em contextos onde a tomada de decisão ocorre ao longo do tempo. Um dos algoritmos mais conhecidos nessa categoria é o *Q-Learning*.

Nesse contexto, ao associar as diferentes abordagens de aprendizado de máquina ao problema da classificação dos sinais vibroartrográficos para a identificação de lesões no joelho, observa-se que a estratégia mais adequada é o uso do aprendizado supervisionado. Essa abordagem, conforme proposto e implementado também por Silva e Camargo (2025) em seus estudos, consiste na construção de modelos treinados com dados previamente rotulados, os quais, uma vez validados, podem ser aplicados na identificação automática de lesões em novos pacientes.

Portanto, para a aplicação dessa abordagem supervisionada, diferentes algoritmos classificadores podem ser utilizados. Uma questão relevante discutida por Mishra (2024) refere-se à capacidade de generalização do modelo, que pode apresentar problemas de *overfitting*<sup>7</sup> ou *underfitting*<sup>8</sup>. Assim, o processo de treinamento deve ser conduzido de modo a evitar ambos os extremos: o modelo não deve apresentar desempenho significativamente superior nos dados de treinamento em relação aos dados de teste, tampouco apresentar resultados insatisfatórios em ambos os conjuntos.

### 3.4.1 Árvore de Decisão

Os algoritmos de árvore de decisão são uma das abordagens mais estudadas e utilizadas no aprendizado de máquina, destacando-se por sua simplicidade e facilidade de interpretação<sup>9</sup> (BLOCKKEEL *et al.*, 2023). Apesar de sua estrutura intuitiva e simples, esses algoritmos possuem grande capacidade para resolver problemas de classificação, sendo ferramentas amplamente aplicáveis na análise de dados.

<sup>7</sup> *Overfitting* ocorre quando o modelo se ajusta excessivamente aos dados de treinamento, perdendo a capacidade de generalizar para novos dados.

<sup>8</sup> *Underfitting* ocorre quando o modelo não consegue captar os padrões subjacentes nem mesmo nos dados de treinamento.

<sup>9</sup> Tem sua estrutura se assemelhando a de uma árvore real e de fácil representação gráfica, sendo então facilmente compreendida e interpretada por humanos (BLOCKKEEL *et al.*, 2023).

O funcionamento de uma árvore de decisão consiste na predição da saída  $y$  de uma função  $f(x)$ , por meio de uma sequência de decisões baseada em um modelo previamente treinado. Essa sequência é representada como uma estrutura hierárquica que se assemelha a uma árvore, sendo que as decisões consideradas mais importantes<sup>10</sup> ficam mais próximas ao nó raiz. Estruturalmente, a árvore de decisão é composta por três tipos principais de nós (o principal e suas ramificações) (CHARBUTY; ABDULAZEEZ, 2021):

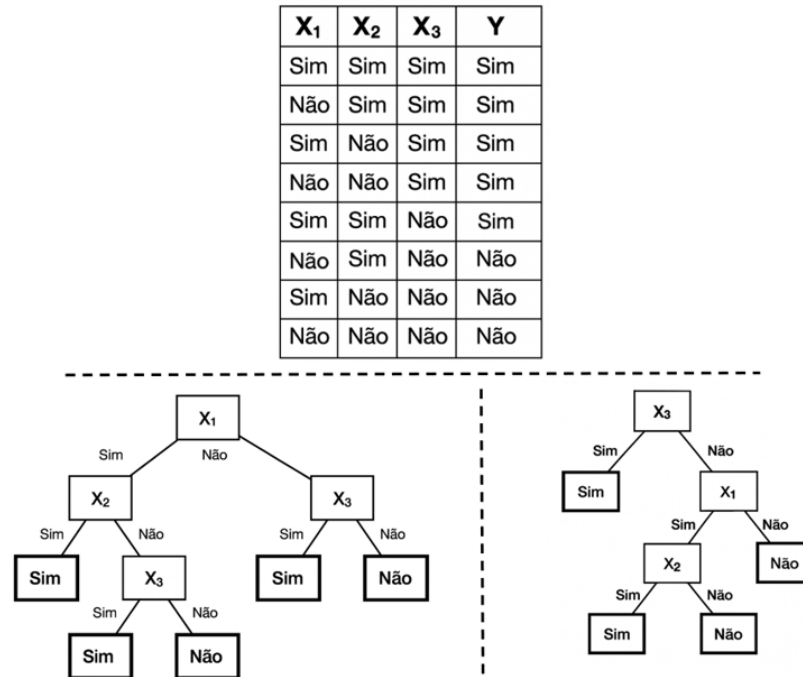
- **Nó raiz:** ponto inicial da árvore, onde ocorre a primeira divisão do conjunto de dados com base em uma condição de decisão;
- **Nós de decisão:** nós intermediários, que representam pontos de avaliação de novas condições (onde decisões são tomadas), direcionando a amostra a diferentes ramos da árvore. São análogos aos galhos de uma árvore real;
- **Nós folha:** nós finais da estrutura, que fornecem a saída predita para  $y$ , seja uma classe (em problemas de classificação) ou um valor numérico (em regressão). São comparáveis às folhas de uma árvore natural.

Esse processo de divisão sucessiva do espaço de atributos permite que a árvore de decisão construa regras claras e interpretáveis, que buscam se aproximar ao máximo da função original e gerar um modelo que possa ser aplicado a outros conjuntos de dados semelhantes. Um aspecto destacado por Blockeel *et al.* (2023) é que não existe uma única árvore de decisão (modelo) que represente uma função  $f(x)$ . Portanto, diferentes configurações e critérios de divisão podem levar à obtenção de modelos distintos.

Esse comportamento é ilustrado na Figura 4, onde uma função  $f(x) = X_1 \wedge X_2 \vee X_3$  é representada por sua tabela verdade e dois diferentes modelos de árvore de decisão. Isso demonstra como diferentes modelos treinados podem representar a mesma função original, e por isso é necessário realizar a validação e comparação entre as métricas dos mesmos, para que se obtenha o modelo mais otimizado (CHARBUTY; ABDULAZEEZ, 2021).

<sup>10</sup> Isto é, os atributos mais relevantes (que seriam os quais subdividiriam o dataset em pares, no caso binário) são sequenciados de forma a serem interpretados primeiramente pelo modelo.

**Figura 4 – A função booleana  $Y = X_1 \wedge X_2 \vee X_3$  e duas árvores de decisão que representam suas soluções.**



Fonte: Adaptado de (BLOCHEEL *et al.*, 2023).

Para definir a estrutura do modelo, como ilustrado na Figura 4, dois parâmetros são fundamentais: a profundidade máxima da árvore e o número de ramificações (folhas) por nó. A profundidade indica o máximo de decisões que o modelo pode tomar até se alcançar a saída  $y$ , o que acaba influenciando diretamente o risco de *overfitting* (quando muito profunda) ou *underfitting* (quando rasa). Já o número de ramificações (folhas que saem de cada nó) define quantas classes podem ser separadas em cada nó, ou seja, quantas decisões são possíveis de serem tomadas. Por isso, ela impacta diretamente na complexidade do modelo, especialmente em tarefas de classificação multiclasse (BLOCHEEL *et al.*, 2023).

### 3.4.2 Máquina de Vetores de Suporte (SVM)

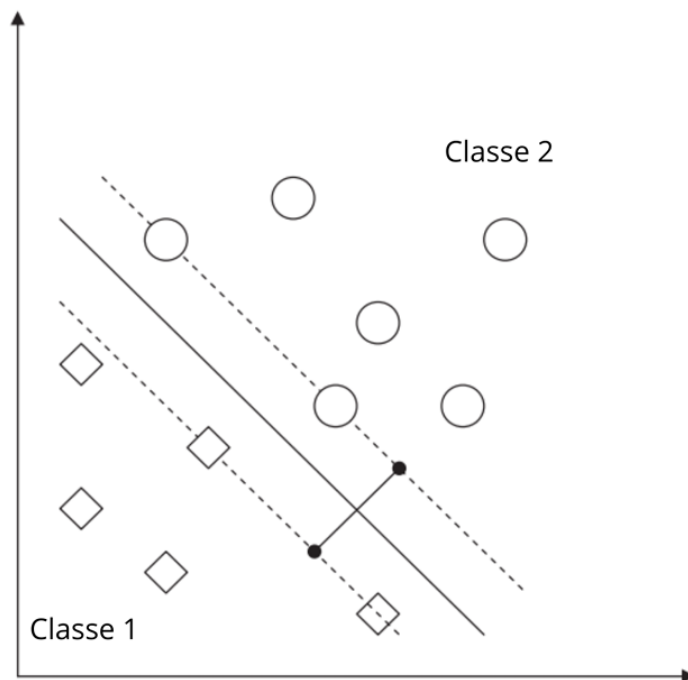
As Máquinas de Vetores de Suporte (SVM) são algoritmos de aprendizado de máquina supervisionado, inicialmente propostos para tarefas de detecção de padrões e classificação de dados. A abordagem se destaca pela capacidade de construir classificadores eficazes com uma estrutura relativamente simples e com baixa dependência de atributos, o que reduz significativamente a demanda por recursos computacionais (AMARAPPA; SATHYANARAYANA, 2014).

O método trata os dados de entrada como vetores  $n$ -dimensionais em um espaço  $\mathbb{R}^n$ .

Com base nisso, o SVM busca separar os dados por meio de um hiperplano<sup>11</sup>  $(n - 1)$ -dimensional (denominado plano canônico<sup>12</sup>) construído a partir dos vetores de suporte e das margens entre as classes.

A definição das margens para o hiperplano canônico ocorre por meio de dois hiperplanos paralelos a ele, conforme ilustrado na Figura 5. O objetivo é posicionar o hiperplano de forma que a distância entre ele e os pontos mais próximos de cada classe (os vetores de suporte) seja maximizada. Assim, os pontos localizados entre o hiperplano e os hiperplanos marginais (representados por linhas pontilhadas na Figura 5) pertencem a uma região de incerteza ou instabilidade na classificação (SUBASI, 2013).

**Figura 5 – Classificação de dados pelo modelo SVM em um espaço 2D.**



Fonte: Adaptado de Subasi (2013).

Apesar de a Figura 5 ilustrar o uso do SVM em um plano bidimensional com duas classes, o método não se limita a esse contexto. Amarappa e Sathyanarayana (2014) destacam que o SVM pode ser aplicado a problemas de classificação multiclasse, desde que cada instância possa pertencer a apenas uma única classe ao mesmo tempo. Na maioria dos casos, a solução adotada consiste em transformar o problema multiclasse em múltiplos problemas binários. Para isso, duas abordagens principais são utilizadas: *one-versus-all* e *one-versus-one*. A estratégia

<sup>11</sup> Um hiperplano é uma generalização de um plano em espaços de dimensão superior. Em um espaço  $\mathbb{R}^n$ , trata-se de um subespaço de dimensão  $n - 1$  que pode ser usado para separar conjuntos de dados no SVM.

<sup>12</sup> O plano canônico, no contexto do SVM, é o hiperplano de separação ideal, posicionado de forma a maximizar a distância (margem) entre os dados das diferentes classes (AMARAPPA; SATHYANARAYANA, 2014).

*one-versus-all* é a mais comum, consistindo na criação de um classificador para cada classe, resultando em  $n$  classificadores para um total de  $n$  classes. Já a abordagem *one-versus-one* define um classificador para cada par de classes, totalizando  $n(n-1)/2$  classificadores, o que a torna computacionalmente mais custosa e lenta do que a anterior (AMARAPPA; SATHYANARAYANA, 2014).

### 3.4.3 Ensemble Learning

Os métodos ensemble são abordagens de aprendizado de máquina que combinam múltiplos modelos de classificação (também chamados de modelos de base) para gerar uma previsão final mais robusta, precisa e confiável. Essa técnica se baseia na premissa de que diferentes modelos podem apresentar pontos fortes e fracos distintos, e que, ao combiná-los, é possível reduzir tanto o viés quanto a variância, além de aumentar a capacidade de generalização (MAHAJAN *et al.*, 2023).

Embora o ensemble não garanta matematicamente uma performance superior ao melhor modelo individual, estudos demonstram que ele reduz significativamente a probabilidade de se obter um modelo com desempenho ruim. Essa vantagem decorre da diversidade dos classificadores e da forma como as previsões são combinadas (TEIXEIRA, 2022).

De modo geral, conforme Teixeira (2022), os métodos *ensemble* de classificação podem ser divididos em três categorias principais:

- **Bagging (Bootstrap Aggregating):** Divisão da base de treinamento em subconjuntos, cada um sendo treinado por diferentes algoritmos, cujas previsões são posteriormente combinadas por meio de votação<sup>13</sup> para determinar a classificação final.
- **Boosting:** Execução repetida sobre o mesmo conjunto de treinamento, aplicando diversos *weak learners* (modelos fracos, que apresentam desempenho apenas ligeiramente superior ao acaso). Em seguida, atribuem-se pesos aos *weak learners* com base em seu desempenho, de forma que aqueles com melhores resultados contribuam mais para a predição final.
- **Stacking:** Método que, em vez de utilizar diretamente os dados originais, aprende um conjunto de metadados derivados ou transformados.

<sup>13</sup> Votação - é o processo de combinar as previsões de múltiplos modelos

Essa divisão conceitual permite compreender melhor as estratégias de combinação e o potencial de cada método, facilitando a escolha do *ensemble* mais adequado de acordo com o problema em estudo.

### 3.5 Métricas de avaliação dos Classificadores

Quando se desenvolve um novo modelo de aprendizado de máquina, torna-se fundamental validar seu desempenho por meio da aplicação em um conjunto de dados de teste. Essa etapa permite a comparação das métricas obtidas com as de modelos previamente estabelecidos, facilitando a seleção daqueles que apresentam desempenho mais robusto e eficaz. Modelos que não atingem níveis satisfatórios podem ser descartados, enquanto aqueles com resultados promissores podem ser ajustados e refinados (RAINIO *et al.*, 2024).

No caso do aprendizado supervisionado, a validação do desempenho do modelo é realizada como parte do processo de desenvolvimento, que compreende tanto o treinamento quanto a validação. Para isso, o conjunto inicial de dados é subdividido em dois subconjuntos distintos: um para o treinamento e outro para a validação. Após o treinamento, o modelo realiza previsões sobre o conjunto de validação, cujas saídas são então comparadas com os rótulos reais conhecidos. A partir disso, métricas são calculadas, oferecendo diferentes perspectivas sobre o desempenho do modelo em relação a características específicas. Como discutido por Rainio *et al.* (2024), a análise e a interpretação adequadas dessas métricas são cruciais, especialmente em aplicações na área médica, como é o caso da aplicação da vibroartrografia, para realizar uma comparação com os resultados obtidos pelo diagnóstico humano.

No contexto das métricas de avaliação, Rainio *et al.* (2024) propõe, para cenários de classificação multiclasse, a utilização de uma matriz de confusão de dimensões  $n \times n$ , onde  $n$  corresponde ao número de classes analisadas, conforme ilustrado na Tabela 2. O autor ainda reforça que, embora os termos Verdadeiros Positivos (VP), Falsos Positivos (FP), Falsos Negativos (FN) e Verdadeiros Negativos (VN) sejam originalmente definidos para problemas binários, eles podem ser estendidos para o contexto multiclasse ao se considerar cada classe individualmente contra todas as demais.

**Tabela 2 – Matriz de confusão para classificação de  $n$  classes.**

Real / Previsto	$C_1$	$C_2$	$\dots$	$C_n$	Total ( $N_i$ )
$C_1$	$a_{11}$	$a_{12}$	$\dots$	$a_{1n}$	$N_1$
$C_2$	$a_{21}$	$a_{22}$	$\dots$	$a_{2n}$	$N_2$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$
$C_n$	$a_{n1}$	$a_{n2}$	$\dots$	$a_{nn}$	$N_n$
Total ( $P_j$ )	$P_1$	$P_2$	$\dots$	$P_n$	$N$

Fonte: Autoria própria.

Nessa matriz de confusão, o termo  $C_i$  representa a classe  $i$ , sendo a classe de referência, enquanto o termo  $C_j$  (coluna) representa a classe predita. O elemento  $a_{ij}$  indica o número de amostras da classe real  $C_i$  que foram classificadas pelo modelo como pertencentes à classe  $C_j$ . O valor  $N_i$  corresponde ao total de amostras reais da classe  $C_i$ , ou seja, a soma de todos os elementos da linha  $i$  da matriz. Por fim,  $P_j$  representa o total de amostras classificadas pelo modelo como sendo da classe  $C_j$ , obtido pela soma de todos os elementos da coluna  $j$ .

Rainio *et al.* (2024) destaca que, mesmo em problemas multiclasse, é possível empregar as mesmas métricas tradicionalmente utilizadas em classificações binárias — como acurácia, precisão, sensibilidade (*recall*) e especificidade — as quais também foram adotadas por Silva e Camargo (2025) em seus estudos.

Portanto, para ser possível aplicar essas métricas citadas acima, é necessário encontrar os valores de VP, VN, FP e FN com base na matriz para  $n$  classes. Assim sendo, Rainio *et al.* (2024) propõe as seguintes equações:

$$VP_i = a_{ii}, \quad VN_i = \sum_{j \neq i} \sum_{h \neq i} a_{jh}, \quad FN_i = \sum_{j \neq i} a_{ij}, \quad FP_i = \sum_{j \neq i} a_{ji}, \quad i = 1, \dots, n. \quad (23)$$

### 3.5.1 Acurácia

A acurácia é uma métrica que descreve a porcentagem de acertos gerais de um classificador. Ela é especialmente útil para conjuntos de dados balanceados, ou seja, que possuem quantidades semelhantes de amostras para as diferentes classes. No entanto, em conjuntos desbalanceados, ela pode ser enganosa, tendendo a favorecer as classes majoritárias, ocultando o desempenho insatisfatório em classes minoritárias (SATHYANARAYANAN; TANTRI, 2024).

Além disso, ela pode fornecer uma falsa impressão de desempenho elevado quando o número de verdadeiros negativos (VN) é alto. Isso porque contabiliza todos os acertos (ver-

dadeiros positivos e verdadeiros negativos), sem diferenciar entre tipos de erro, reforçando a necessidade do uso de mais de uma métrica para avaliação (RAINIO *et al.*, 2024). Assim, a acurácia é definida como a razão entre a soma dos acertos (VP e VN) e o número total de amostras, como mostrado na equação abaixo:

$$A = \frac{VP + VN}{VP + FP + FN + VN} 100\%. \quad (24)$$

### 3.5.2 Precisão

Em conjuntos de dados com mais de uma classe, a precisão é uma métrica especialmente útil para complementar e contextualizar os resultados fornecidos pela acurácia, ajudando inclusive a corrigir interpretações equivocadas (SATHYANARAYANAN; TANTRI, 2024). A precisão avalia a proporção de predições corretas entre todas as predições classificadas como positivas, ou seja, corresponde à razão entre os verdadeiros positivos (VP) e a soma entre verdadeiros positivos e falsos positivos (VP + FP), conforme a equação a seguir:

$$S = \frac{VP}{VP + FP} 100\%. \quad (25)$$

### 3.5.3 Sensibilidade

A sensibilidade, também conhecida como *recall* ou taxa de verdadeiros positivos (TPR)<sup>14</sup>, é uma métrica que serve para expressar a capacidade do modelo em prever corretamente os casos positivos nos dados. Ou seja, ela representa a proporção de verdadeiros positivos (VP) preditos em relação ao total de amostras que realmente pertencem à classe positiva (VP + FN). Sathyanarayanan e Tantri (2024) ainda destacam que essa é uma métrica interessante a ser utilizada em casos médicos, onde os casos de falsos negativos devem ser minimizados, pois a não identificação de uma lesão pode trazer consequências problemáticas.

Assim, a sensibilidade pode ser obtida conforme a seguinte equação:

$$S = \frac{VP}{VP + FN} 100\%. \quad (26)$$

<sup>14</sup> Do inglês *True Positive Rate (TPR)*.

### 3.5.4 Especificidade

A especificidade, também conhecida como taxa de verdadeiros negativos (TNR)<sup>15</sup>, é uma métrica que serve para completar a sensibilidade. Ela indica o quão bem o modelo proposto prediz corretamente resultados negativos, ou seja, expressa a porcentagem de todos os valores realmente negativos previstos (VN) em relação ao total de amostras negativas (VN + FP) (SATHYANARAYANAN; TANTRI, 2024). Assim, ela é expressa de acordo com a equação apresentada abaixo:

$$E = \frac{VN}{VN + FP} 100\%. \quad (27)$$

---

<sup>15</sup> Do inglês *True Negative Rate (TNR)*.

## 4 METODOLOGIA

Neste Capítulo são apresentados os procedimentos metodológicos adotados para a análise dos sinais vibroartrográficos utilizados neste estudo, bem como o desenvolvimento do aplicativo de apoio ao processo. A abordagem metodológica geral empregada na coleta de dados é inicialmente apresentada, conforme os estudos desenvolvidos por Funaki e Lorentz (2022) e Silva e Camargo (2025). Em seguida, é detalhado o procedimento utilizado no processamento dos dados coletados e na obtenção dos coeficientes da Série de Fourier Multidimensional, finalizando com a descrição do processo de reconstrução dos sinais originais a partir dos coeficientes previamente determinados.

Na etapa de classificação das lesões, são descritos os procedimentos de treinamento e validação dos modelos de aprendizado de máquina. Por fim, é apresentado o desenvolvimento do aplicativo de apoio, incluindo sua arquitetura, usabilidade e integração com o sistema.

### 4.1 Procedimento de coleta de dados

O processo de aquisição de dados adotado neste estudo segue a abordagem proposta por Funaki e Lorentz (2022) e aprimorada por Silva e Camargo (2025) em estudos anteriores, uma vez que serão utilizados os dados previamente coletados por estes autores.

