

UNIVERSIDADE DE CAXIAS DO SUL
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
COORDENADORIA DE PESQUISA E PÓS-GRADUAÇÃO *STRICTO SENSU*
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA E CIÊNCIA DOS
MATERIAIS

**Contribuição para o desenvolvimento de um *software* de diagnóstico
baseado em sensores magnetoelásticos**

Alessandro Josué da Silva Andreatta

CAXIAS DO SUL

2024

Alessandro Josué da Silva Andreatta

Contribuição para o desenvolvimento de um *software* de diagnóstico baseado em sensores magnetoelásticos

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Engenharia e Ciência dos Materiais, pelo Programa de Pós-Graduação em Engenharia e Ciência dos Materiais da Universidade de Caxias do Sul.

Orientadora: Profa. Dra. Janete E. Zorzi

Coorientador: Prof. Dr. Cesar Aguzzoli

CAXIAS DO SUL

2024

Dados Internacionais de Catalogação na Publicação (CIP)
Universidade de Caxias do Sul
Sistema de Bibliotecas UCS - Processamento Técnico

A557c Andreatta, Alessandro Josué da Silva

Contribuição para o desenvolvimento de um software de diagnóstico baseado em sensores magnetoelásticos [recurso eletrônico] / Alessandro Josué da Silva Andreatta. – 2024.

Dados eletrônicos.

Dissertação (Mestrado) - Universidade de Caxias do Sul, Programa de Pós-Graduação em Engenharia e Ciência dos Materiais, 2024.

Orientação: Janete Eunice Zorzi.

Coorientação: Cesar Aguzzoli.

Modo de acesso: World Wide Web

Disponível em: <https://repositorio.ucs.br>

1. Materiais. 2. Magnetostricção. 3. Magnetismo. 4. Elasticidade. 5. Maquinas - Projetos. I. Zorzi, Janete Eunice, orient. II. Aguzzoli, Cesar, coorient. III. Título.

CDU 2. ed.: 620.1

Catalogação na fonte elaborada pela(o) bibliotecária(o)
Carolina Machado Quadros - CRB 10/2236

Alessandro Josué da Silva Andreatta

Contribuição para o desenvolvimento de um *software* de diagnóstico baseado em sensores magnetoelásticos

Dissertação apresentada como requisito parcial para obtenção do título de Mestre em Engenharia e Ciência dos Materiais, pelo Programa de Pós-Graduação em Engenharia e Ciência dos Materiais da Universidade de Caxias do Sul.

Banca Examinadora:

Dra. Janete Eunice Zorzi
Orientadora
Universidade de Caxias do Sul

Dr. Cesar Aguzzoli
Coorientador
Universidade de Caxias do Sul

Dr. Giovani Luís Rech

Dr. André Luís Martinotto
Universidade de Caxias do Sul

Dr. Leandro Luís Corso
Universidade de Caxias do Sul

AGRADECIMENTOS

A Deus, pelo dom da vida.

Agradeço à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de mestrado, essencial para o desenvolvimento desta pesquisa. Também estendo meus agradecimentos à Universidade de Caxias do Sul (UCS) por proporcionar a oportunidade de realizar este estudo.

À professora Dra. Janete Eunice Zorzi, expresse meus sinceros agradecimentos pela paciência, orientação e apoio constantes ao longo deste trabalho. Seu amor pela pesquisa se manifestou em cada conversa sobre o projeto.

Reconheço igualmente o professor Dr. Cláudio Antônio Perottoni pelo valioso suporte e orientação, fundamentais para direcionar meus esforços de pesquisa corretamente.

Ao professor Dr. Cesar Aguzzoli, agradeço pela generosidade em disponibilizar os recursos necessários para a conclusão deste projeto e pela coorientação prestada.

À professora Dra. Mariana Roesch-Ely, agradeço por prover materiais, espaços e tecnologia, sem os quais este trabalho não seria possível.

À Luiza Felippi de Lima, cujo empenho na geração dos materiais essenciais para esta pesquisa foi inestimável. Sua contribuição com valiosas percepções impactou positivamente a realização das atividades.

À minha família, pela paciência nos momentos de ausência, e especialmente ao amor da minha vida, José Luiz, por ser meu maior incentivador e apoiador em todas as circunstâncias.