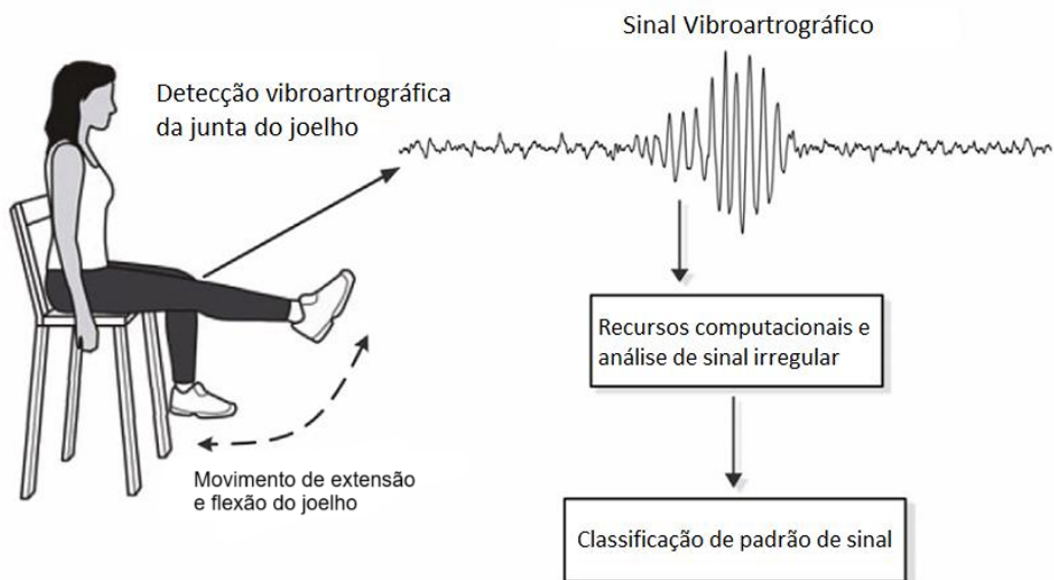
O procedimento de aquisição dos dados vibroartrográficos é realizado durante consultas com pacientes voluntários no Hospital do Trabalhador, de Curitiba. Esses pacientes encontram-se em fase pré-operatória, ou seja, possuem suas lesões já confirmadas por métodos diagnósticos independentes (seja por artroscopia, ressonância magnética ou algum outro), conforme relatado por Funaki e Lorentz (2022).

O posicionamento dos sensores no joelho do paciente segue o mesmo padrão descrito por Funaki e Lorentz (2022). Tal abordagem permite a continuidade na utilização dos dados coletados desde as fases iniciais do projeto, contribuindo para o aumento do número de amostras disponíveis para o treinamento e a validação dos modelos de identificação. Portanto, três módulos do elemento sensor inercial MPU-9250 são utilizados, sendo posicionados da seguinte maneira: um fixado na linha média da patela, e os outros dois ficam nos côndilos femorais, nas regiões medial e lateral do joelho.

Como o processamento dos dados parte da premissa de que os sinais de vibração

serão aproximadamente periódicos, é necessário um breve treinamento prévio com os pacientes antes da coleta propriamente dita. Nesse treinamento inicial, o médico orienta o paciente com comandos específicos e executa ciclos de demonstração. Assim, possibilitando que o paciente compreenda a sequência completa de cada movimento a ser executado, com o objetivo de melhorar o processo de aquisição, conforme ilustrado na Figura 6.

**Figura 6 – Movimento realizado pelo paciente durante o processo de treinamento e aquisição dos dados.**



Fonte: Adaptado de Wu *et al.* (2016).

Os movimentos realizados consistem, basicamente, na extensão<sup>1</sup> e flexão<sup>2</sup> do joelho, cada um com duração predefinida de 5 segundos, executados em velocidade constante. O tempo é indicado ao paciente por meio de contagem em voz alta realizada pelo examinador, que também pode fornecer instruções adicionais durante o movimento, sempre que considerar necessário para aprimorar o processo de aquisição dos dados. Ao todo, durante a aquisição válida dos dados, são realizados 3 ciclos do movimento completo, totalizando 30 segundos (SILVA; CAMARGO, 2025).

Para viabilizar e padronizar o processo de aquisição, Silva e Camargo (2025) também desenvolveram uma interface homem-máquina (IHM)<sup>3</sup>, que é utilizada pelo médico ao longo do exame (desde as etapas iniciais de coleta de informações até a aquisição efetiva dos dados).

<sup>1</sup> É o movimento de alongamento da articulação do joelho, aumentando o ângulo entre a coxa e a perna.

<sup>2</sup> É o movimento de dobramento da articulação do joelho, diminuindo o ângulo entre a coxa e a perna

<sup>3</sup> Uma interface homem máquina (IHM) é um sistema ou dispositivo desenvolvido de forma a facilitar a interação entre um humano e uma máquina, seja para controle, monitoramento ou até a operação de ações em sistemas (VILLANI *et al.*, 2017).

Antes do procedimento de aquisição válida, o processo tem início com a coleta de informações referentes ao paciente. Essas são informações consideradas relevantes por Silva e Camargo (2025) em seus estudos, tais como: nome, idade, gênero, peso, lesão diagnosticada, método diagnóstico utilizado e demais características. Essa etapa é realizada por meio de uma tela inicial da interface homem-máquina (IHM), conforme ilustrado na Figura 7 a seguir:

**Figura 7 – Tela de coleta de informações e configuração da aquisição da IHM desenvolvida por Silva e Camargo (2025).**

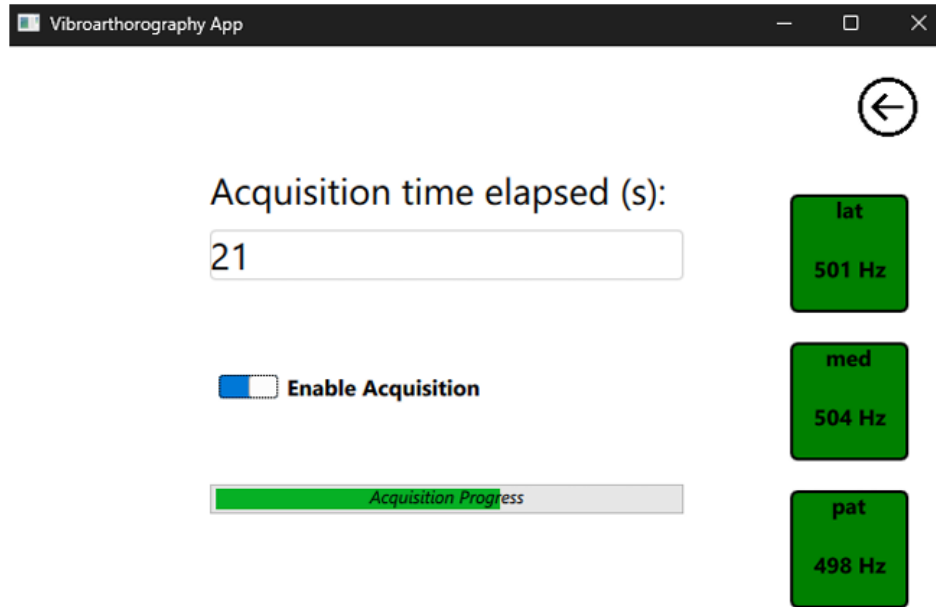
The screenshot shows the 'Vibroarthrography App' window. It features a 'Try Reconnect' button and a Wi-Fi icon with a red 'X' indicating a connection issue. The form is organized into two columns. The left column contains text input fields for 'Doctor', 'Patient Id', 'Weight (kg):', 'Age:', 'Height (m):', 'Diagnosys method:', and 'What injury:'. The right column contains a numeric input for 'Acquisition Time Trigger (s):' with the value '38', and dropdown menus for 'Gender:' (set to 'Male'), 'Knee:' (set to 'Right'), and 'Knee Hair:' (set to 'Very low'). Below these is a section for folder selection with the text 'Please, select the destination folder:', a text box showing 'C:/Users/Usuarios/OneDrive/Documentos/proje', and a 'Browse' button. A large green 'Start Operation' button is positioned at the bottom right of the form.

**Fonte: (SILVA; CAMARGO, 2025).**

Após os testes iniciais realizados com o paciente e o preenchimento das informações, o examinador dá início ao processo de aquisição por meio da IHM. Então, pressionando o botão verde “*Start Operation*”, presente na Figura 7 para prosseguir à tela de aquisição (Figura 8), o procedimento é iniciado assim que o paciente disser que está preparado. Os dados são aquisitados por um tempo médio de 38 segundos, de forma que, idealmente, o paciente deve permanecer imóvel durante os 5 segundos iniciais e os 3 segundos finais.

Passado esse tempo inicial, o paciente começa o movimento de extensão, seguido da flexão do joelho, da mesma forma como realizado nos testes preparatórios. Durante a execução, o médico continua realizando a contagem em voz alta e pode fornecer orientações adicionais, caso identifique necessidade de ajustes para melhorar a qualidade do movimento e, consequentemente, do processo de aquisição.

**Figura 8 – Tela de aquisição da IHM desenvolvida por Silva e Camargo (2025), com indicadores de status.**



Fonte: (SILVA; CAMARGO, 2025).

A tela exibida na Figura 8 é apresentada durante o processo de uma aquisição válida dos dados. Nela, é possível que o examinador acompanhe o tempo decorrido, de forma a facilitar sua contagem, bem como monitorar a frequência de aquisição em cada um dos módulos sensores inerciais MPU-9250. Assim, é possível identificar eventuais falhas de leitura ou desconexões dos sensores durante o processo.

## 4.2 Características dos Dados Coletados

Os dados, coletados pelos dispositivos sensores, seguindo o procedimento explicado na Seção 4.1, são armazenados em dois arquivos: um “.csv” e um “.json”. O arquivo “.json”, com exemplo genérico na Figura 9, contém as informações referentes ao paciente recolhidas durante o exame. Já o arquivo “.csv” armazena os dados de vibração registrados por cada um dos módulos sensores inerciais MPU-9250, conforme exemplo parcial apresentado na Figura 10.



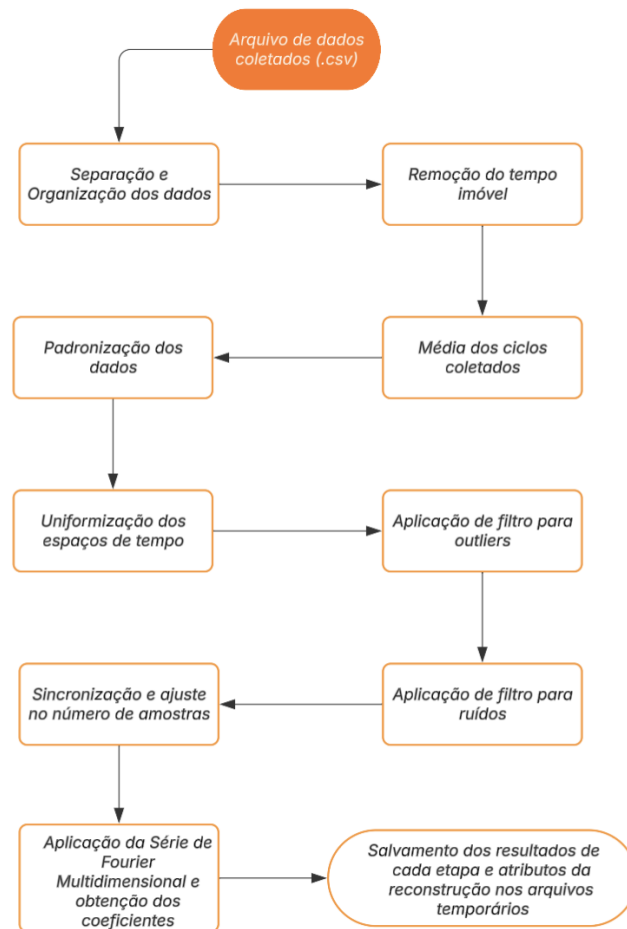
Conforme explicado pelos autores, o arquivo “.csv” contém dados já armazenados com os valores convertidos para unidades físicas: aceleração linear em  $m/s^2$ , medida pelo acelerômetro, e velocidade angular em  $^\circ/s$ , medida pelo giroscópio. A partir da análise do formato atual desses arquivos, busca-se compreender como os dados estão organizados e de que forma podem ser utilizados para a reconstrução do sinal original.

### 4.3 Organização e Processamento dos Dados

Nessa etapa inicial, todo o procedimento envolvendo os dados coletados (abrangendo desde a interpretação e o pré-processamento até o processamento propriamente dito) foi desenvolvido em MATLAB®. À medida que o projeto avançou, houve uma transição para o Visual Studio Code, com foco na utilização de Python e de suas bibliotecas voltadas ao processamento de sinais. Essa mudança se justifica pelo fato de que a linguagem Python constitui um ambiente de programação amplamente adotado em aplicações científicas e de engenharia, oferecendo uma estrutura orientada a objetos e suporte eficiente a operações vetoriais e matriciais por meio de pacotes como *NumPy*, *Pandas* e *SciPy*. Além disso, ferramentas como *Matplotlib* e *Plotly* permitem uma visualização gráfica eficiente dos resultados obtidos, contribuindo para a interpretação dos fenômenos analisados (QUINTANILHA; RUPPERT, 2023).

O fluxograma apresentado na Figura 11 ilustra as principais etapas desenvolvidas na parte inicial do tratamento dos dados. Nele é ilustrada desde a organização estrutural das informações contidas no arquivo “.csv” até a reconstrução dos sinais originais por meio da Série de Fourier, incluindo também a visualização gráfica dos resultados e comparações entre diferentes configurações.

**Figura 11 – Fluxograma que representa as etapas para o processamento dos dados.**



**Fonte: Autoria própria.**

#### 4.3.1 Análise Estrutural dos Dados

Para dar início ao tratamento dos dados, foi analisada a estrutura presente no arquivo “.csv” ilustrado na Figura 10. Observou-se que, em um primeiro momento, o arquivo não apresenta uma organização intuitiva, sendo pouco evidente a correspondência entre as colunas, os eixos de medição e as posições dos sensores. Assim, a etapa inicial do tratamento consistiu na separação e organização dos dados em uma estrutura mais adequada para o processamento posterior.

Esse procedimento foi implementado em linguagem Python, utilizando a biblioteca *Pandas* para leitura e manipulação tabular dos dados. As informações foram reorganizadas com base nos nomes das colunas, que indicam o tipo de variável (*ax*, *ay*, *az*, *gx*, *gy*, *gz*), e na coluna “*Who*”, responsável por identificar a posição do sensor (*lat*, *med* ou *pat*).

A partir dessa análise, os dados foram estruturados em uma lista de dicionários<sup>6</sup>, na qual

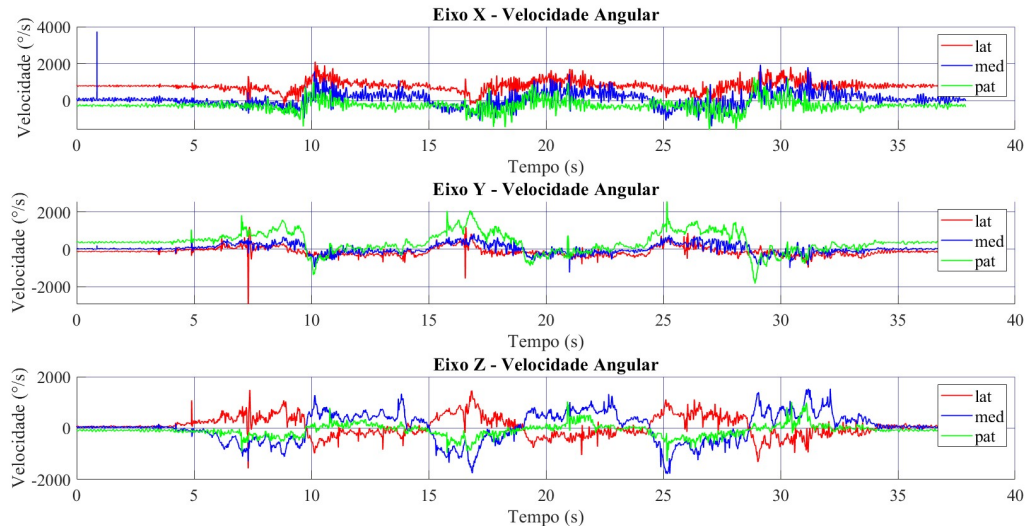
<sup>6</sup> Estrutura de dados que armazena informações em pares chave–valor, permitindo acesso rápido aos valores por

cada dicionário representa um sinal individual. Cada elemento contém o nome do canal “*name*”, o vetor de tempos “*t*” e os respectivos valores medidos “*data*”. Essa organização modular tornou o conjunto de dados mais legível e facilitou o acesso e a manipulação dos sinais nas etapas subsequentes do processamento de dados, permitindo que cada série temporal fosse tratada de forma independente.

As Figuras 12 e 13 apresentam os sinais brutos de velocidade angular e aceleração linear, coletados pelos sensores inerciais nas posições lateral (lat), medial (med) e patelar (pat). O objetivo dessa visualização inicial é verificar a coerência das escalas, identificar ruídos e *outliers*, além de reconhecer padrões que possam influenciar as etapas subsequentes de processamento.

Na Figura 12, que ilustra os sinais de velocidade angular obtidos pelos giroscópios nos três eixos cartesianos, são exibidas variações expressivas entre sensores e entre eixos. É evidente a presença de ruídos de alta frequência, saltos abruptos e valores atípicos, os quais comprometem a qualidade do sinal bruto e indicam a necessidade de filtragem, correção de desvios e padronização.

**Figura 12 – Sinais de velocidade angular originais aqisitados pelos giroscópios.**

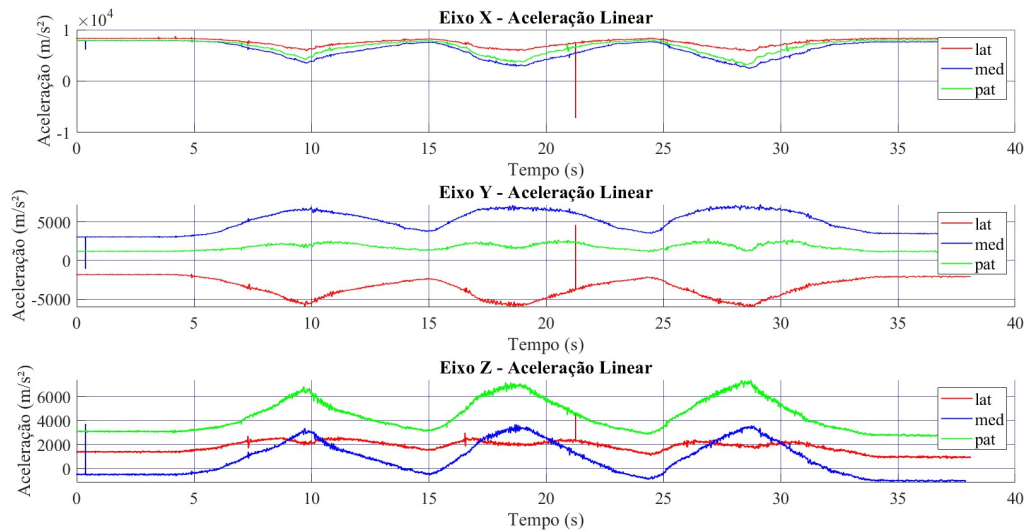


**Fonte: Autoria própria.**

A Figura 13 mostra os sinais de aceleração linear captados pelos acelerômetros. Embora visualmente mais estáveis, esses dados ainda apresentam variações de amplitude entre as posições dos sensores e padrões característicos que requerem atenção. Essas observações reforçam a importância de um pré-processamento criterioso, visando à obtenção de métricas confiáveis para meio de suas chaves.

análise quantitativa.

**Figura 13 – Sinais de aceleração linear originais adquiridos pelos acelerômetros.**



**Fonte: Autoria própria.**

A análise inicial dos sinais evidencia a presença de *outliers*, ruídos e diferenças significativas nas escalas, especialmente nos dados provenientes dos giroscópios. Essas características são devidamente consideradas na etapa de pré-processamento.

#### 4.3.2 Etapas de Pré-Processamento

Conforme abordado na Seção 2.2, antes da aplicação efetiva da Série de Fourier aos sinais coletados, há a necessidade de realizar um pré-processamento nos dados originais. Sendo assim, nessa etapa buscou-se preparar os sinais para que possam ser analisados de forma mais eficaz, aplicando-se procedimentos como: (i) padronização, para ajustar a amplitude e evitar distorções na análise; (ii) interpolação, para garantir o espaçamento uniforme dos dados; e (iii) filtragem, com a finalidade de reduzir ruídos e eliminar possíveis medições inválidas (CRENNA *et al.*, 2021).

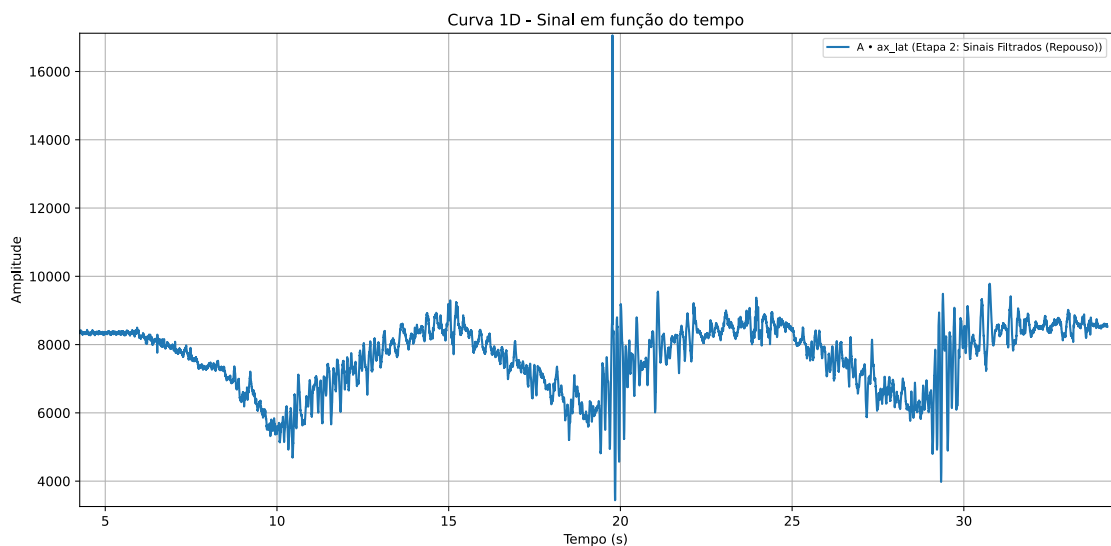
##### 4.3.2.1 Movimento

A primeira etapa de pré-processamento, antes da aplicação da Série de Fourier sobre o conjunto de dados, foi a realização da filtragem para os momentos em que o paciente permaneceu com o joelho imóvel durante a coleta. Conforme explicado na Seção 4.1, o procedimento de aquisição determinava que o paciente deve permanecer em repouso pelos 5 segundos iniciais,

seguido de 3 ciclos de 10 segundos para o movimento completo de extensão e flexão, totalizando um tempo de aquisição de 38 segundos. Dessa forma, idealmente, nos 3 segundos finais, o joelho do paciente também deve permanecer imóvel.

Dessa forma, buscou-se identificar e extrair os períodos de imobilidade, a fim de evitar que esses trechos fossem considerados nas análises de frequência. Para garantir que essa etapa não gere impactos nos dados correspondentes ao ciclo de execução do movimento, foi implementado um filtro que recorta um intervalo contínuo de tempo fixo, definido como parâmetro da função e configurado, por padrão, nos 4 segundos iniciais e finais. A Figura 14 ilustra um exemplo de sinal resultante após o processo de remoção do tempo imóvel.

**Figura 14 – Sinal de aceleração do sensor lateral em x após remoção do tempo sem movimento (4s).**



**Fonte: Autoria própria.**

#### 4.3.2.2 Média dos ciclos

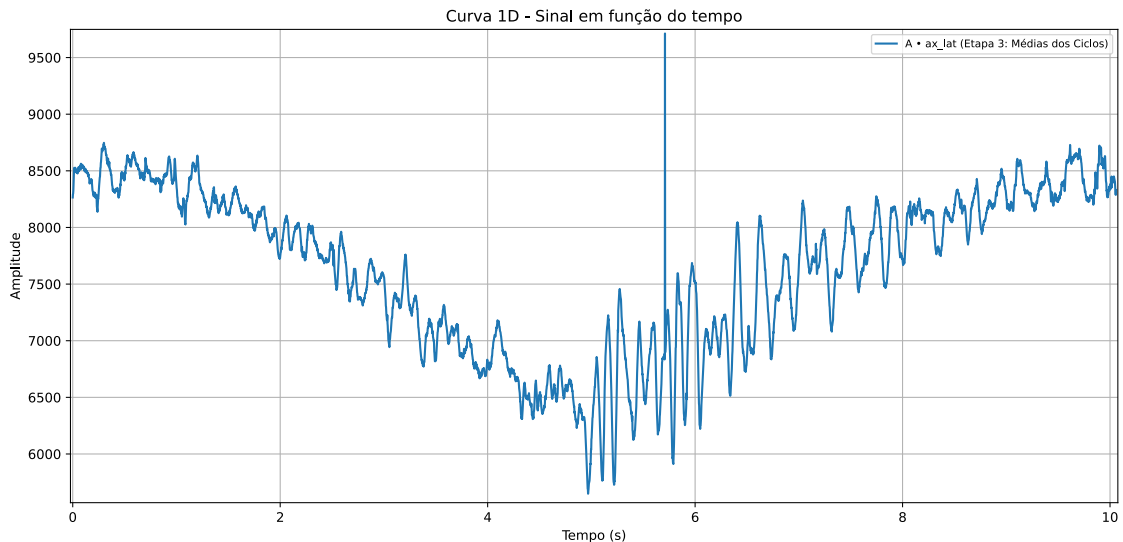
Com os dados referentes apenas ao período de movimento, foi realizada uma média ponto a ponto entre os três ciclos de execução. O objetivo da etapa é garantir que os contornos fechados obtidos após a reconstrução por Fourier representem um único ciclo, evitando a formação de *loops* no espaço de fases<sup>7</sup>. Para isso, foi desenvolvida uma função que calcula a média ponto a ponto entre os  $n$  ciclos de movimento coletados.

Dessa forma, a análise e a reconstrução passam a considerar apenas o valor médio de cada ponto, o que reduz o custo computacional e atua como um filtro inicial para atenuar ruídos

<sup>7</sup> O espaço de fases consiste na representação conjunta de variáveis dinâmicas relacionadas, como aceleração linear e velocidade angular, permitindo visualizar a evolução do movimento como uma trajetória no plano (SOUDAN; DIERCKX, 1979).

e valores de possíveis *outliers* aleatórios. A Figura 15 ilustra um exemplo de sinal resultante após o processo de média dos ciclos.

**Figura 15 – Sinal de aceleração do sensor lateral em x após a média dos 3 ciclos.**



**Fonte: Autoria própria.**

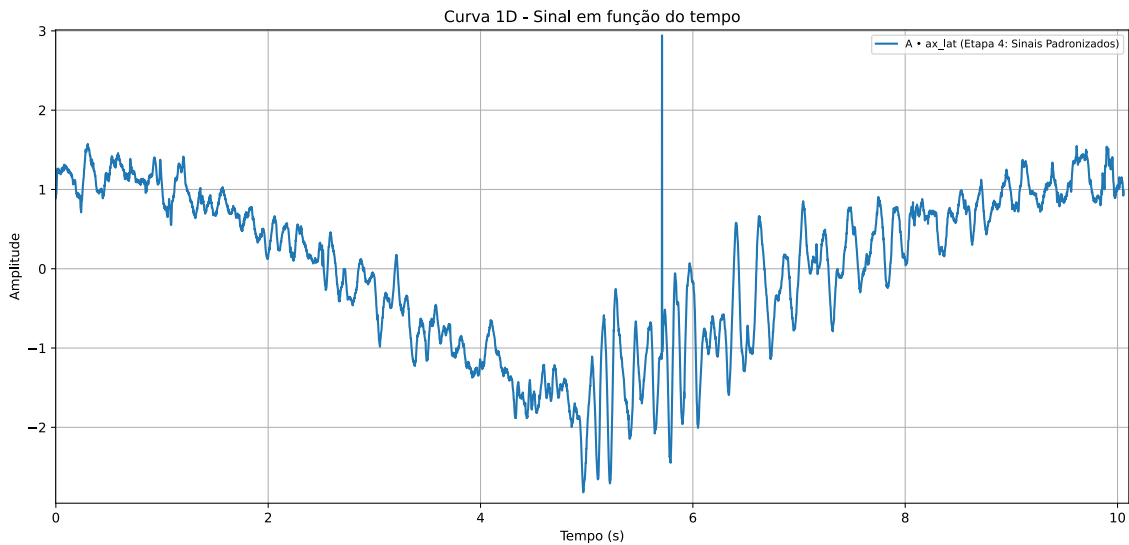
#### 4.3.2.3 Padronização

A partir dos dados ajustados para representar um único ciclo do sinal, a etapa seguinte consistiu na padronização dos dados. Como os sinais coletados apresentam diferentes escalas de amplitude, tornou-se necessário realizar ajustes nos valores, uma vez que os sinais de diferentes sensores e eixos serão posteriormente comparados. Com isso, evitam-se distorções nas análises e garante-se que todos os dados contribuam com pesos equivalentes.

Então, pela observação de que os sinais originais assumem tanto valores positivos quanto negativos e possuem alguns *outliers*, optou-se pela padronização do tipo *z-score*<sup>8</sup>, que preserva a natureza oscilatória do sinal original. A Figura 16 ilustra um exemplo de sinal resultante após o processo de padronização dos dados.

<sup>8</sup> Esse método consiste em centralizar os dados em torno da média e escalá-los pela sua variância, resultando em uma nova distribuição com média zero e desvio padrão unitário (HENDARI *et al.*, 2021).

**Figura 16 – Sinal de aceleração do sensor lateral em x após a padronização.**



**Fonte: Autoria própria.**

#### 4.3.2.4 Uniformização

Durante a análise inicial dos dados, observou-se que as amostras coletadas não possuíam um registro temporal individual (*timestamp*<sup>9</sup>), pois eram transmitidas em pacotes de 100 valores por variável do sensor. Diante disso, foi necessário uniformizar os intervalos de tempo entre as amostras, a fim de garantir consistência na representação temporal dos sinais e permitir uma análise adequada.

Garantir um espaçamento uniforme ( $T_s$ ) entre os pontos é essencial não apenas para a aplicação da série de Fourier, mas também para a execução de filtros e etapas de pré-processamento subsequentes. Esses procedimentos se baseiam em operações como integrais e somatórios, os quais exigem um passo de tempo constante ( $T_s$ ) (JAMES, 2002).

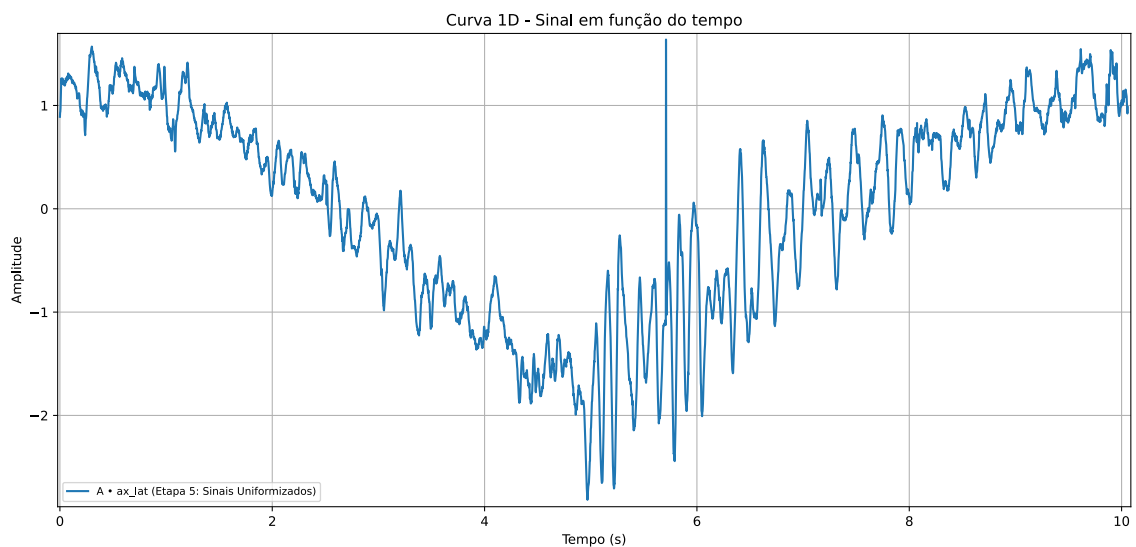
Para garantir a consistência temporal dos dados, foi necessário realizar uma uniformização dos intervalos entre as amostras, utilizando interpolação com base na frequência de amostragem de  $500\text{ Hz}$ <sup>10</sup>. A partir da diferença entre os tempos registrados da primeira e da última amostra, foi determinado o tempo total de coleta. Com esse valor e a frequência de amostragem que foi definida, calculou-se o número esperado de amostras. Assim, foi gerado um vetor de tempo uniformemente espaçado, indo do instante inicial até o final da coleta, contendo exatamente o número de pontos estimado.

<sup>9</sup> *Timestamp* é o registro do instante exato de tempo associado a uma medida.

<sup>10</sup> Valor definido por Silva e Camargo (2025) como a frequência ajustada no software embarcado para a aquisição dos valores dos sensores.

Para o tratamento de possíveis lacunas ou redundâncias nos dados, realizou-se a interpolação dos sinais por meio de funções da biblioteca *SciPy*, utilizando o método *spline*<sup>11</sup>. Embora o sinal seja discreto (resultante de uma amostragem uniforme de um fenômeno originalmente contínuo), a interpolação por *spline* permite reconstruir uma estimativa suave entre os pontos amostrados. Dessa forma, obtém-se uma representação mais regular do sinal, sem introduzir variações abruptas que poderiam distorcer a análise subsequente. A Figura 17 ilustra um exemplo de sinal resultante após a etapa de uniformização.

**Figura 17 – Sinal de aceleração do sensor lateral em x após a uniformização.**



**Fonte: Autoria própria.**

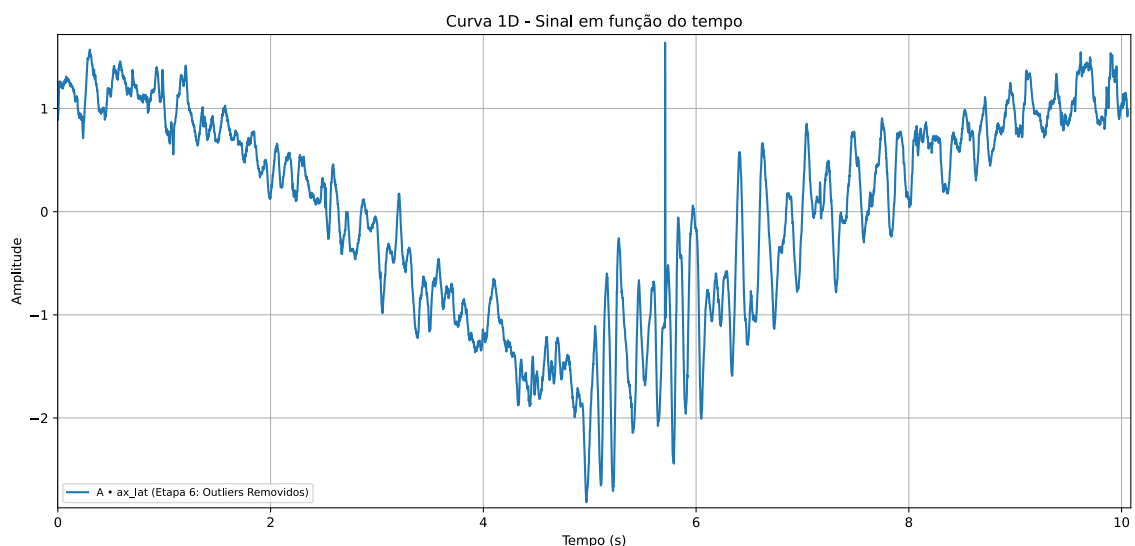
#### 4.3.2.5 Filtro de Outliers

Com os dados que já se tornaram uniformemente espaçados e padronizados, iniciou-se a aplicação de filtros, com o objetivo de atenuar ruídos e eliminar medidas inconsistentes. Nesta etapa, foram identificados e removidos os *outliers*, com base em um limite de valor absoluto máximo considerado aceitável para os dados padronizados. Esse *threshold* foi definido com o objetivo de eliminar amostras que poderiam comprometer a qualidade da análise, por representarem variações extremas não condizentes com o comportamento esperado dos sinais. Como os dados removidos precisavam ser substituídos por novos valores, foi realizado o preenchimento das lacunas por meio de interpolação, utilizando funções *spline*, já apresentado anteriormente.

<sup>11</sup> A interpolação *spline* consiste em aproximar um conjunto de pontos por meio de polinômios definidos por trechos, garantindo suavidade e continuidade de derivadas entre os segmentos (The MathWorks, Inc., 2025).

O limite para detecção de *outliers* foi definido, em uma primeira abordagem, de forma empírica, com base na inspeção visual dos dados padronizados. Nessa etapa preliminar, verificou-se que valores que ultrapassavam determinado limiar exibiam comportamento discrepante, o que motivou sua exclusão para evitar distorções na análise. Posteriormente, esse procedimento foi substituído por uma abordagem estatística, utilizando a mediana e o desvio absoluto da mediana (MAD). Esse método permite estabelecer um *threshold* de detecção fundamentado no comportamento local do sinal, tornando o processo não dependente de inspeção visual (GOIC *et al.*, 2012). A Figura 18 ilustra um exemplo de sinal resultante após a aplicação do filtro de *outliers*.

**Figura 18 – Sinal de aceleração do sensor lateral em x após o filtro de outliers.**



**Fonte: Autoria própria.**

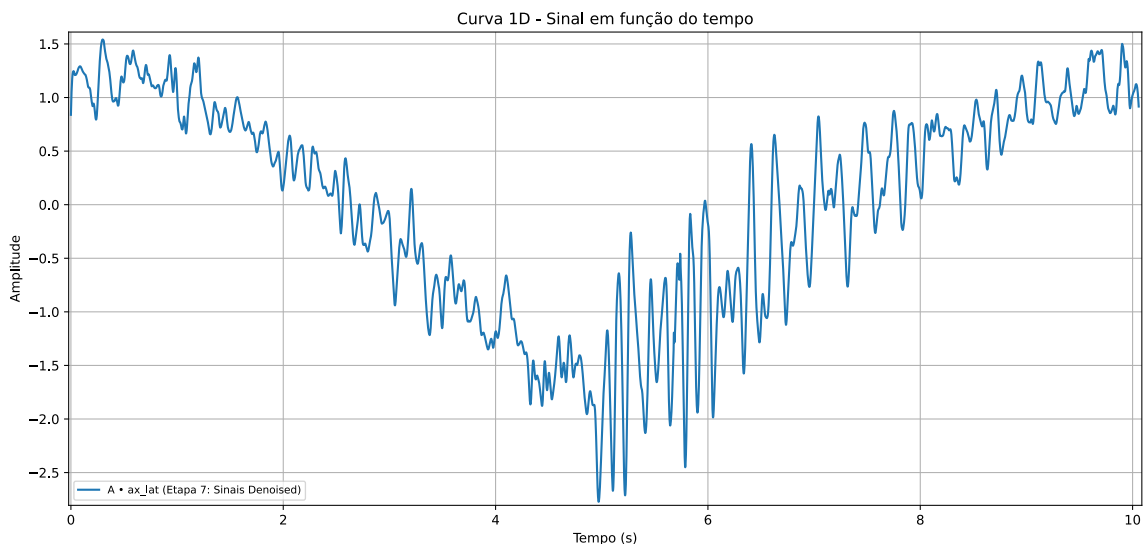
#### 4.3.2.6 Filtro de Ruídos

O segundo passo foi realizar a aplicação de um filtro para redução de ruídos no sinal. Considerando que o objetivo principal era preservar as características essenciais do sinal original, atenuando apenas os ruídos mais evidentes, optou-se pela utilização do filtro Savitzky-Golay.

A escolha desse filtro se justifica pelo fato de que, diferentemente de outros filtros como o FIR, o Savitzky-Golay promove uma suavização sem comprometer a integridade das principais características do sinal. A abordagem de ajustes polinomiais e suavização digital permite reduzir o ruído, bem como preservar detalhes importantes (RAHMAN *et al.*, 2019). Esse aspecto o torna útil para a aplicação na vibroartrografia, pois a forma do sinal precisa ser mantida para análises subsequentes.

Para aplicar o filtro no ambiente de programação, utilizou-se a função *savgol\_filter()* da biblioteca *SciPy*. Assim como em outras implementações, foi necessário definir dois parâmetros principais: a ordem polinomial do filtro e o tamanho da janela de aplicação. A janela estabelece quantos pontos do sinal são considerados a cada iteração (devendo ser um número ímpar, pois o ponto central é tomado como referência), enquanto a ordem define o grau do polinômio ajustado em cada subconjunto de dados. Com base em análises empíricas, adotou-se uma ordem polinomial igual a 4 e uma janela de 31 amostras, valores que proporcionaram uma suavização eficaz sem distorcer as características originais dos sinais. A Figura 19 ilustra um exemplo de sinal resultante após a aplicação do filtro de ruídos.

**Figura 19 – Sinal de aceleração do sensor lateral em x após o filtro de ruídos.**



**Fonte: Autoria própria.**

#### 4.3.2.7 Ajuste e Sincronização dos Sinais

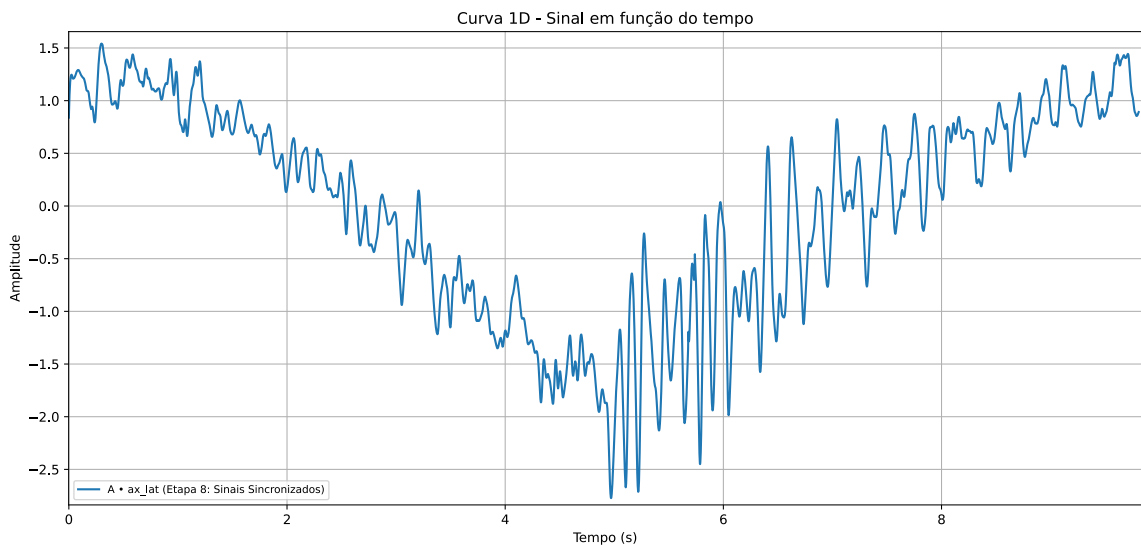
Observou-se que cada variável presente no arquivo de dados original não apresentava, necessariamente, o mesmo número de amostras. Ainda, devido às operações realizadas durante o pré-processamento, como a remoção dos períodos de imobilidade, os sinais poderiam não estar iniciando e terminando no mesmo instante de tempo.