RESUMO

Nas últimas décadas, ocorreram várias epidemias de vírus mortais com grande impacto na saúde pública global. Exemplos incluem o Ebola, o Marburg, o Nipah, o Zika, a febre hemorrágica da Crimeia-Congo, Dengue e, mais recentemente, o coronavírus SARS-CoV-2, que causou a pandemia de COVID-19. A rápida disseminação desses vírus e a falta de infraestrutura de saúde adequada em algumas regiões tornaram difícil conter as epidemias. A mobilização global foi necessária para conter a disseminação e fornecer assistência médica adequada às pessoas afetadas. A detecção rápida e o desenvolvimento de medicamentos e vacinas eficazes são cruciais para combater futuras epidemias. Sensores feitos com materiais magnetoelásticos e modelos de aprendizado de máquina podem ser úteis nessa corrida. Neste trabalho, a análise de um conjunto de dados coletados em um analisador de rede, a partir de sensores magnetoelásticos, foi utilizada para treinar um algoritmo de aprendizado de máquina. Um modelo de classificação foi desenvolvido para se obter o diagnóstico correto da presença de patógenos nos sensores de teste, com a exportação do mesmo para uso em dados externos. O modelo desenvolvido foi capaz de distinguir entre dados de controle e de teste, demonstrando precisão na identificação da presença de patógenos. A aplicação de aprendizado de máquina permitiu a detecção de pequenas variações nos sinais dos sensores, permitindo o uso para diagnóstico. A quantidade de dados disponíveis para o treinamento do modelo limitou parcialmente o aprofundamento da análise, não comprometendo, contudo, a técnica aplicada.

Palavras-chave: epidemia, vírus, COVID-19, sensores magnetoelásticos, aprendizado de máquina, diagnóstico.

ABSTRACT

In recent decades, there have been several epidemics of deadly viruses with significant impact on global public health. Examples include Ebola, Marburg, Nipah, Zika, Crimean-Congo hemorrhagic fever, Dengue, and more recently, the SARS-CoV-2 coronavirus, which caused the COVID-19 pandemic. The rapid spread of these viruses and the lack of adequate healthcare infrastructure in some regions have made it difficult to contain the epidemics. Global mobilization was necessary to contain the spread and provide adequate medical assistance to affected individuals. Rapid detection and the development of effective drugs and vaccines are crucial to combating future epidemics. Sensors made from magnetoelastic materials and machine learning models can be useful in this endeavor. In this work, the analysis of a dataset collected from a network analyzer using magnetoelastic sensors was employed to train a machine learning algorithm. A classification model was developed to accurately diagnose the presence of pathogens in test sensors, with the capability of exporting the model for use with external data. The developed model was able to distinguish between control and test data, demonstrating precision in identifying the presence of pathogens. The application of machine learning enabled the detection of small variations in the sensor signals, facilitating its use for diagnostic purposes. The amount of data available for model training partially limited the depth of the analysis; however, it did not compromise the applied technique.

Keywords: epidemic, viruses, COVID-19, magnetoelastic sensors, machine learning, diagnose.

LISTA DE FIGURAS

Figura 1 - Representação da influência dos domínios magnéticos na deformação de materiais ferromagnéticos.....	15
Figura 2 - Representação de um sistema de detecção de frequência de biossensores. O sinal vibracional de entrada se propaga longitudinalmente. Adaptado da Ref. [46].....	19
Figura 3 - Sensores magnetoelásticos similares aos usados neste trabalho para diagnóstico de SARS-CoV-2. As dimensões são 5 mm x 1 mm. Adaptado da Ref. [47].	25
Figura 4 - Analizador de rede Keysight E5061B.....	26
Figura 5 - Modelo esquemático do sistema de detecção utilizando sensores magnetoelásticos e analisador de rede. Adaptado da Ref. [74].	26
Figura 6 - Modelo esquemático do processo de aquisição de dados.	27
Figura 7 - Dados iniciais do sinal versus frequência, identificados pelo hash.....	28
Figura 8 - Representação do conjunto de dados utilizado para análise, mostrando a coluna de hash usada para correlacionar diferentes conjuntos de dados.....	28
Figura 9 - Variação da frequência no tempo.	30
Figura 10 - Fluxograma do funcionamento do programa desenvolvido em Python.	32
Figura 11 - Posição do ponto mínimo de frequência e sinal.	33
Figura 12 - Comparação entre dados reais e ajustados com LMFIT e SMC.	34
Figura 13 - Índice de acurácia entre LMFIT e a simulação de Monte Carlo.....	34
Figura 14 - Amostra do conjunto de dados de médias minimizadas.....	35
Figura 15 - Deslocamento das frequências. O tempo do ensaio é demonstrado pelo tamanho do marcador. S4, S5 e S6 se referem aos sensores 4, 5 e 6, do grupo de teste. Sensor 2 no detalhe mostrando alguma variação de frequência.	35
Figura 16 - Comparação entre a variação de frequência entre dois sensores (controle e teste).	36
Figura 17 - Ajuste dos dados de frequência pelo tempo utilizando a equação de Lagergren.	37
Figura 18 - Deslocamentos de frequência pelo tempo para o sensor 5 (teste).....	38
Figura 19 - Ajuste das frequências à equação de Lagergren para o sensor 5. O comentário adicional faz referência a $f(t=0) - f(t=30)$, como meio de se observar uma aproximação fora do ajuste.	39
Figura 20 - Agrupamento das curvas de deslocamento por grupo.	40
Figura 21 - Agrupamento das curvas de deslocamento por grupo em gráfico de caixa.....	41
Figura 22 - Distribuição do deslocamento de frequências por grupo em gráfico de violino e caixa para os dados gerados pelo Bootstrap.	41
Figura 23 - Matriz de confusão: XGBoost.	42