Dessa forma, para possibilitar a reconstrução conjunta dos sinais, considerando as  $n$  dimensões analisadas, foi necessário realizar o ajuste do comprimento dos vetores e a sincronização dos tempos de início e fim. Então, inicialmente, foi realizada a sincronização temporal dos sinais, ajustando o intervalo de análise para o período compreendido entre o maior tempo de início e o menor tempo de fim entre todas as variáveis amostradas. Em seguida, o número de

amostras foi padronizado, definindo como referência a menor quantidade de amostras encontrada entre o conjunto dos sinais.

Assim, essa etapa assegurou a consistência temporal entre os sinais e a padronização do número de amostras, viabilizando tanto as análises multivariadas subsequentes, quanto as comparações diretas entre as diferentes variáveis. A Figura 20 mostra um exemplo de sinal após a sincronização, etapa que encerra o pré-processamento dos dados.

**Figura 20 – Sinal de aceleração do sensor lateral em x após ajuste e sincronização.**



**Fonte: Autoria própria.**

#### 4.4 Aplicação da Série de Fourier e Reconstrução do Sinal

Com as etapas de pré-processamento concluídas, foram obtidos os vetores contendo os valores adquiridos e seus respectivos tempos. Para ser possível a aplicação da Série de Fourier, foi necessário determinar o período de cada um dos sinais coletados. Portanto, optou-se por adotar o tempo total ao final da última etapa de pré-processamento como referência para o período a ser utilizado na análise de cada sinal pela Série de Fourier. Essa decisão se deve ao fato de que já é realizada a média entre o  $n$  ciclos de aquisição, e o tempo imóvel já é ajustado pelas etapas de pré-processamento, de forma que ao final deste processo o sinal resultante já possui um único ciclo de movimento.

Assim, foi desenvolvida uma função que aplica a Série de Fourier discretizada multi-dimensional aos sinais. Nessa abordagem, todos os componentes do sinal (organizados em  $D$  dimensões após a sincronização) são processados conjuntamente sobre a mesma base temporal.

A função retorna os coeficientes  $A_0$ ,  $A_h$  e  $B_h$ , organizados de forma compatível com o caso multidimensional<sup>12</sup>, e fornece uma prévia do sinal reconstruído. Para a execução dessa função, além dos vetores do sinal original  $x$ , do tempo  $t$  e do período  $T$ , é necessário fornecer o número de harmônicos considerados no somatório<sup>13</sup>, o que define o grau de truncamento da série.

Então, foram realizadas algumas análises da reconstrução para diferentes números de harmônicos. Como Funaki e Lorentz (2022) utilizaram 50 harmônicos para a realização dos seus estudos, esse valor foi adotado como ponto de partida. No entanto, observou-se que a reconstrução inicial apresentava uma certa margem de erro em relação ao sinal original. Diante disso, buscando inicialmente ter a comprovação de que o algoritmo desenvolvido funcionava e realizava a reconstrução correta do sinal, o número de harmônicos considerados na série foi progressivamente aumentado, buscando aproximar o sinal reconstruído do original, bem como manter um desempenho de processamento do *software* adequado.

Portanto, definiu-se o número de harmônicos a ser utilizado nessa etapa inicial para comprovação da eficácia da reconstrução como 1000. Esse valor mostrou-se suficiente para garantir uma aproximação satisfatória entre o sinal reconstruído e o sinal original, sem causar impacto significativo no tempo de execução da aplicação. Contudo, para aplicação em etapas posteriores, especialmente na classificação dos sinais por meio dos algoritmos de aprendizado de máquina, esse número de harmônicos precisou ser reduzido consideravelmente, em função das limitações de processamento computacional e da disponibilidade restrita de amostras.<sup>14</sup>

Portanto, foram então realizadas reconstruções dos sinais em diferentes formatos ponto a ponto: (i) reconstrução individual de cada variável, permitindo a análise isolada de suas componentes ao longo do tempo e comprovação da eficácia da reconstrução; (ii) reconstrução da trajetória bidimensional, obtida pela associação de pares formados pela posição do sensor e pelo eixo cartesiano correspondente; (iii) reconstrução da trajetória bidimensional, resultante da associação de triplas de posição do sensor e eixo cartesiano correspondente.

As análises detalhadas dos resultados obtidos por meio dessas diferentes reconstruções serão apresentadas no Capítulo 5

<sup>12</sup> O termo constante  $A_0$  é representado por um vetor de dimensão  $D$ , enquanto  $A_h$  e  $B_h$  são matrizes de dimensão  $H \times D$ , nas quais cada coluna correspondendo a um componente do sinal.

<sup>13</sup> É preciso determinar o número de harmônicos no somatório da série até seu truncamento. Valores maiores resultam em uma reconstrução mais fiel ao sinal original, conforme discutido na Seção 2.2.1.

<sup>14</sup> O número de harmônicos utilizados está diretamente relacionado à quantidade de atributos fornecidos aos classificadores, uma vez que todas as componentes harmônicas  $A_0$ ,  $A_h$  e  $B_h$  são utilizadas como variáveis de entrada.

#### 4.4.1 Cálculo dos Erros Nas Reconstruções

A avaliação da qualidade das reconstruções é realizada por meio do *Normalized Root Mean Square Error* (NRMSE), uma métrica amplamente empregada em contextos de regressão e reconstrução de sinais (GOMES *et al.*, 2025). O RMSE quantifica o erro médio entre os valores reais  $x_n$  e os valores reconstruídos  $\hat{x}_n$ , atribuindo maior peso a discrepâncias de maior magnitude em virtude da operação de quadrado aplicada aos resíduos. Sua formulação é expressa por:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (x_n - \hat{x}_n)^2}, \quad (28)$$

em que  $N$  representa o número total de amostras,  $x_n$  representa o valor real e  $\hat{x}_n$  representa o valor reconstruído. Adicionalmente, utiliza-se o erro percentual normalizado  $\text{RMSE}_{\%}$ , a partir da variância  $\sigma_x$ , que permite uma interpretação relativa da precisão da reconstrução, tornando a métrica independente da escala do sinal. NRMSE é definido como:

$$\text{NRMSE} = \frac{\text{RMSE}}{\sigma_x}, \quad (29)$$

Essa normalização é especialmente relevante para comparações entre diferentes sinais ou condições experimentais, uma vez que fornece uma métrica adimensional da qualidade da reconstrução (GOMES *et al.*, 2025).

A escolha do NRMSE como métrica de avaliação justifica-se por sua elevada sensibilidade a discrepâncias de maior magnitude. No contexto deste trabalho, o RMSE representa uma estimativa consistente da fidelidade das reconstruções em relação aos sinais originais, permitindo quantificar de forma precisa os desvios existentes.

#### 4.5 Desenvolvimento e Treinamento dos Classificadores

Uma etapa do desenvolvimento consistiu na definição dos algoritmos de aprendizado de máquina mais apropriados ao problema em questão, na preparação dos dados utilizados para o treinamento e na escolha dos parâmetros a serem otimizados durante o ajuste dos modelos. Esse processo envolveu, ainda, a validação do desempenho dos classificadores, com o objetivo de assegurar sua capacidade de generalização. Cada modelo foi concebido considerando as particularidades do conjunto de dados disponível, caracterizado por um número limitado de amostras,

alta dimensionalidade resultante da decomposição harmônica e a presença de inconsistências estruturais nos arquivos originais.

Durante o desenvolvimento dos classificadores, foi inicialmente avaliada a possibilidade de utilizar um número maior de harmônicos na decomposição dos sinais vibroartrográficos. No entanto, optou-se por limitar essa quantidade a dez harmônicos, de modo a manter uma relação mais equilibrada entre o número de atributos e a quantidade de amostras disponíveis. Considerando as dezoito variáveis analisadas (sensores acelerômetros e giroscópios, distribuídos em três eixos e três posições), cada harmônico gera dois coeficientes ( $A_h$  e  $B_h$ ) correspondentes aos termos de seno e cosseno, além de um termo médio ( $A_0$ ). Assim, o conjunto total de características alcançou 378 para dez harmônicos<sup>15</sup>. Caso o número fosse ampliado para cem harmônicos, o total atingiria 3.618 atributos, um valor significativamente superior ao número de amostras válidas disponíveis. De acordo com Subramanian e Simon (2013), essa desproporção entre o número de variáveis e o tamanho da amostra tende a aumentar o risco de *overfitting*, comprometendo a capacidade de generalização dos modelos. Além disso, alguns registros apresentavam a ausência de dados em determinados sensores, o que exigiu tratamento adicional, havendo a substituição desses valores por estimativas calculadas a partir da mediana.

Ademais, é importante destacar que o conjunto de dados (*dataset*) utilizado foi construído a partir dos arquivos brutos registrados pelos sensores durante a execução dos exames. Esses arquivos foram submetidos a etapas de pré-processamento e reconstrução dos sinais por meio da Série de Fourier Multidimensional, considerando dez harmônicos, conforme mencionado anteriormente. A partir dessa reconstrução, foram extraídos os coeficientes empregados como atributos de entrada para os modelos de aprendizado. Os dados resultantes foram consolidados como atributos no *dataset*, em formato “.csv”, composto por 54 amostras válidas, que serviram de base para o treinamento e a validação dos classificadores.

Como procedimento geral para o treinamento e a seleção dos classificadores, adotou-se uma metodologia comum baseada em múltiplas execuções independentes. Inicialmente, o mesmo tratamento foi aplicado ao *dataset* em todos os algoritmos: remoção de amostras sem *label*, padronização dos *labels* em duas classes (*Normal* e *Lesão*), conversão dos atributos para formato numérico, substituição de valores infinitos por ausentes e imputação dos valores faltantes por meio

<sup>15</sup> O número total de atributos ( $A_t$ ) gerados a partir da Série de Fourier Multidimensional é determinado pela relação  $A_t = (N_{\text{sensores}} \cdot N_{\text{posições}} \cdot N_{\text{eixos}}) \cdot (2h + 1)$ , em que  $N_{\text{sensores}}$  representa o número de sensores utilizados,  $N_{\text{posições}}$  corresponde às diferentes posições de coleta dos sinais,  $N_{\text{eixos}}$  indica o número de eixos cartesianos considerados (geralmente  $x$ ,  $y$  e  $z$ ), e  $h$  é o número de harmônicos empregados na reconstrução. O termo  $(2h + 1)$  refere-se às componentes harmônicas  $A_0$ ,  $A_h$  e  $B_h$  que compõem o conjunto de atributos fornecidos aos classificadores.

da mediana de cada atributo. Em seguida, a mesma metodologia de treinamento foi empregada para os demais modelos, de forma a possibilitar uma comparação mais franca, mantendo idênticos o processo de busca de hiperparâmetros, o número de partições da validação cruzada, o número de simulações e os demais fatores que influenciam a construção dos classificadores.

Em cada execução, o *dataset* inicial foi dividido aleatoriamente em 75% das amostras para treinamento e 25% para teste. Apenas o subconjunto de treinamento foi utilizado na etapa de ajuste de hiperparâmetros, utilizando validação cruzada estratificada repetida (*RepeatedStratifiedKfold*, com 2 partições e 2 repetições). O método de treinamento adotado foi o *GridSearchCV*, utilizando a acurácia balanceada como métrica de otimização. Para cada classificador, foi explorado um conjunto pré-definido de combinações de hiperparâmetros e, em cada execução, o modelo selecionado foi aquele que obteve a maior acurácia balanceada média na validação cruzada interna. Em seguida, esse modelo foi reajustado com o conjunto de treinamento e avaliado na partição de teste correspondente. Após 100 execuções, obteve-se um conjunto de 100 modelos por classificador, e o modelo final adotado em cada caso foi aquele que obteve a maior acurácia balanceada no conjunto de teste, buscando selecionar o modelo que melhor generalizou os resultados.

Esse procedimento teve como objetivo maximizar a capacidade de generalização dos classificadores e permitir uma comparação adequada. Assim, reduzindo a probabilidade de ocorrência de *overfitting* ou *underfitting*, garantindo que o modelo resultante possua a combinação mais adequada de hiperparâmetros para o contexto dos dados médicos empregados.

#### 4.5.1 Algoritmo da Árvore de Decisão

O primeiro classificador desenvolvido foi por meio do uso do algoritmo de Árvore de Decisão, com o objetivo de distinguir sinais provenientes de um joelho saudável e de um joelho com lesão. O processo envolveu diferentes etapas de preparação, treinamento, validação e seleção do melhor modelo.

O modelo foi estruturado em um *pipeline*<sup>16</sup> supervisionado, implementado com a *scikit-learn* (biblioteca para treinamento de ML), tendo como estimador base o *DecisionTreeClassifier*<sup>17</sup>, configurado com ponderação automática das classes para compensar possíveis desequilíbrios amostrais. A etapa de ajuste de hiperparâmetros foi conduzida com *GridSearchCV*, que realizou

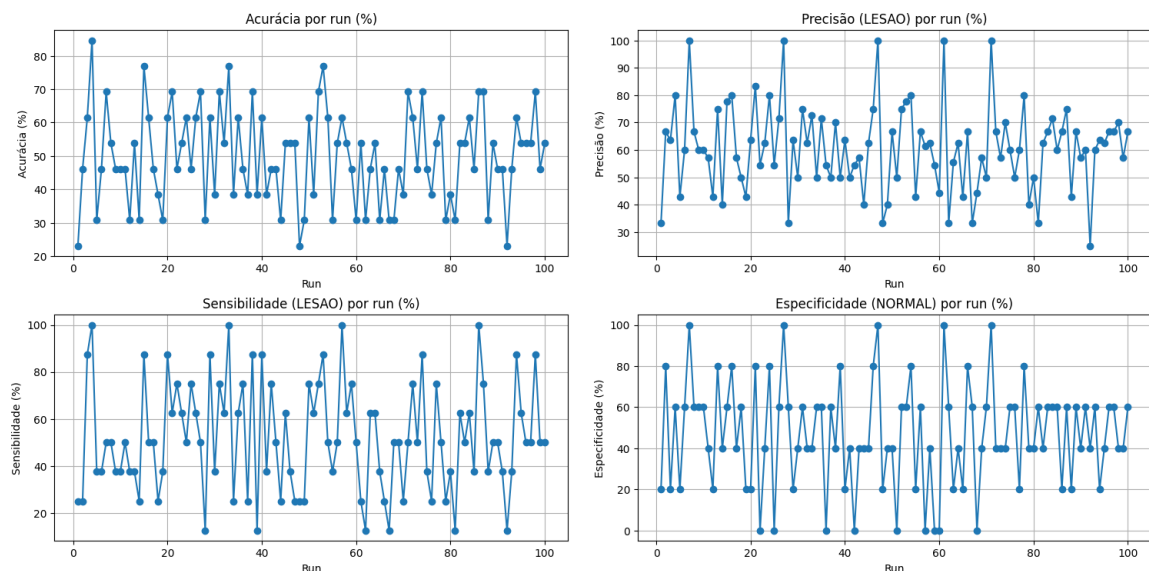
<sup>16</sup> Estrutura que encadeia todas as etapas do processo de aprendizado supervisionado.

<sup>17</sup> Algoritmo de aprendizado supervisionado do *scikit-learn* que constrói uma árvore de decisão para classificar dados com base em regras derivadas dos atributos de entrada.

busca exaustiva sobre um conjunto pré-definido de parâmetros. O espaço de busca incluiu a profundidade máxima da árvore, o número mínimo de amostras por divisão e por folha, bem como o parâmetro de poda por complexidade mínima (*ccp\_alpha*), responsável por controlar o grau de regularização do modelo. A avaliação de desempenho foi conduzida por meio da técnica de validação cruzada estratificada repetida (*RepeatedStratifiedKFold*), assegurando a representatividade de ambas as classes em cada subdivisão do conjunto de dados.

Em cada execução, foram calculadas as métricas de avaliação acurácia, sensibilidade, especificidade e precisão, sempre considerando a classe *Lesão* como positiva. Esses resultados foram registrados em um arquivo de resumo, possibilitando a comparação entre as diferentes execuções e a identificação do modelo com o melhor desempenho em termos de acurácia balanceada. O desempenho obtido pelo melhor modelo em cada uma das 100 execuções realizadas é apresentado na Figura 21, onde mostra a variação das métricas ao longo das diferentes execuções. Essa análise permite observar a sensibilidade do modelo às distintas partições do conjunto de dados, ressaltando a necessidade de múltiplas execuções para uma avaliação mais robusta de sua estabilidade.

**Figura 21 – Resumo das 100 execuções para a Árvore de Decisão.**

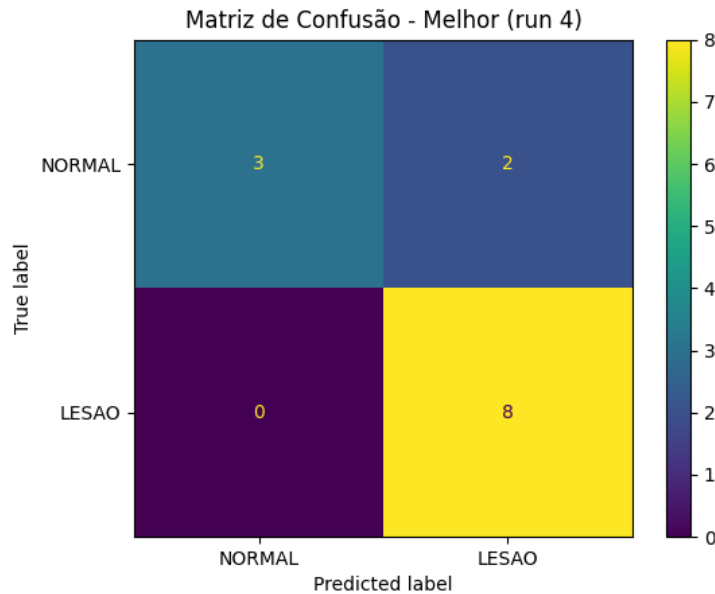


**Fonte: Autoria própria.**

O modelo com maior acurácia balanceada foi obtido na execução de número 4. Para ele, foram geradas a matriz de confusão e o relatório de classificação, permitindo uma análise detalhada dos acertos e erros de predição. A Figura 22 apresenta a matriz de confusão referente à validação do melhor modelo sobre o conjunto de teste. Por fim, esse modelo de melhor desempe-

nho foi salvo em formato binário “.pkl”, possibilitando sua reutilização em testes posteriores e sua integração direta ao aplicativo desenvolvido, sem necessidade de novo treinamento.

**Figura 22 – Matriz de confusão com os dados de teste para o modelo de Árvore de Decisão.**



**Fonte: Autoria própria.**

A partir da matriz de confusão foram obtidas as métricas tradicionais de avaliação. Os resultados de cada classificador, assim como a comparação entre as execuções, são apresentados no Capítulo 5.

#### 4.5.2 Algoritmo SVM

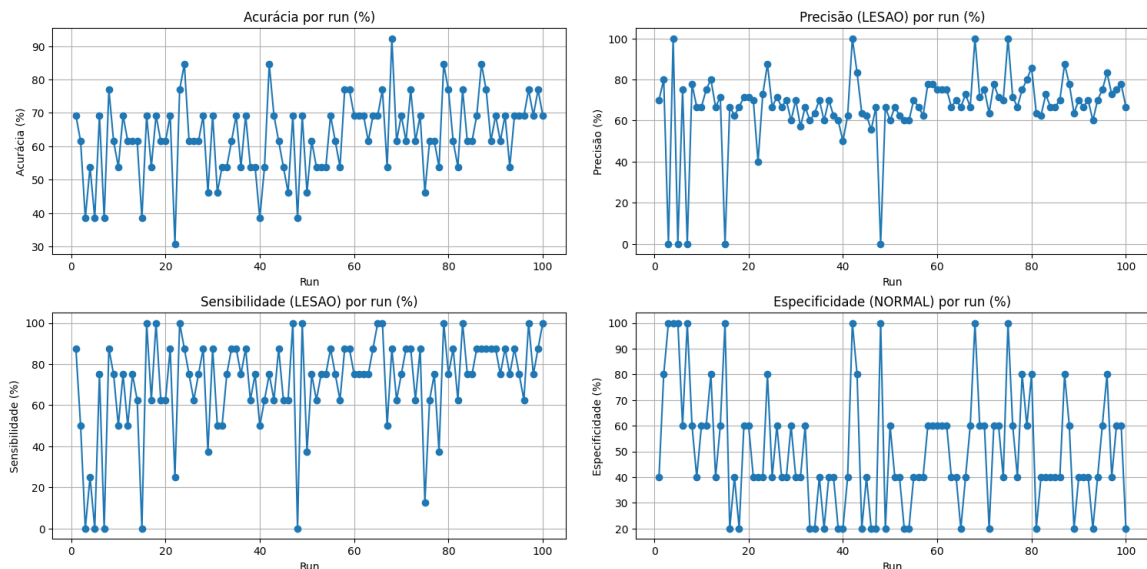
Além da Árvore de Decisão, empregou-se também o algoritmo de Máquina de Vetores de Suporte (SVM) para a tarefa de classificação binária entre amostras *Normal* e *Lesão*. Para esse classificador, o treinamento do algoritmo de aprendizado de máquina foi também organizado em um *pipeline* composto por duas etapas: a imputação e o classificador SVC (classe específica do SVM para classificação) com *class\_weight='balanced'* para mitigar possíveis desequilíbrios entre classes. O parâmetro *class\_weight='balanced'* ajusta automaticamente os pesos atribuídos a cada classe com base na frequência das amostras no conjunto de treinamento (Scikit-learn Developers, 2025).

Quanto à otimização de hiperparâmetros, foram avaliados dois núcleos: linear e RBF. Para o núcleo linear variou-se o parâmetro de regularização  $C$  em 0.01, 0.1, 1, 5, 10. Para o núcleo RBF variou-se  $C$  no mesmo conjunto e o parâmetro  $\gamma$  em scale, auto, 0.1, 0.01, 0.001. O

critério de seleção foi a acurácia balanceada. A validação empregou, conforme o padrão adotado, *RepeatedStratifiedKfold*<sup>18</sup> com 2 partições e 2 repetições, assegurando estratificação por classe em cada divisão.

Para reduzir a variância associada à aleatoriedade das partições, a seleção do melhor modelo foi baseada em  $N = 100$  execuções independentes, cada uma com uma nova divisão estratificada entre treinamento e teste (75%–25%). Em cada execução, o modelo SVM foi ajustado com os hiperparâmetros ótimos, e as métricas foram calculadas para a classe positiva no conjunto de treinamento. As métricas obtidas pelo melhor modelo em cada uma das 100 execuções realizadas são apresentadas na Figura 23.

**Figura 23 – Resumo das 100 execuções para o SVM.**

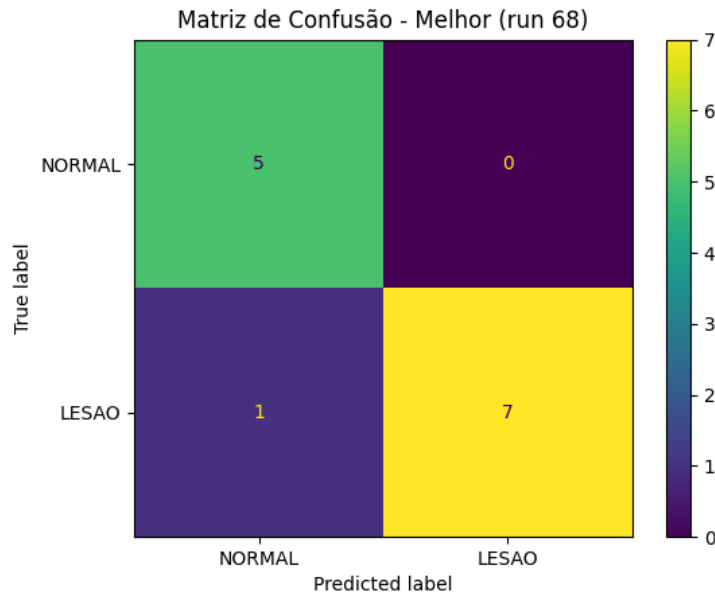


**Fonte: Autoria própria.**

Os resultados foram consolidados em um arquivo de resumo, e o modelo final foi definido como aquele com a maior acurácia balanceada no conjunto de teste entre as 100 execuções. O modelo com maior acurácia balanceada foi obtido na execução de número 68. Para ele, foram geradas a matriz de confusão e o relatório de classificação, permitindo a análise dos erros por classe. A Figura 22 apresenta a matriz de confusão referente à validação do melhor modelo sobre o conjunto de teste. Por fim, o modelo de melhor desempenho foi também salvo em formato binário “.pkl”, para ser utilizado posteriormente.

<sup>18</sup> Realiza validação cruzada estratificada repetida, preservando a proporção entre classes e aumentando a robustez da avaliação do modelo.

**Figura 24 – Matriz de confusão com os dados de teste para o modelo de SVM.**



**Fonte: Autoria própria.**

Além disso, é importante salientar que, no núcleo linear,  $C$  controla o compromisso entre margem ampla e erros de classificação. No núcleo RBF,  $\gamma$  regula a influência local dos vetores de suporte no espaço transformado. A busca conjunta de  $C$  e  $\gamma$  visa equilibrar a variância, prevenindo tanto subajuste quanto sobreajuste.

#### 4.5.3 Algoritmo Ensemble

Outro classificador foi desenvolvido com base no algoritmo *Random Forest* (que é um algoritmo específico de *ensemble*, mais precisamente do tipo *Bagging*), com o objetivo de aprimorar a capacidade discriminativa entre amostras normais e com lesão. O método baseia-se na combinação de múltiplas árvores de decisão independentes, cada uma treinada a partir de subconjuntos aleatórios de amostras e atributos, de forma a reduzir a variância do modelo e aumentar sua robustez frente a ruídos e variações nos dados.

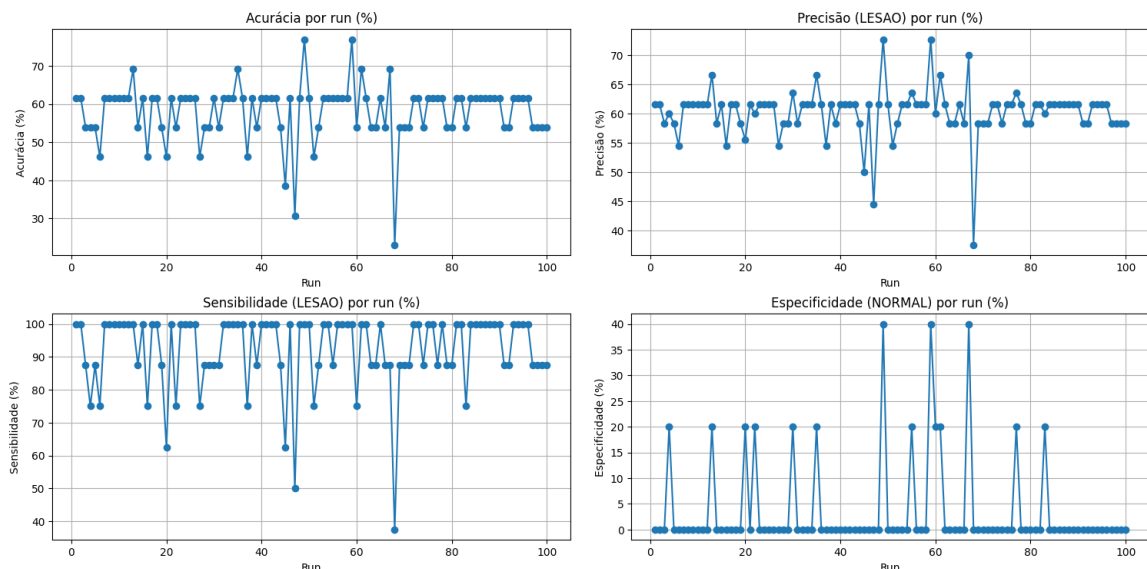
De forma análoga aos demais classificadores, esse algoritmo passou pelas mesmas etapas de pré-processamento e, em seguida, foi organizado em um *pipeline* supervisionado composto por duas fases: imputação dos valores ausentes e aplicação do classificador *RandomForestClassifier*. O modelo foi configurado com ponderação automática das classes (*balanced\_subsample*), que recalcula os pesos de cada classe em cada árvore, reduzindo o impacto do desbalanceamento durante o treinamento. Além disso, empregou-se amostragem com reposição (*bootstrap*) e a

estimativa de desempenho *out-of-bag*, na qual a acurácia é calculada avaliando cada amostra apenas nas árvores em que ela não foi utilizada para treinamento.

A etapa de otimização dos hiperparâmetros foi conduzida, assim como para os demais, por meio do método *GridSearchCV*, que selecionou as combinações dentro de um espaço de busca pré-definido. Foram ajustados parâmetros como o número de árvores (*n\_estimators*), profundidade máxima (*max\_depth*), número mínimo de amostras por divisão e folha (*min\_samples\_split* e *min\_samples\_leaf*), proporção de atributos utilizados em cada divisão (*max\_features*) e fração de amostras empregadas em cada estimador (*max\_samples*). A métrica de avaliação adotada foi a acurácia balanceada, e a validação do modelo utilizou o método *RepeatedStratifiedKfold*, assegurando a representatividade das classes em cada subdivisão.

O processo foi repetido também em 100 execuções independentes, com diferentes variações de amostras entre treinamento e teste, a fim de avaliar a estabilidade e a generalização do modelo. Para cada execução, foram calculadas métricas de desempenho, incluindo acurácia, sensibilidade, especificidade e precisão para a classe positiva (*Lesão*), e o resumo desses desempenhos pode ser observado na Figura 25.

**Figura 25 – Resumo das 100 execuções para o Ensemble.**

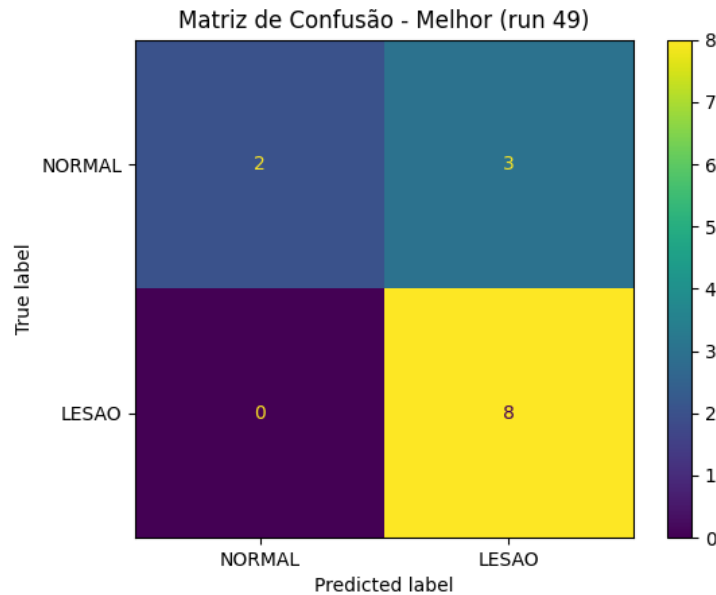


**Fonte: Autoria própria.**

O modelo com maior acurácia balanceada foi obtido na execução de número 49. Esse modelo de melhor desempenho foi selecionado como modelo final, sendo posteriormente avaliado por meio da matriz de confusão e do relatório de classificação. A Figura 26 apresenta a matriz de confusão referente ao desempenho do modelo sobre o conjunto de teste, sendo ele salvo em

formato binário “.pkl”.

**Figura 26 – Matriz de confusão com os dados de teste para o modelo de Ensemble.**

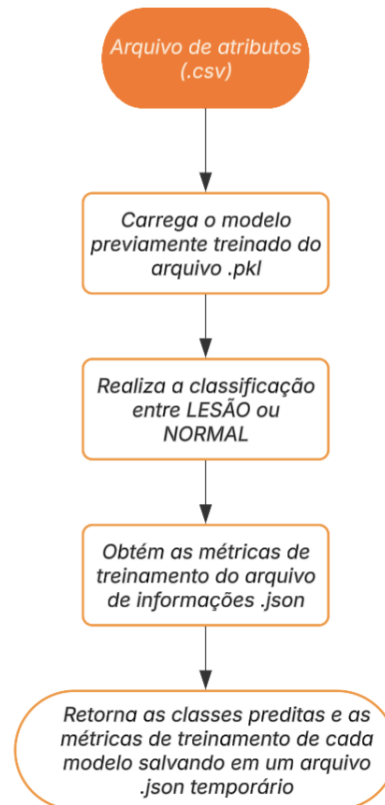


Fonte: Autoria própria.

#### 4.6 Classificação e Geração de Métricas

De forma geral, a etapa de classificação fundamenta-se na aplicação final dos modelos de aprendizado de máquina sobre os sinais previamente processados. Inicialmente, o conjunto de atributos obtido após as etapas de pré-processamento e reconstrução é consolidado em um arquivo “.csv”, o qual é utilizado como entrada para o modelo previamente treinado e armazenado em um arquivo “.pkl”. Em seguida, esse modelo realiza a classificação automática dos dados, distinguindo entre as categorias *Lesão* e *Normal*, conforme os padrões aprendidos durante o treinamento supervisionado. A Figura 27 descreve as etapas envolvidas no processo de classificação.

**Figura 27 – Fluxo de processamento de classificação.**



**Fonte: Autoria própria.**

Após a execução da classificação, são obtidas as métricas de desempenho correspondentes ao modelo, abrangendo acurácia, precisão, sensibilidade e especificidade, calculadas a partir das informações contidas em arquivos “.json”. Por fim, as classes previstas e as respectivas métricas são registradas em arquivos temporários “.json”, assegurando o registro, a rastreabilidade e a análise comparativa dos resultados obtidos em cada execução.

#### **4.7 Aplicativo de Apoio à Análise dos Sinais de Vibroartrografia**

Buscando integrar todas as etapas desenvolvidas anteriormente, foi desenvolvido um aplicativo *desktop*, para o sistema operacional *Windows*, destinado à análise de sinais vibroartrográficos. O sistema permite desde a visualização dos arquivos de dados adquiridos até a configuração das etapas de pré-processamento e dos parâmetros utilizados na reconstrução dos sinais. Além disso, o aplicativo disponibiliza uma interface que facilita a visualização das curvas clínicas resultantes e inclui uma seção voltada à visualização e sugestão dos classificadores utilizados na etapa de diagnóstico.

A interface gráfica interativa do aplicativo foi desenvolvida utilizando tecnologias web (HTML, CSS e JavaScript), permitindo a criação de telas responsivas e dinâmicas. O núcleo de processamento dos dados, responsável pelas etapas de pré-processamento dos sinais, reconstrução via série de Fourier multidimensional e classificação das lesões, foi implementado em Python, aproveitando a mesma estrutura dos módulos desenvolvidos previamente. Para integrar essas duas camadas, foi empregada uma interface intermediária baseada no *framework* Electron, que possibilita a criação de aplicações *desktop* multiplataforma utilizando tecnologias web, mantendo a comunicação com *scripts* Python executados em segundo plano (Electron, 2025).

Ainda, buscando proporcionar maior flexibilidade e acessibilidade, o aplicativo foi desenvolvido para ser utilizado em formato executável independente, dispensando qualquer instalação adicional. Para isso, utilizou-se a distribuição *open source* WinPython, disponível no GitHub, que fornece um ambiente Python gratuito, completo e portátil para a plataforma *Windows*. Essa distribuição permite empacotar o interpretador e todas as bibliotecas necessárias dentro do próprio diretório do aplicativo, garantindo que ele possa ser executado em qualquer computador compatível, mesmo sem o Python ou outras dependências previamente instaladas. Assim, o profissional precisa apenas executar o arquivo principal do sistema, enquanto todas as versões de bibliotecas e módulos permanecem controladas localmente (WinPython, 2025).

#### 4.7.1 Arquitetura

A estrutura do aplicativo foi organizada de forma modular, separando a interface, a lógica de controle e o núcleo de processamento. Para garantir maior organização interna, manutenibilidade e escalabilidade, o desenvolvimento foi realizado integralmente na interface do Visual Studio Code, com controle de versão pelo GitHub.

A camada de interface gráfica foi desenvolvida com HTML e CSS, responsáveis pela estruturação visual e estilização das telas, respectivamente. A camada de controle, por sua vez, foi desenvolvida com o uso de JavaScript, gerenciando os eventos e a lógica de interação entre os elementos da interface. Dessa maneira, cada tela presente no aplicativo é estruturada em uma pasta individual, conforme ilustrado na Figura 28, que é composta por um conjunto independente de arquivos: um HTML para a estrutura, um CSS para estilização e tornar a interface mais agradável, e um JavaScript para o controle da comunicação com os demais módulos da aplicação.

**Figura 28 – Estrutura de diretórios do aplicativo.**

```

VIBROARTHROGRAPHY/
|
+-- App_Desktop/
|   +-- analysis/
|       |   +-- app.py
|       |   \-- app_prediction.py
|   +-- Biblioteca_clinica/
|       |   +-- Deslocamento da patela/
|       |   +-- LCA/
|       |   +-- Menisco/
|       |   +-- Multiligamentar/
|       |   +-- Normal/
|       |   +-- Protese de Joelho/
|       |   \-- Reto Femoral/
|   +-- imgs/
|   +-- models/
|   +-- runs/
|   +-- Tela_inicial/
|   +-- Tela_examenes/
|   +-- Tela_Detalhes/
|   +-- Tela_configs/
|   +-- Tela_explorar/
|   \-- Tela_prediction/
\-- main.js

```

**Fonte: Autoria própria.**

O núcleo de processamento, implementado em Python, é composto por dois arquivos principais que se comunicam diretamente com o aplicativo: “app.py” e “app\_prediction.py”. O primeiro é responsável por receber o conjunto de dados e aplicar as etapas de pré-processamento, reconstrução e análise dos sinais vibroartrográficos. Ao final da execução, ele gera os resultados intermediários e os salva temporariamente na pasta “runs/”, além de criar um arquivo “manifest.json” contendo o mapeamento completo da análise. Também é gerado um arquivo “attributes.csv”, que armazena os valores para cada um dos harmônicos extraídos da reconstrução e serve como base para os classificadores.

O segundo arquivo, “app\_prediction.py”, é responsável por aplicar os modelos de classificação previamente treinados. Ele utiliza os atributos salvos em “attributes.csv” e executa os modelos sobre esses dados, retornando o resultado da classificação (indicando se o exame corresponde ao de um joelho normal ou com presença de lesão), bem como as métricas obtidas no treino de cada classificador.

#### 4.7.2 Integração

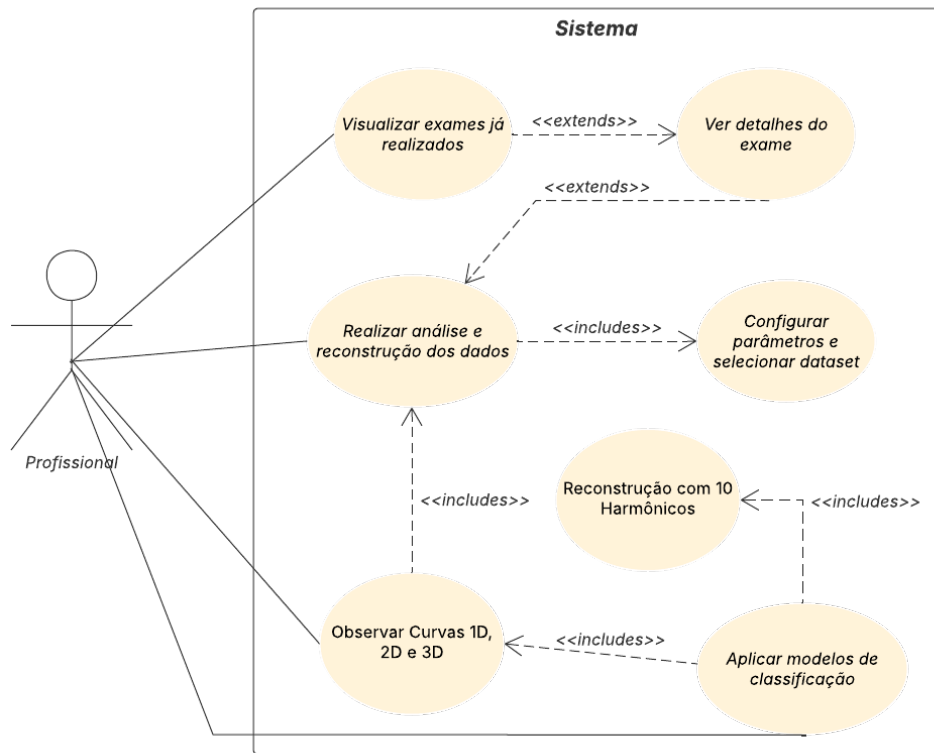
A integração entre as camadas internas, especificamente entre a interface desenvolvida em plataforma Electron e o núcleo de processamento em Python, foi estruturada para garantir a execução coordenada das etapas de análise definidas pelo usuário. Essa comunicação é realizada por meio do mecanismo de IPC (*Inter-Process Communication*) do Electron, que utiliza o módulo *child\_process* e o método *spawn* para executar os *scripts* Python em processos independentes. Os parâmetros configurados na interface são enviados ao núcleo de processamento por meio de arquivos temporários no formato “.json”, o que permite a execução do código Python com os valores definidos pelo usuário, sem interromper o funcionamento do aplicativo principal.

Após a execução, o Electron coleta as saídas geradas pelos *scripts*, como os arquivos “.json” e “.csv”, e as disponibiliza para acesso pela interface gráfica. Esses arquivos são armazenados temporariamente na pasta “runs/”, o que possibilita o carregamento posterior dos resultados de análise e a geração dos gráficos no ambiente da aplicação.

Dessa maneira caracteriza-se o fluxo modular: o *script* em Python é acionado apenas uma vez durante o processo de análise, e todos os gráficos são gerados a partir dos arquivos resultantes armazenados na pasta “runs/”. Uma segunda possibilidade de acionamento do núcleo Python ocorre na etapa de classificação das lesões, utilizando o arquivo de atributos gerado durante a análise e armazenado nessa mesma pasta. Nessa etapa, o resultado da classificação é também salvo em um arquivo temporário “.json”, que contém a indicação do classificador, classe predita e métricas de treinamento. Esse arquivo é acessado pela interface gráfica, que apresenta os diagnósticos obtidos por cada um dos classificadores.

Tratando da integração sob a perspectiva do usuário, a Figura 29 apresenta o diagrama de casos de uso do aplicativo. Esse diagrama resume as principais interações entre o usuário e o sistema, demonstrando as ações disponíveis na interface e as correspondentes operações executadas pelo núcleo de processamento.

**Figura 29 – Diagrama de Casos de Uso do Aplicativo.**



**Fonte: Autoria própria.**

De forma geral, o diagrama expresso na Figura 29 representa o fluxo de utilização do aplicativo, abrangendo desde a seleção e configuração dos dados até a geração dos resultados e a aplicação dos classificadores.

## 5 RESULTADOS E DISCUSSÃO

Após a aplicação dos processos de filtragem dos sinais e obtidos os coeficientes padronizados da série de Fourier Multidimensional para os sinais amostrados, o processo de reconstrução do sinal a partir dos parâmetros foi projetado graficamente em espaços bidimensionais e tridimensionais a fim de facilitar a visualização das assinaturas cinemáticas de lesão. Os resultados foram analisados tanto em função do tempo, como em função dos referenciais cartesianos de cada sensor.

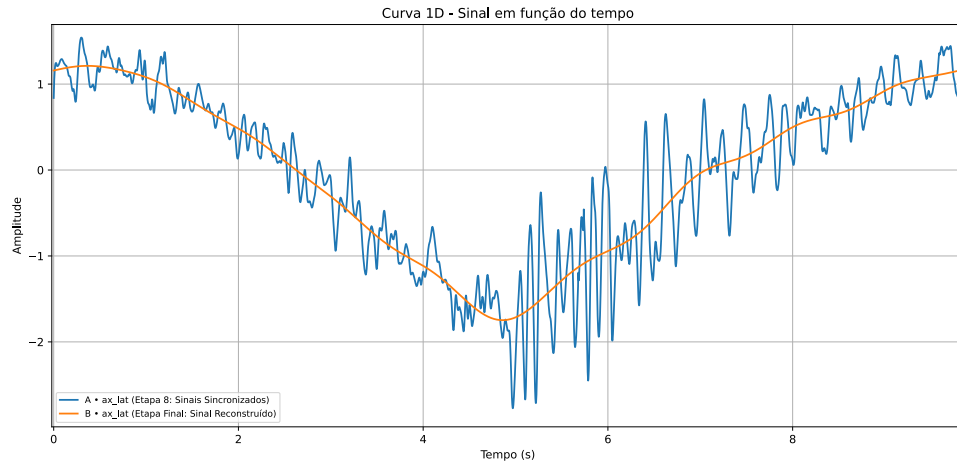
Nas representações, foram adotadas abreviações relativas ao sensor, eixo cartesiano e região onde está posicionado. Por exemplo, a componente  $x$  do acelerômetro localizada na região patelar foi representada como  $ax_{pat}$ . No caso do giroscópio, o mesmo padrão é adotado, como no caso da componente  $y$  do sensor localizada na região medial, que será representada como  $gy_{med}$ . O mesmo se aplica para ambos os sinais da região lateral, representados como  $lat$ .

Após a validação computacional do método de reconstrução via Série de Fourier, a solução foi empregada como uma funcionalidade disponível na interface de análises em tempo real desenvolvida. Essa interface é uma ferramenta que tem o objetivo de integrar todos os avanços desenvolvidos no projeto e possibilitar a análise e comparação de diferentes exames e parâmetros de forma amigável ao usuário.

### 5.1 Reconstrução das componentes cartesianas

Os sinais foram inicialmente analisados individualmente em função do tempo. Este experimento permitiu que fossem identificadas as componentes mais relevantes durante o movimento, ou seja, aquelas que possuem as maiores variações no período amostrado. Foram obtidos os gráficos para a reconstrução de todas as acelerações e velocidades angulares, levando em consideração diferentes quantidades de harmônicos na decomposição de Fourier. As diferenças na precisão das reconstruções ficam visualmente representadas na Figura 30 e na Figura 31. Em ambos os casos, o sinal original está representado na cor azul e a reconstrução na cor laranja.

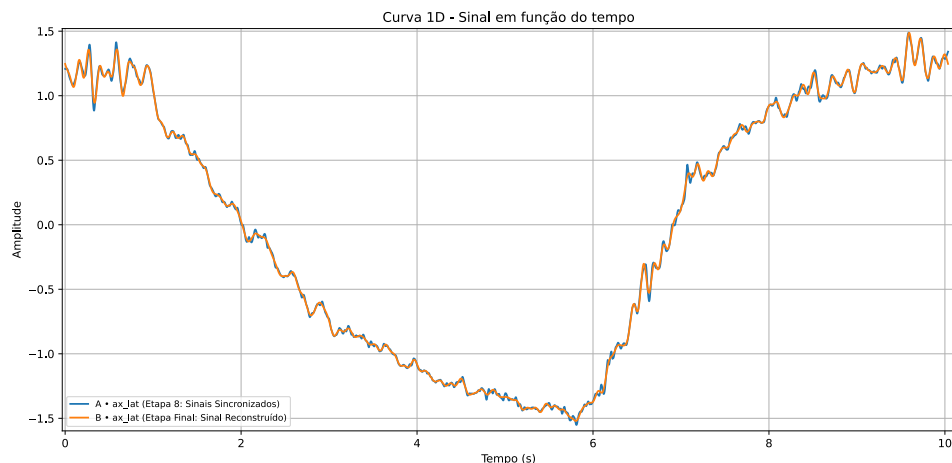
**Figura 30 – Exemplo de reconstrução individual realizado para a componente aceleração do sensor lateral no eixo x, considerando 10 harmônicos.**



**Fonte: Autoria própria.**

A reconstrução, a partir da adoção de 10 harmônicos, segue a tendência do gráfico e reconstrói parcialmente o sinal, mas não é sensível a picos e regiões de maior frequência. Estas características não são tão evidentes quando observamos o gráfico que utilizou 100 harmônicos para a decomposição do sinal, uma vez que este representou com maior precisão os dados originais. Contudo, utilizando 10 harmônicos, é possível obter uma representação que pode ser utilizada no treinamento dos classificadores com menor custo computacional, o que representa uma grande vantagem na operação do sistema.

**Figura 31 – Exemplo de reconstrução individual realizado para a componente aceleração do sensor lateral no eixo x, considerando 100 harmônicos.**



**Fonte: Autoria própria.**

Após a obtenção das reconstruções, foram realizados testes variando o número de harmônicos utilizados no processo, a fim de avaliar sua influência sobre a qualidade da reconstrução

dos sinais. Em seguida, foram calculados os valores de NRMSE em relação aos sinais originais normalizados. A Tabela 3 apresenta os erros obtidos nas reconstruções com diferentes números de harmônicos (100 e 1000), quantificando as análises realizadas a partir das representações gráficas.

**Tabela 3 – Comparação do NRMSE entre reconstrução com 100 e 1000 harmônicos - Sinal por posição do sensor e eixo cartesiano.**

Sinal	Erro (%) - 100 harmônicos	Erro (%) - 1000 harmônicos
ax_lat	8.4977	0.7965
ax_med	5.0408	1.3904
ax_pat	4.7590	0.4152
ay_lat	7.3480	0.6551
ay_med	6.6569	1.6496
ay_pat	15.5882	0.9253
az_lat	10.7317	2.2862
az_med	4.9994	0.4334
az_pat	5.2194	0.3048
gx_lat	53.1852	2.9327
gx_med	57.5183	2.6428
gx_pat	66.0564	3.7096
gy_lat	49.5962	6.1782
gy_med	39.3192	2.7162
gy_pat	15.4294	0.4909
gz_lat	29.1977	1.2289
gz_med	11.0165	0.4904
gz_pat	30.6928	2.9772

**Fonte: Autoria própria.**

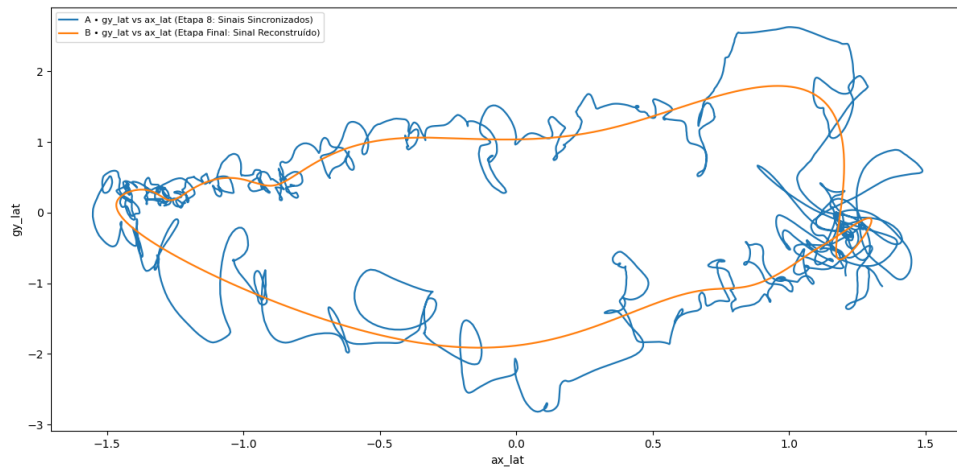
Quando analisados os 18 gráficos gerados nesta etapa, observa-se um padrão muito próximo da periodicidade nas representações dos acelerômetros, enquanto os sinais dos giroscópios apresentam diversas regiões de alta frequência, o que resulta em erros mais elevados. Isto ocorre devido às características físicas distintas dos sensores. Movimentos biomecânicos tendem a gerar variações rápidas na velocidade angular, o que cria componentes de maior frequência. Estas componentes podem não ser necessariamente ruído, e não foram filtradas a fim de manter a fidelidade ao sinal original. Este dispositivo é muito sensível às variações de temperatura, o que também contribui para a geração de ruídos.

## 5.2 Reconstrução do sinal em duas dimensões

Nesta etapa, foi realizada uma comparação entre as qualidades das reconstruções para diferentes quantidades de harmônicos e avaliação dos benefícios na inspeção visual do sinal reconstruído com menor complexidade. Comparando-se um sinal de acelerômetro em função de um sinal de giroscópio com sua reconstrução, foi possível visualizar graficamente a capacidade

de manutenção das características mais significativas do sinal.

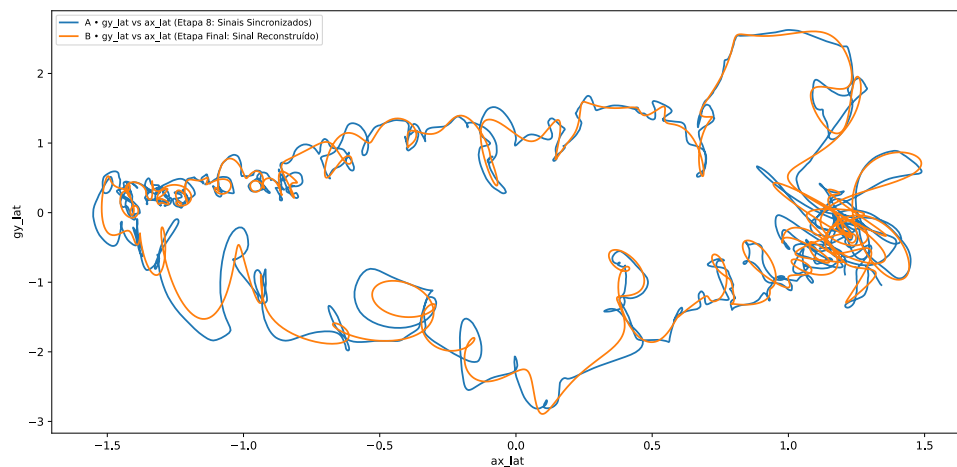
**Figura 32 – Sinal reconstruído a partir dos coeficientes da Série de Fourier - 10 harmônicos.**



**Fonte: Autoria própria.**

Na Figura 32, que contempla a contribuição dos sinais de acelerômetros e giroscópios na região lateral, o método demonstrou-se eficiente na manutenção das características da função original com a utilização de 10 harmônicos. Todavia, não foi possível mensurar os efeitos da suavização das componentes de maior frequência até que fossem aplicados os classificadores, uma vez que estas poderiam conter assinaturas características de lesões. A seguir, a Figura 33 demonstra o mesmo sinal reconstruído, a partir da decomposição com 100 harmônicos. Esta mostrou-se uma representação muito próxima do sinal original, comprovando novamente a eficácia da composição da Série de Fourier para um caso de maior dimensão.

**Figura 33 – Sinal reconstruído a partir dos coeficientes da Série de Fourier - 100 harmônicos.**

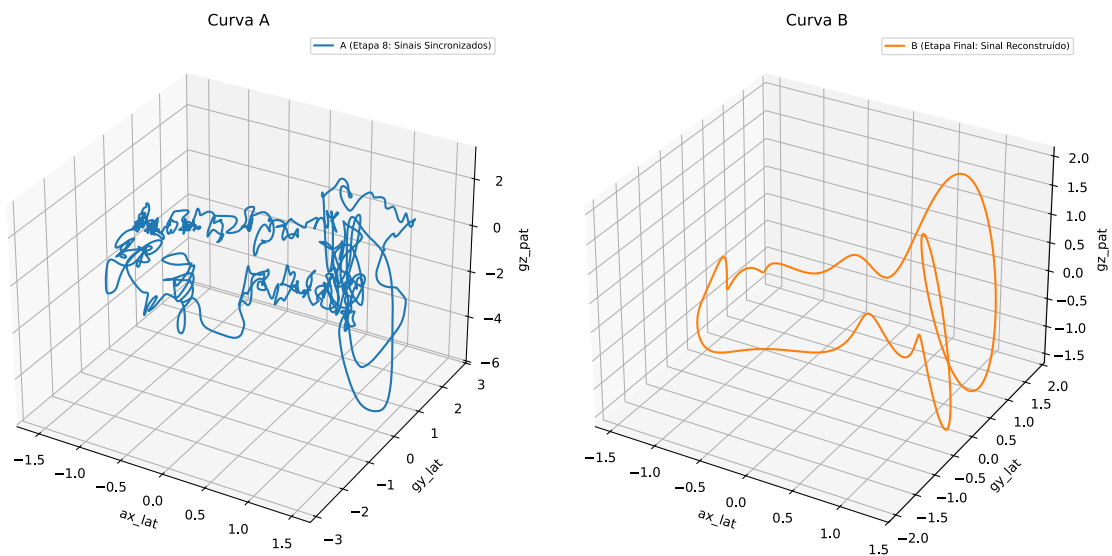


**Fonte: Autoria própria.**

### 5.3 Reconstrução dos sinais com 3 variáveis

Neste caso, o procedimento foi realizado a partir dos resultados de três componentes distintas, buscando melhorias na análise de padrões de assinatura das lesões. As curvas tridimensionais geradas consideram as três coordenadas cartesianas (X, Y, Z), sem levar em conta o eixo temporal. Com o objetivo de facilitar a visualização, foram elaborados gráficos separados para os sinais originais e reconstruídos. A Figura 34 ilustra um exemplo referente ao sinal de um giroscópio e um acelerômetro na região lateral.

**Figura 34 – Sinais cartesianos dos sensores e sua reconstrução com  $h = 10$ .**



**Fonte: Autoria própria.**

Nas representações tridimensionais, é possível rotacionar as figuras para melhor visualização. Também fica evidente a manutenção das características da série, como o fato da curva ser fechada. A reconstrução com menor número de harmônicos no caso com três dimensões simplifica a análise das assinaturas, podendo ser utilizada para comparar diretamente diferentes condições de lesão.

### 5.4 Classificadores

A etapa de classificação representou a aplicação prática dos modelos de aprendizado de máquina sobre os sinais previamente reconstruídos, possibilitando avaliar a capacidade dos algoritmos em reconhecer padrões vibroartrográficos associados à presença ou ausência de lesões no joelho.

De forma geral, observou-se que os três classificadores (Árvore de Decisão, SVM e método *Ensemble*) conseguiram distinguir adequadamente as duas classes analisadas, mesmo diante da limitada qualidade e quantidade de dados disponíveis. As subseções a seguir discutem os fatores que influenciaram esses resultados, as restrições impostas pelo conjunto de dados e as métricas utilizadas para expressar a confiabilidade de cada classificador.

#### 5.4.1 Limitações dos Dados

Embora todas as etapas tenham sido conduzidas com o intuito de garantir resultados consistentes, algumas limitações inerentes ao conjunto de dados podem ter impactado a robustez dos modelos de classificação. Durante a aquisição dos sinais, foram identificadas inconsistências ocasionais em sensores específicos, comprometendo a qualidade de parte das amostras. Além disso, o número reduzido de observações limita a capacidade de generalização dos modelos e aumenta a suscetibilidade ao *overfitting*.

Observou-se também uma falta de padronização nos rótulos das classes originais, o que exigiu intervenções adicionais na fase de pré-processamento. Tais fatores devem ser considerados na interpretação dos resultados, pois afetam diretamente a distribuição das amostras e, conseqüentemente, a estabilidade das métricas de desempenho.

#### 5.4.2 Análise dos Resultados de Classificação

Os experimentos de classificação foram conduzidos com o objetivo de avaliar a capacidade dos modelos em distinguir entre sinais vibroartrográficos provenientes de joelhos normais e lesionados. Após a padronização do número de simulações, do método de validação cruzada e do processo de busca de hiperparâmetros, foi possível realizar uma análise equitativa entre os classificadores, assegurando que as diferenças observadas refletissem as particularidades de cada paradigma de aprendizado. As métricas de desempenho obtidas (acurácia, precisão, sensibilidade e especificidade) estão apresentadas na Tabela 4, que sintetiza os resultados obtidos para cada modelo com o *dataset* de validação.

**Tabela 4 – Métricas de Desempenho dos classificadores.**

<b>Classificador</b>	<b>Acurácia</b>	<b>Precisão</b>	<b>Sensibilidade</b>	<b>Especificidade</b>
Árvore de Decisão	0,846	0,800	<b>1,000</b>	0,600
<i>Ensemble</i>	0,769	0,727	<b>1,000</b>	0,400
SVM	<b>0,923</b>	<b>1,000</b>	0,875	<b>1,000</b>

**Fonte: Autoria própria.**

De forma geral, os modelos apresentaram resultados consistentes, com acurácia mínima de aproximadamente 77%, indicando que, mesmo com um conjunto de dados limitado e sujeito a variações de qualidade, os algoritmos analisados foram capazes de identificar padrões relevantes nos sinais reconstruídos.

A Árvore de Decisão apresentou desempenho equilibrado, com acurácia de 84,6%, precisão de 80%, sensibilidade de 100% e especificidade de 60%. Esse comportamento evidencia boa capacidade de generalização, mesmo com número reduzido de amostras, e reflete a natureza simbólica do modelo, cuja estrutura interpretável favorece a explicabilidade dos resultados e a possível aplicação em sistemas de *Edge Computing*<sup>1</sup>.

O método *Ensemble*, assim como a Árvore de Decisão, alcançou sensibilidade máxima (100%), identificando corretamente todas as amostras lesionadas e evitando falsos negativos, característica de extrema importância em contextos clínicos, pois um falso negativo (lesão predita como normal) poderia levar à ausência de tratamento adequado. Contudo, apresentou menor capacidade de generalização, com acurácia de 76,9% e especificidade de 40%, o que indica tendência a classificar amostras normais como patológicas. O modelo *Ensemble*, inclusive, apresentou discrepância entre o bom desempenho no treinamento (no qual alcançou 100,0% para todas as métricas), em relação às métricas apresentadas com o conjunto de teste, o que sugere *overfitting*, comportamento comum em modelos compostos por múltiplos estimadores quando há limitação no volume de dados.

O SVM apresentou o melhor desempenho global entre os classificadores avaliados, com acurácia de 92,3%, precisão e especificidade de 100%, e sensibilidade de 87,5%. Esses resultados demonstram a capacidade do modelo em generalizar adequadamente o aprendizado, separando as classes com elevada precisão. Tal desempenho está associado à natureza linear do algoritmo, que busca maximizar as margens de separação entre as classes, tornando-o menos suscetível ao *overfitting* em conjuntos de dados pequenos e com fronteiras bem definidas.

As matrizes de confusão apresentadas na seção 4.7, nas Figuras 22, 24 e 26, complementam a interpretação das métricas, evidenciando o comportamento individual de cada modelo. A Árvore de Decisão exibiu diagonal dominante, com poucas confusões entre as classes. O *Ensemble* manteve a predominância da diagonal principal, mas apresentou maior dispersão na classe normal, o que explica sua especificidade reduzida. O SVM apresentou matriz predominantemente diagonal, refletindo clara separação entre as classes e ausência de falsos positivos.

<sup>1</sup> Refere-se ao processamento e análise de dados próximos à sua origem, em dispositivos ou sensores, reduzindo latência, uso de banda e dependência da nuvem.

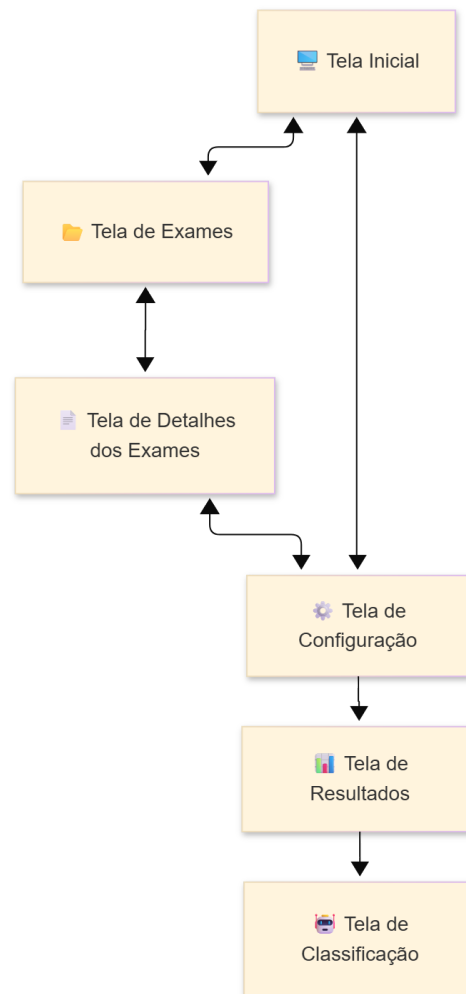
Por fim, destaca-se que cada classificador possui características próprias que podem ser exploradas conforme o contexto de aplicação. A Árvore de Decisão, por ser um modelo simbólico, oferece transparência e facilidade de interpretação, o que pode ser útil quando a explicabilidade é um requisito. O SVM apresentou o melhor equilíbrio entre desempenho e capacidade de generalização, mostrando-se particularmente adequado para bases de dados menores e bem estruturadas. Já o modelo baseado em agregação (*Ensemble*), embora teoricamente mais robusto por combinar múltiplos aprendizados, neste estudo apresentou sinais de sobreajuste aos dados de treinamento, resultando em desempenho inferior na etapa de validação.

Esses resultados confirmam as tendências observadas nas métricas numéricas e reforçam o bom desempenho geral dos modelos, com destaque para o SVM pela sua elevada precisão e estabilidade.

## **5.5 Aplicativo desenvolvido e suas funcionalidades**

A fim de ilustrar a organização interna do sistema e o percurso de navegação do profissional no aplicativo, a Figura 35 apresenta todas as telas que o compõem, bem como o diagrama de fluxo entre elas.

**Figura 35 – Estrutura de telas do aplicativo.**



**Fonte: Autoria própria.**

### 5.5.1 Tela Inicial

Ao iniciar a aplicação, o profissional é direcionado à tela inicial, ilustrada na Figura 36, na qual há duas opções principais: visualizar uma lista com os exames já realizados, por meio do botão de “Exames Realizados”, ou prosseguir diretamente para a configuração das etapas e parâmetros de pré-processamento e processamento dos sinais, pelo botão “Explorar Resultados”.

**Figura 36 – Tela inicial do aplicativo.**

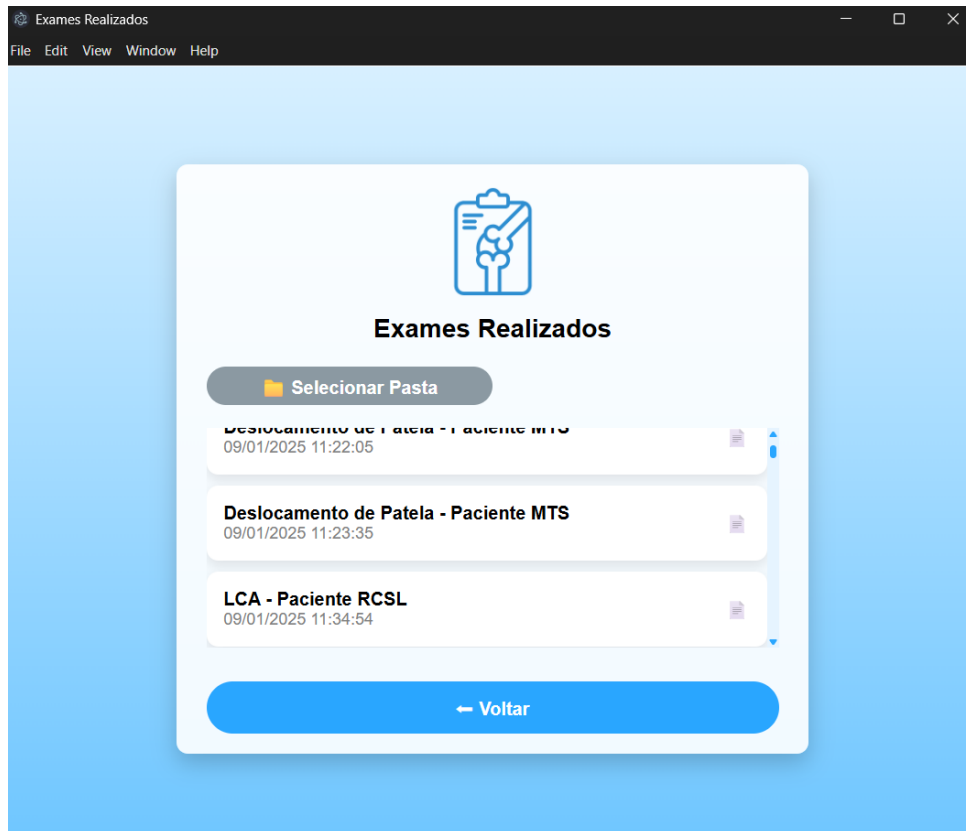
**Fonte: Autoria própria.**

### 5.5.2 Tela de Exames

A “Tela de Exames” permite ao profissional visualizar uma lista com os exames previamente realizados pelo software de aquisição que possuam seus respectivos arquivos de informações (“.json”) e dados (“.csv”) armazenados em uma mesma pasta. Para listar os exames disponíveis, o profissional deve apenas selecionar o diretório de interesse por meio do botão “Selecionar Pasta”.

Conforme ilustrado na Figura 37, a listagem apresenta, para cada exame, o diagnóstico obtido e o identificador do paciente no título do item, seguidos pela data e o horário de realização logo abaixo. A partir dessa tela, o profissional pode selecionar um exame para acessar seus detalhes, seguindo assim para a “Tela de Detalhes do Exame”, ou retornar à “Tela Inicial” por meio do botão “Voltar”.

**Figura 37 – Tela de Exames do aplicativo.**



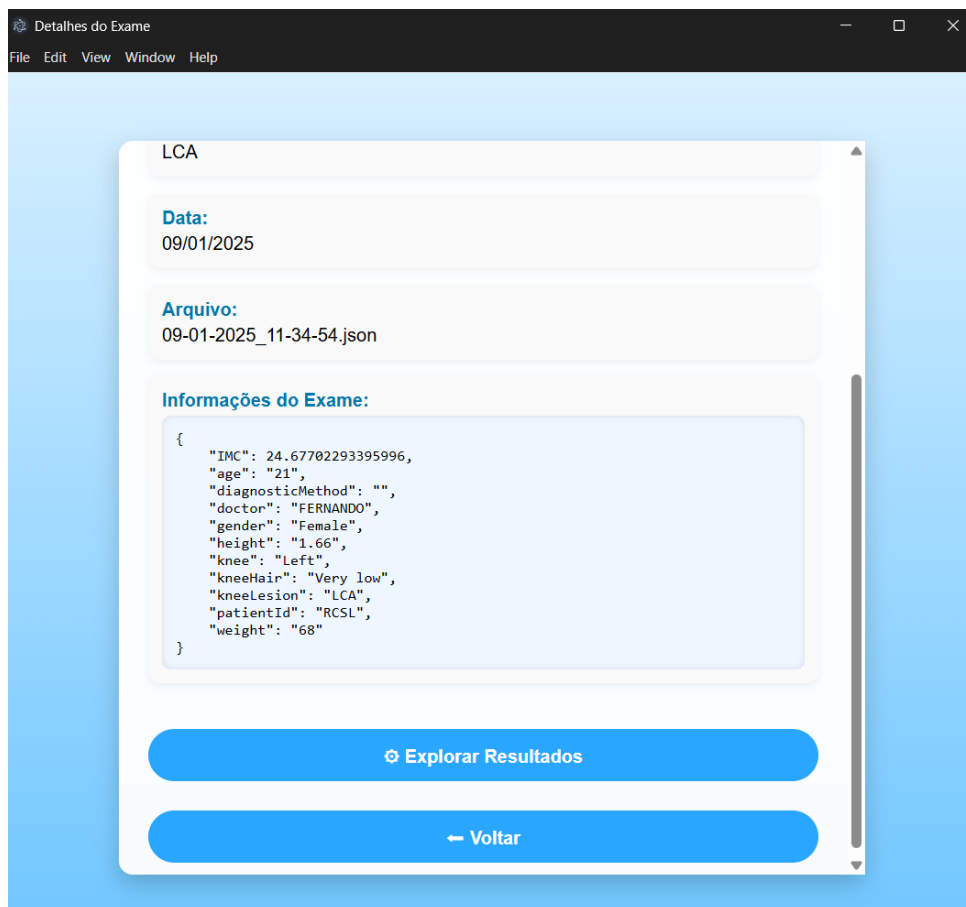
**Fonte: Autoria própria.**

### 5.5.3 Tela de Detalhes dos Exames

A “Tela de Detalhes dos Exames” exibe as informações coletadas do paciente durante a consulta de aquisição dos dados, armazenadas em seu respectivo arquivo de informações “.json”. Como apresentado na Figura 38, essa tela apresenta campos de destaque para o identificador do paciente, o diagnóstico da lesão obtido por método auxiliar, a data do exame e o nome do arquivo acessado. Além disso, o conteúdo completo do arquivo é exibido em formato legível, permitindo a verificação detalhada das informações associadas ao exame e ao paciente.

A partir dessa tela, o profissional pode retornar à “Tela de Exames”, por meio do botão “Voltar”, para consultar outros exames disponíveis, ou avançar para a “Tela de Configurações”, onde são definidos os parâmetros e etapas de processamento que serão aplicados aos dados do exame selecionado.

**Figura 38 – Tela de Detalhes do Exame do aplicativo.**



**Fonte: Autoria própria.**

#### 5.5.4 Tela de Configurações

A “Tela de Configurações” representa a interface na qual o profissional define as etapas e os parâmetros que serão aplicados pelo código em Python durante o processamento dos sinais vibroartrográficos. Essa tela organiza as opções em três blocos principais: dataset, pré-processamento e reconstrução. Todos os campos de configuração são inicializados com valores padrão, que refletem os parâmetros utilizados na rotina atual de coleta de dados. No entanto, esses valores podem ser ajustados conforme a necessidade do profissional, como ilustrado na Figura 39.

No bloco de dataset, o profissional pode selecionar o arquivo de dados “.csv” a ser analisado, por meio do botão “Escolher”. Caso o acesso tenha sido realizado a partir da “Tela de Detalhes dos Exames”, o arquivo correspondente é automaticamente carregado. Em seguida, o bloco de pré-processamento permite definir quais etapas serão aplicadas e ajustar seus respectivos parâmetros, incluindo padronização, filtragem de ruídos, remoção de períodos de imobilidade,

uniformização temporal das amostras, filtragem de *outliers* e cálculo da média entre múltiplos ciclos de aquisição. Por fim, o bloco de reconstrução possibilita configurar o número de harmônicos utilizados e o método de reconstrução do sinal, podendo o profissional escolher entre a Série de Fourier Simples, a Série Bidimensional (ou Paramétrica) proposta por Funaki e Lorentz (2022), ou a Série de Fourier Multidimensional, desenvolvida neste estudo.

**Figura 39 – Tela de Configurações do aplicativo.**

A interface do aplicativo 'Configurar Análise' apresenta as seguintes configurações visíveis:

- Dataset:** Campo para 'Arquivo CSV' com o texto 'Selecione um arquivo...' e um botão 'Escolher'. Um link 'Selecione um arquivo .csv.' está disponível.
- Pré-processamento:**
  - Botões de alternância: Padronizar, Reduzir ruído, Uniformizar tempo, Remover trechos parados, Remover outliers, Média dos ciclos (todos ativados).
  - Entradas numéricas:
    - Frequência de amostragem (Hz): 500
    - Número de ciclos: 3
    - Tempo parado (s): 4
    - Amostras por aquisição: 100
    - Ordem do filtro Savitzky-Golay: 4
    - Janela do filtro Savitzky-Golay: 31
- Reconstrução (Fourier):**
  - Número de harmônicos: 10
  - Método de reconstrução: Série de Fourier Multidimensional (selecionado no menu suspenso)

Na base da interface, há botões para 'Voltar', 'Restaurar padrão', 'Explorar resultados' e 'Analisar agora'.

**Fonte: Autoria própria.**

Com o dataset definido e as configurações de pré-processamento e reconstrução ajustadas, o profissional pode iniciar a análise dos sinais por meio do botão “Analisar agora”. Nesse processo, as configurações definidas na interface são convertidas em um arquivo temporário no formato “.json” e enviadas ao interpretador Python através de chamadas assíncronas do módulo *spawn* do Electron. Dessa forma, o script principal em Python é executado com os parâmetros configurados no *frontend*.

Ao término da execução, caso o processamento seja concluído com sucesso, o botão “Explorar Resultados” é automaticamente habilitado, permitindo ao profissional abrir uma nova tela dedicada à análise dos resultados e dos contornos obtidos na reconstrução dos sinais.

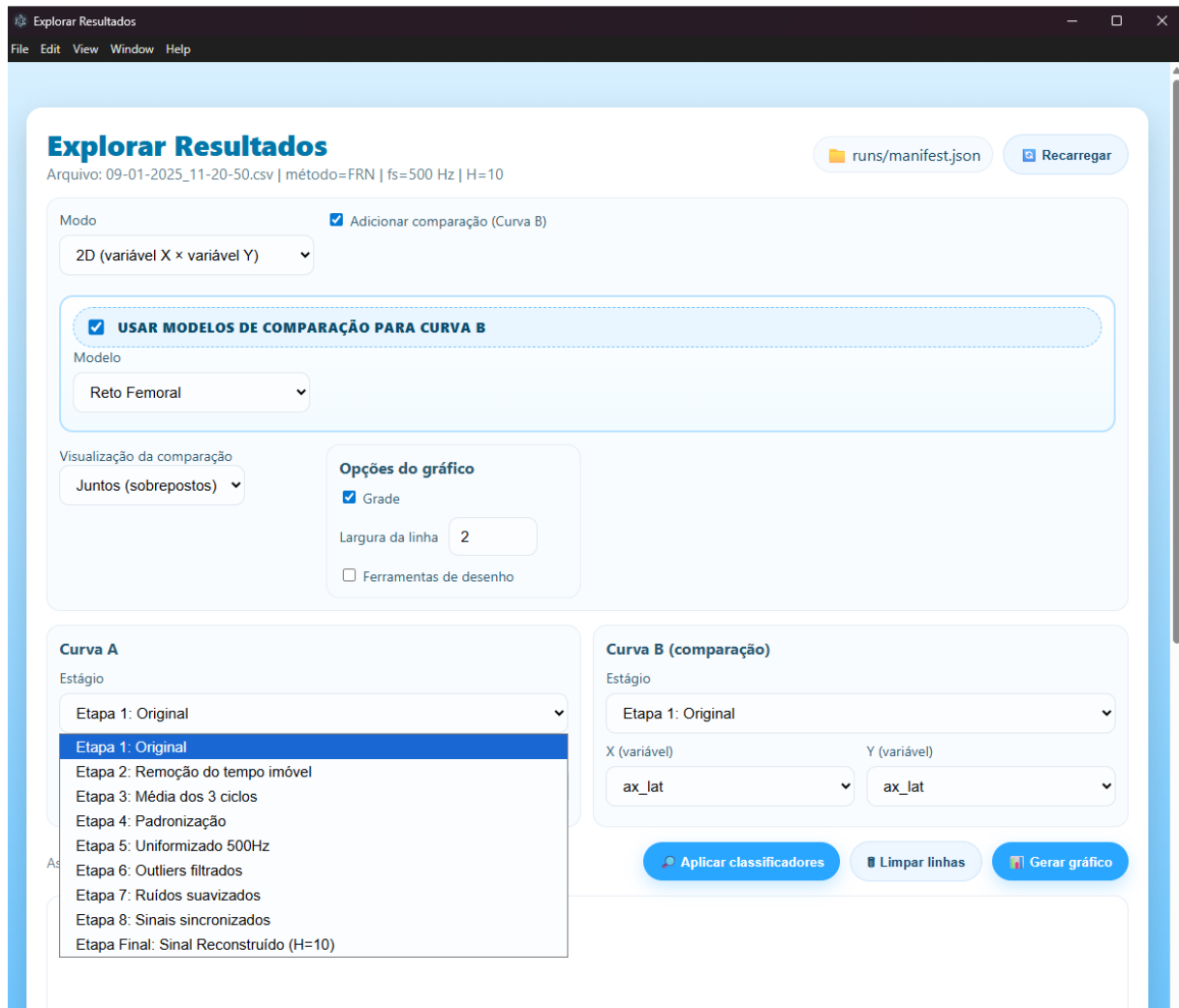
Além disso, essa tela inclui ferramentas adicionais para aumentar a praticidade no uso do sistema, como o gerenciamento de *presets*. Por meio dos botões “Salvar Preset” e “Carregar Preset”, o profissional pode armazenar suas configurações atuais e reutilizá-las posteriormente, facilitando a repetição de experimentos com os mesmos parâmetros. Também é possível restaurar os valores padrão utilizados na inicialização da página, por meio do botão “Restaurar Padrão”. Por fim, o profissional, assim como nas demais telas, tem a opção de retornar à tela de onde veio com o uso do botão “Voltar”.

#### 5.5.5 Tela de Resultados

A “Tela de Resultados” é aberta em uma nova janela após a conclusão da análise realizada pelo profissional na “Tela de Configurações”. Nela, o profissional, especialmente o profissional responsável pela avaliação clínica, pode explorar e investigar os resultados obtidos a partir da reconstrução dos sinais pelo método selecionado. Como ilustrado na Figura 40, essa interface disponibiliza diversas configurações e parâmetros ajustáveis, buscando oferecer uma ferramenta flexível e adaptável às diferentes necessidades de cada análise.

Na parte superior da tela, são exibidas informações gerais sobre o processamento realizado, como o nome do arquivo analisado, o método de reconstrução adotado, a frequência de amostragem dos dados e o número de harmônicos empregados na reconstrução. À direita do cabeçalho, o botão “Recarregar” permite que o profissional realize novas análises em paralelo, uma vez que os arquivos utilizados como base para gerar os gráficos são atualizados apenas quando esse botão é pressionado.

**Figura 40 – Tela de Resultados do aplicativo.**

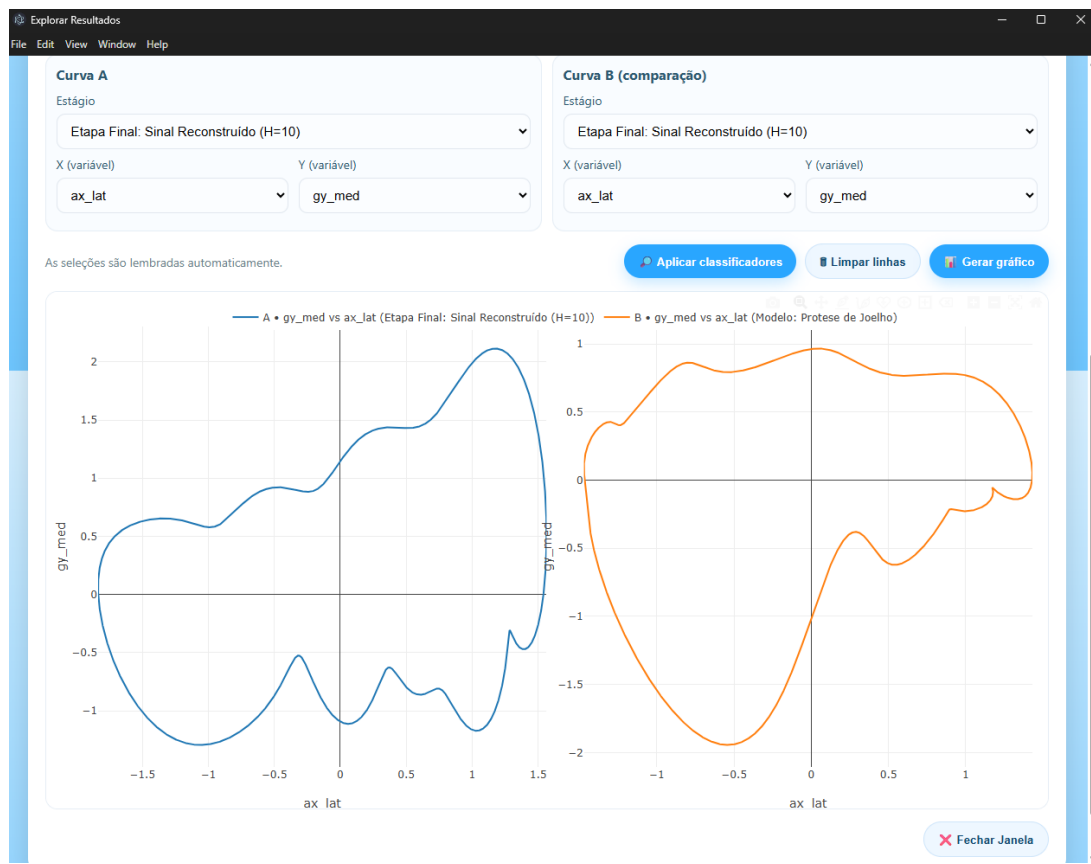


**Fonte: Autoria própria.**

Tratando do propósito principal, a visualização e comparação dos contornos reconstruídos, há a possibilidade de gerar gráficos 1D, 2D e 3D, representando uma variável em função do tempo ou mais variáveis entre si. Ainda, há a opção para o profissional de incluir uma curva adicional para comparação, exibindo-as lado a lado ou sobrepostas no mesmo gráfico. Além disso, o sistema permite combinar livremente qualquer eixo de aceleração linear ou velocidade angular, conforme a necessidade da análise.

Como uma ferramenta auxiliar e de base comparativa, é disponibilizada a alternativa de comparar os sinais obtidos com modelos de referência pré-existent, disponíveis para diferentes tipos de lesão (como deslocamento de patela, lesão de menisco, LCA, multiligamentar, prótese de joelho e reto femoral), bem como com dados representativos de joelhos normais.

**Figura 41 – Gráficos na Tela de Resultados do aplicativo.**



**Fonte: Autoria própria.**

Durante a exploração dos resultados, o aplicativo permite a seleção da etapa específica de pré-processamento cujos sinais se deseja visualizar, desde o sinal original até a reconstrução final, bem como escolher as variáveis envolvidas (combinando sensores, posicionamentos e eixos cartesianos) para compor o gráfico. Para a inspeção visual do gráfico, a interface conta com ferramentas de desenho interativas, que permitem ao profissional destacar regiões de interesse diretamente sobre os gráficos, traçar linhas ou contornos, inserir formas geométricas ou delimitar intervalos relevantes. Esses recursos foram implementados com o objetivo de disponibilizar diferentes ferramentas que podem vir a auxiliar na análise qualitativa dos sinais reconstruídos, favorecendo a interpretação clínica e o registro de observações diretamente no ambiente de análise.

Com os parâmetros definidos e as variáveis selecionadas, o profissional pode gerar os contornos correspondentes por meio do botão “Gerar gráfico”. Além disso, conforme apresentado na Figura 41, há a opção de aplicar os classificadores desenvolvidos no projeto diretamente sobre o sinal reconstruído. Essa funcionalidade é apenas habilitada quando a reconstrução do sinal é realizada utilizando 10 harmônicos, uma vez que todos os modelos de classificadores foram

treinados com essa configuração.

Ao profissional acionar o botão “Aplicar classificadores”, os modelos treinados são executados utilizando um *dataset* previamente salvo durante a análise dos sinais, realizando a classificação entre joelhos normais e com lesão. Assim que o processo é concluído, uma nova janela é aberta exibindo o resultado da predição, permitindo ao profissional visualizar o diagnóstico sugerido pelos classificadores.

#### 5.5.6 Tela de Classificação

A última tela presente no aplicativo, a “Tela de Classificação”, é aberta em uma nova janela logo após a conclusão da predição realizada na “Tela de Resultados”. Nessa interface, são exibidos os diagnósticos obtidos pelos três classificadores implementados no estudo (Árvore de Decisão, SVM e *Ensemble*), conforme ilustrado na Figura 42. Esse diagnóstico predito busca oferecer uma interpretação complementar dos resultados, combinando as diferentes abordagens de classificação desenvolvidas e auxiliando o profissional na tomada de decisão quanto ao diagnóstico final.

**Figura 42 – Tela de Classificação do aplicativo.**

MODELO	DIAGNÓSTICO	ACURÁCIA	PRECISÃO	SENSIBILIDADE	ESPECIFICIDADE
DecisionTree	Normal	84.6%	80.0%	100.0%	60.0%
EnsembleRF	Normal	76.9%	72.7%	100.0%	40.0%
SVM	Normal	92.3%	100.0%	87.5%	100.0%

**Fonte: Autoria própria.**

Cada modelo apresenta o resultado da predição para o exame analisado, indicando se os dados foram classificados como de um paciente normal ou com a presença de lesão. Além do diagnóstico, a tela exibe também as principais métricas de desempenho obtidas durante o

processo de treinamento dos classificadores: acurácia, precisão, sensibilidade e especificidade. Essa abordagem foi realizada com o objetivo de trazer transparência e permitir ao profissional interpretar a confiabilidade e a robustez de cada modelo.

## 6 CONCLUSÃO

A fim de avaliar a viabilidade do diagnóstico de lesões articulares no joelho a partir da reconstrução de sinais via Série de Fourier, foram analisados dados previamente coletados de pacientes com lesões classificadas e em fase pré operatória. Os sinais foram então decompostos em componentes harmônicas fundamentais utilizando uma aproximação periódica para aplicação da Série, o que resultou na extração de um vetor de características de ordem  $n$ . Foi observada a relação direta entre o aumento de  $n$  e a qualidade da reconstrução do sinal. Os algoritmos de aprendizado de máquina: SVM, Ensemble e Decision Tree, foram então treinados a partir da reconstrução dos sinais realizada com  $n = 10$ . Por fim, foi desenvolvida uma interface voltada ao uso clínico do sistema, facilitando a interação e tornando o processo de análise acessível.

A utilização da Série de Fourier  $n$ -dimensional demonstrou-se eficiente na representação dos sinais em distintos níveis de granularidade, evidenciando que a preservação de informações relevantes para tarefas de classificação pode ser alcançada mesmo com uma limitação no número de harmônicos. Quando se utilizou, por exemplo, 1000 harmônicos, observou-se um erro de reconstrução inferior a 1% em diversas componentes do sinal. Dessa forma, constatou-se que o aumento linear no número de harmônicos favorece uma reconstrução mais precisa dos sinais adquiridos, embora esse ganho de acurácia esteja atrelado a um aumento também linear no custo computacional, proporcional ao tamanho do vetor de atributos.

Dentre os classificadores avaliados, o modelo baseado em Máquinas de Vetores de Suporte (SVM) apresentou o desempenho mais expressivo, alcançando acurácia de 92,3%, com precisão e especificidade de 100% e sensibilidade de 87,5%. Esses resultados atestam a robustez da abordagem adotada e sua potencialidade e facilidade de uso para aplicação em sistemas de diagnóstico assistido por computador.

Apesar dos resultados promissores, limitações foram identificadas ao longo do estudo. Entre os principais desafios, destacam-se o número reduzido de amostras disponíveis, a inconsistência nas *labels* das medições e a presença de ruídos em determinados sensores, notadamente nos giroscópios, que se mostraram mais suscetíveis a variações ambientais, além da influência das componentes gravitacionais observada nos acelerômetros. Essas dificuldades inviabilizaram a identificação do tipo da lesão, inicialmente pretendido como um dos objetivos da pesquisa. Contudo, as estratégias adotadas para filtragem, sincronização temporal e padronização dos rótulos mostraram-se eficazes na atenuação desses fatores adversos, contribuindo para a estabilidade

e confiabilidade das análises subsequentes.

Portanto, o método demonstrou equilíbrio entre precisão e eficiência computacional, tornando-se uma alternativa promissora para aplicações clínicas futuras. Recomenda-se, como continuidade deste trabalho, a ampliação do conjunto amostral e a inclusão de diferentes condições biomecânicas no banco de dados, a fim de aprimorar a capacidade preditiva dos classificadores. Neste sentido, melhorias no processo de aquisição, como o aperfeiçoamento do procedimento de coleta, redução de influências ambientais no hardware e organização do banco de dados coletados, contribuiriam no fortalecimento da aplicação da vibroartrografia como ferramenta complementar no diagnóstico de lesões articulares.

## REFERÊNCIAS

AMARAPPA, S.; SATHYANARAYANA, S. V. Data classification using support vector machine (svm) – a simplified approach. **International Journal of Electronics and Computer Science Engineering (IJECS)**, v. 3, p. 435–445, 2014.

BACHES, Pedro; OLIVEIRA, Diego Escudeiro de; MUSSATTO, Guilherme do Amaral; HORITA, Melanie Mayumi; SALAS, Victor Eduardo Roman; JORGE, Rafael Baches. Diagnóstico das lesões meniscais em rampa. **Revista brasileira de ortopedia**, v. 59,5, p. e702–e706, 12 2024. Disponível em: <https://www.thieme-connect.de/products/ejournals/abstract/10.1055/s-0044-1791791>.

BASAVARAJU, K. S.; KUMAR, T. K.; REDDY, K. A. Analysis of vibroarthrographic signals for knee abnormality detection using fbse-ewt hierarchical frequency zone clustering approach. **IETE Journal of Research**, Feb. 25 2025.

BASAVARAJU, Krishna Sundeep; KUMAR, T. Kishore; REDDY, K. Ashoka. Vibroarthrographic signal classification for knee joint disorder detection using tunable q-factor wavelet transform based on entropy measures. **Engineering, Technology & Applied Science Research**, v. 15, n. 1, p. 19953–19958, 2025. Disponível em: <https://doi.org/10.48084/etasr.9245>.

BLOCKEEL, Hendrik; DEVOS, Laurens; FRÉNEY, Benoît; NANFACK, Géraldine; NIJSSEN, Siegfried. Decision trees: from efficient prediction to responsible ai. **Frontiers in Artificial Intelligence**, v. 6, July 2023. Disponível em: <https://www.frontiersin.org/ARTICLES/10.3389/frai.2023.1124553>.

BOASHASH, Benny. **Time-Frequency Signal Analysis and Processing: A Comprehensive Reference**. Elsevier, 1992. ISBN 978-0-7506-2129-6. Disponível em: <https://www.elsevier.com/books/time-frequency-signal-analysis-and-processing/boashash/978-0-7506-2129-6>.

BORZUCKA, Dorota; KRĘCISZ, Krzysztof; BĄCZKOWICZ, Dawid. Influence of external load during back squats on knee joint arthrokinematics analyzed by vibroarthrography. **Research Square**, 2024. Disponível em: <https://doi.org/10.21203/rs.3.rs-4516377/v1>.

CAPLE, J.; BYRD, J.; STEPHAN, C. Elliptical fourier analysis: fundamentals, applications, and value for forensic anthropology. **International Journal of Legal Medicine**, Springer, v. 131, n. 6, p. 1675–1690, Feb 2017.

CHARBUTY, B.; ABDULAZEEZ, A. Classification based on decision tree algorithm for machine learning. **Journal of Applied Science and Technology Trends**, v. 2, n. 01, p. 20–28, 2021.

CRENNA, Francesco; ROSSI, Giovanni Battista; BERARDENGO, Marta. Filtering biomechanical signals in movement analysis. **Sensors**, MDPI, v. 21, n. 13, p. 4580, 2021. Disponível em: <https://www.mdpi.com/1424-8220/21/13/4580>.

Electron. **Electron: Build cross-platform desktop apps with JavaScript, HTML, and CSS**. 2025. <https://www.electronjs.org>. Accessed: 2025-10-21.

FILHO, Marcantonio Machado da Cunha Cavalcanti; DOCA, Daniel; COHEN, Moisés; FERRETTI, Mário. Atualização no diagnóstico e tratamento das lesões condrais do joelho. **Revista Brasileira de Ortopedia**, v. 47, n. 1, p. 12–20, 2012. Disponível em: <https://doi.org/10.1590/S0102-36162012000100001>.

FILHO, Umberto Paulo de Castro Alves; LEAL, Liomar Couto; FERNANDES, Aloísio Alan Costa; MARQUES, Petruska de Oliveira; SILVA, Robson Rocah da. Concordância entre artroscopia e ressonância magnética para avaliação das lesões do joelho. **Revista Baiana de Saúde Pública**, 2011. Disponível em: <https://doi.org/10.22278/2318-2660.2010.v34.n1.a88>.

FUNAKI, Lucas Eidi; LORENTZ, Matheus de Oliveira. Vibroartrografia de joelho. **Deteção De Lesões De Joelho Por Técnicas De Vibroartrografia E Machine Learning**, 2022. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/32833>.

GOIC, Gaëtan Le; BROWN, Christopher A; FAVRELIERE, Hugues; SAMPER, Serge. Outlier filtering: A new method for improving the quality of surface measurements. **ResearchGate**, v. 24, n. 1, p. 2–4, 2012.

GOMES, José Wivo; SANTOS, João Evangelista Dantas dos; SANTOS, Rodolfo de Sousa. Avaliação de modelos de previsão de séries temporais, aplicado às despesas de uma prefeitura no interior do ceará. **Brazilian Journal of Production Engineering**, v. 11, n. 2, p. 2–4, 2025. Disponível em: <https://doi.org/10.47456/bjpe.v11i2.47358>.

HAIDER, Nishi Shahnaj. Signal processing. **SciELO Brazil**, v. 61, n. 1, p. 3–6, 2018.

HATZE, H. The use of optimally regularized fourier series for estimating higher-order derivatives of noisy biomechanical data. **Journal of Biomechanics**, v. 14, p. 13–18, 1981.

HEATH, Steve. **Embedded Systems Design**. 2. ed. Oxford: Newnes, 2003.

HENDERI, H.; WAHYUNINGSIH, T.; RAHWANTO, E. Comparison of min-max normalization and z-score normalization in the k-nearest neighbor (knn) algorithm to test the accuracy of types of breast cancer. **International Journal of Informatics and Information Systems**, v. 4, n. 1, p. 13–20, 2021.

JAMES, J. F. **A Student's Guide to Fourier Transforms**. [S.l.]: Cambridge University Press, 2002.

KRECISZ, Krzysztof; BACZKOWICZ, Dawid. Analysis and multiclass classification of pathological knee joints using vibroarthrographic signals. **Computer Methods and Programs in Biomedicine**, v. 154, p. 37–44, 2018. ISSN 0169-2607. Disponível em: <https://www.sciencedirect.com/science/ARTICLE/pii/S0169260717301827>.

KREYSZIG, Erwin. **Matemática Superior para Engenharia**. 9. ed. [S.l.]: s.n., 2009. v. 2.

KUHL, F. P.; GIARDINA, C. R. Elliptic fourier features of a closed contour. **Computer Graphics and Image Processing**, Elsevier, v. 18, n. 3, p. 236–258, 1982. ISSN 0146-664X. Disponível em: [https://doi.org/10.1016/0146-664X\(82\)90034-X](https://doi.org/10.1016/0146-664X(82)90034-X).

KUMAR, Prathap; KUMAR, Arun; VENKATESH. Healthy gait: Review of anatomy and physiology of knee joint. **International Journal of Current Research and Review**, IJCRR, v. 12, n. 8, p. 2–6, 2020. Disponível em: [https://www.researchgate.net/publication/340230697\\_Healthy\\_Gait\\_Review\\_of\\_Anatomy\\_and\\_Physiology\\_of\\_Knee\\_Joint](https://www.researchgate.net/publication/340230697_Healthy_Gait_Review_of_Anatomy_and_Physiology_of_Knee_Joint).

LESTREL, Pete E. Method for analyzing complex two-dimensional forms: Elliptical fourier functions. **American Journal of Human Biology**, Wiley-Liss, Inc., A Wiley Company, v. 1, n. 2, p. 149–164, 1989. ISSN 1042-0533. Disponível em: <https://doi.org/10.1002/ajhb.1310010204>.

MAHAJAN, Palak; UDDIN, Shahadat; HAJATI, Farshid; MONI, Mohammad Ali. Ensemble learning for disease prediction: A review. **Healthcare**, v. 11, June 2023. Disponível em: [https://www.academia.edu/103981282/Ensemble\\_Learning\\_for\\_Disease\\_Prediction\\_A\\_Review](https://www.academia.edu/103981282/Ensemble_Learning_for_Disease_Prediction_A_Review).

MARTINEK, Radek; LADROVA, Martina; SIDIKOVA, Michaela; BEHBEHANI, Jaros Khosrow; KAHANKOVA, Radana; KAWALA-STERNIUK, Aleksandra. Advanced bioelectrical signal processing methods: Past, present and future approach—part i: Cardiac signals. **MDPI**, v. 21, n. 1, p. 3–6, 2021.

MCCOY, Gerald F.; MCCREA, John D.; BEVERLAND, David E.; KERNOHAN, W. George; MOLLAN, Raymond A. Vibration arthrography as a diagnostic aid in diseases of the knee a preliminary report. **The Journal of Bone and Joint Surgery**, v. 69 B, n. 2, 1987. Disponível em: [https://www.researchgate.net/publication/19342699\\_Vibration\\_arthrography\\_as\\_a\\_diagnostic\\_aid\\_in\\_diseases\\_of\\_the\\_knee\\_A\\_preliminary\\_report](https://www.researchgate.net/publication/19342699_Vibration_arthrography_as_a_diagnostic_aid_in_diseases_of_the_knee_A_preliminary_report).

MISHRA, Akanksha. A comprehensive review of artificial intelligence and machine learning: Concepts, trends, and applications. **International Journal of Scientific Research in Science and Technology**, v. 11, n. 5, p. 126–142, September 2024. ISSN 2395-602X. Disponível em: <https://www.ijrst.com>.

MOREIRA, D. B. F. **Classification of knee arthropathy with accelerometer-based vibroarthrography**. 2015. Tese (Tese (Doutorado)) — Universidade do Porto (Portugal), 2015.

NALBAND, S.; SUNDAR, A.; PRINCE, A. A.; AGARWAL, A. Feature selection and classification methodology for the detection of knee-joint disorders. **Comput. Methods Programs Biomed.**, v. 127, p. 94–104, Apr. 2016.

OLIVEIRA, H. L. C. **Análise do assincronismo entre remadores através de sinais biomecânicos e sua influência no desempenho em treinos técnicos de remo olímpico**. 2017. Dissertação (Dissertação de Mestrado) — Universidade Federal do Rio de Janeiro (UFRJ), 2017. Disponível em: <https://pantheon.ufrj.br/bitstream/11422/7856/1/877616.pdf>.

PARZEN, Elias. On estimation of a probability density function and a mode. **The Annals of Mathematical Statistics**, v. 33, n. 3, p. 1065–1076, 1962. Disponível em: <https://projecteuclid.org/euclid.aoms/1177704472>.

PASCHOA, Dr. Bruno. **Anatomia e Estrutura do Joelho**. 2022. Disponível em: <https://www.brunopaschoa.com/anatomia-e-estruturas-do-joelho/>.

PERCOPE, Marco Antônio. **Mais de cinco milhões de pessoas fizeram exames médicos no joelho no SUS só em 2021**. 2022. Disponível em: <https://ufmg.br/comunicacao/noticias/mais-de-cinco-milhoes-de-pessoas-fizeram-exames-medicos-no-joelho-no-sus-so-em-2021>.

PIMENTEL, Marco A.F.; CLIFTON, David A.; CLIFTON, Lei; TARASSENKO, Lionel. Signal processing. **Science Direct**, v. 99, n. 1, p. 215–249, 2014.

QUINTANILHA, Raphael Ferreira; RUPPERT, Guilherme Cesar Soares. **Estudo de Ferramentas para Ciência de Dados Aplicadas a Informações Médicas**. 2023. 2–8 p. Disponível em: <https://www.gov.br/cti/pt-br/publicacoes/producao-cientifica/jicc/xxv-jicc-2023/pdf/jicc-2023-paper-29.pdf>.

RAHMAN, Md. Asadur; RASHID, Mohd Abdur; AHMAD, Mohiuddin. Selecting the optimal conditions of savitzky–golay filter for fnirs signal. **Biocybernetics and Biomedical Engineering**, v. 39, n. 3, p. 624–637, 2019. ISSN 0208-5216. Disponível em: <https://www.sciencedirect.com/science/ARTICLE/pii/S0208521618305667>.

RAINIO, O.; TEUHO, J.; KLÉN, R. Evaluation metrics and statistical tests for machine learning. **Scientific Reports**, Nature Publishing Group, v. 14, p. 6086, 2024.

RANGAYYAN, Rangaraj M.; WU, Yunfeng. Modeling and classification of knee-joint vibroarthrographic signals using probability density functions estimated with parzen windows. **30th Annual International IEEE EMBS Conference Vancouver, British Columbia, Canada**, 2008. Disponível em: <https://ieeexplore.ieee.org/document/4649607>.

SATHYANARAYANAN, S.; TANTRI, B. Roopashri. Confusion matrix-based performance evaluation metrics. **African Journal of Biomedical Research**, v. 27, n. 4S, p. 4023–4031, nov. 2024. Disponível em: <https://africanjournalofbiomedicalresearch.com/index.php/AJBR/ARTICLE/view/4345>.

Scikit-learn Developers. **SVC — Support Vector Classification**. [S.l.], 2025. Available at: <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>. Accessed: 16 Nov. 2025.

SILVA, Paulo Augusto; CAMARGO, Matheus Gomes. Vibroartrografia de joelho. **Diagnóstico de Lesão no Ligamento Cruzado Anterior Por Vibroartrografia e Machine Learning**, 2025. Disponível em: <http://repositorio.utfpr.edu.br/jspui/handle/1/5353>.

SKOCZYLAS, Artur; SZTANGRET, Łukasz; NOWICKI, Michał. Comparison of selected classification methods based on machine learning as a diagnostic tool for knee joint cartilage damage based on generated vibroacoustic processes. **Applied Computer Science**, Lublin University of Technology, v. 19, n. 4, p. 49–65, 2023. Disponível em: <https://ph.pollub.pl/index.php/acs/ARTICLE/view/5558>.

SOUDAN, Karel; DIERCKX, Paul. Calculation of derivatives and fourier coefficients of human motion data, while using spline functions. **Journal of Biomechanics**, v. 12, p. 21–26, 1979.

SOUZA, Hellen Priscila Santos; MORAIS, Aline Cristina Batista Resende de. **Perfil Epidemiológico dos Pacientes com Lesões em Joelhos Atendidos em uma Clínica de Ortopedia na Cidade de Goiânia (GO)**. 2016. 205–207 p. Disponível em: [https://web.archive.org/web/20180417053744id\\_/http://www.revista.ueg.br/index.php/movimenta/ARTICLE/viewFile/3957/3492](https://web.archive.org/web/20180417053744id_/http://www.revista.ueg.br/index.php/movimenta/ARTICLE/viewFile/3957/3492).

STEIN, Elias M.; SHAKARCHI, Rami. **Fourier Analysis: An Introduction**. Princeton, NJ: Princeton University Press, 2003. v. 1. (Princeton Lectures in Analysis, v. 1). ISBN 9780691113845.

SUBASI, Abdulhamit. Classification of emg signals using pso optimized svm for diagnosis of neuromuscular disorders. **Computers in Biology and Medicine**, Elsevier, v. 43, p. 576–586, 2013.

SUBRAMANIAN, Jyothi; SIMON, Richard. Overfitting in prediction models – is it a problem only in high dimensions? **ELSEVIER**, v. 36, n. 2, p. 636–641, 2013.

TEIXEIRA, Gabriel dos Anjos Dantas. **SISTEMA DE DETECÇÃO E CLASSIFICAÇÃO DE NOTÍCIAS FALSAS COM APRENDIZADO DE MÁQUINA**. 2022. <https://bdta.abcd.usp.br/directbitstream/11507488-3dc3-4d25-a0cc-5b33050d763a/GabrieldosAnjosDantasTeixeiraTCC-PTC22.pdf>. Acesso em: 10 jun. 2025.

The MathWorks, Inc. **MATLAB - MathWorks**. 2025. Disponível em: <https://www.mathworks.com/products/matlab.html>. Acesso em: 12 jun. 2025.

TOCQUEVILLE, Sophie de; MARJIN, Mihaela; RUZEK, Michal. A review of the vibration arthrography technique applied to the knee diagnostics. **Applied Sciences** **11**, n. 16, 2021. Disponível em: <https://doi.org/10.3390/app11167337>.

VILLANI, Valeria; SABATTINI, Lorenzo; CZERNIAKI, Julia N.; MERTENS, Alexander; VOGEL-HEUSER, Birgit; FANTUZZI, Cesare. Towards modern inclusive factories: A methodology for the development of smart adaptive human-machine interfaces. *In: 2017 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. [S.l.: s.n.], 2017. p. 1–7.

WEISZ, Ferenc. **Summability of Multi-Dimensional Trigonometric Fourier Series**. Basel: Birkhäuser, 2002. ISBN 978-3-0348-8203-9.

WinPython. **WinPython: Free open-source portable distribution of Python for Windows**. 2025. <https://github.com/winpython/winpython>. Acesso em: 21 out. 2025.

WU, Yufeng; KRISHMAN, Sridhar. Classification of knee-joint vibroarthrographic signals using time-domain and time-frequency domain features and least-squares support vector machine. **16th International Conference on Digital Signal Processing, Santorini, Greece**, 2009. Disponível em: <https://ieeexplore.ieee.org/document/5201156>.

WU, Y. *et al.* Quantification of knee vibroarthrographic signal irregularity associated with patellofemoral joint cartilage pathology based on entropy and envelope amplitude measures. **Comput. Methods Programs Biomed.**, v. 130, p. 1–12, Jul. 2016.