LISTA TABELAS

Tabela 1 - Definição de métricas utilizadas em modelos de classificação.....	22
Tabela 2 - Constituição do conjunto de dados analisado.	28
Tabela 3 - Deslocamentos de frequência total para os sensores de teste e controle..	38
Tabela 4 - Métricas de avaliação dos classificadores.....	42

SIGLAS, SÍMBOLOS, TERMOS E ABREVIACÕES

CCHF – Febre hemorrágica da Crimeia-Congo

SME – Sensores magnetoelásticos

BSME – Biossensores magnetoelásticos

IgG – Imunoglobulina G

μ_{AC} – Permeabilidade magnética

σ – Tensão

λ – Constante de magnetostricção

M – Magnetização

FMG – Vidros metálicos ferromagnéticos (do inglês *Ferromagnetic metallic glasses*)

E – Módulo de elasticidade

ν – Coeficiente de Poisson

ρ – Densidade

PVD – Deposição física por vapor (do inglês *Physical vapor deposition*)

IA – Inteligência artificial

AM – Aprendizado de máquina

RN/RNA – Redes neurais artificiais

AC e DC – Corrente alternada e corrente contínua

SMC – Simulação de Monte Carlo

BNP – *Bootstrap* não paramétrico

IC – Intervalo de confiança

SUMÁRIO

1. INTRODUÇÃO	10
2. OBJETIVO GERAL	14
2.1. OBJETIVOS ESPECÍFICOS	14
3. REFERENCIAL TEÓRICO.....	15
3.1. MAGNETOELASTICIDADE	15
3.2. SENSORES MAGNETOELÁSTICOS.....	16
3.2.1. <i>Aplicações de SME</i>	17
3.2.2. <i>Princípio de funcionamento de SME</i>	18
3.3. INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA	20
3.4. <i>EXTREME GRADIENT BOOSTING</i>	23
3.5. <i>MÉTODO BOOTSTRAP</i>	24
4. MATERIAIS E MÉTODOS	25
4.1. MATERIAIS.....	25
4.1.1. <i>Biossensores</i>	25
4.1.2. <i>Ambiente</i>	26
4.2. MÉTODOS.....	27
5. RESULTADOS E DISCUSSÃO	33
5.1. AJUSTES INICIAIS.....	33
5.2. ANÁLISE DOS DESLOCAMENTOS DE FREQUÊNCIAS	35
5.3. CLASSIFICADOR E ESTIMATIVA DE DIAGNÓSTICO	42
6. CONSIDERAÇÕES FINAIS E CONCLUSÕES	44
REFERÊNCIAS	47

1. INTRODUÇÃO

Nos últimos 10 anos, ocorreram várias epidemias de vírus mortais que tiveram grande impacto na saúde pública global. Em 2014, um surto de ebola na África Ocidental matou mais de 11.000 pessoas, com a maioria dos casos ocorrendo na Guiné, Serra Leoa e Libéria. A rápida disseminação do vírus, juntamente com a falta de infraestrutura de saúde adequada na região, dificultou a contenção da epidemia. Esforços de colaboração global foram necessários para conter a disseminação do vírus e fornecer assistência médica adequada para as pessoas afetadas [1,2].

O vírus Marburg, que é semelhante ao vírus Ebola, causa febre hemorrágica. Este vírus foi identificado pela primeira vez em 1967 na cidade alemã de Marburg e, desde então, houve vários surtos em países africanos, como Uganda, Angola e Congo. A transmissão do vírus ocorre por meio do contato direto com fluidos corporais de pessoas infectadas ou animais. A taxa de mortalidade é alta, variando de 24 % a 88 %, dependendo do surto. Não existe uma vacina ou tratamento específico para o vírus, e o tratamento se concentra no gerenciamento dos sintomas e na prevenção de complicações [3,4].

Outro exemplo de vírus mortal ocorrido nos últimos 10 anos foi o vírus *Nipah*, que surgiu na Malásia em 1998, com surtos subsequentes em outros países, incluindo Bangladesh e Índia. Em maio de 2018, um surto de *Nipah* ocorreu em Kerala, sul da Índia, que resultou em 17 mortes. O vírus é transmitido por morcegos frugívoros e pode ser transmitido de pessoa para pessoa. A rápida identificação e contenção do surto pelo governo indiano e pelas organizações internacionais de saúde ajudaram a impedir a disseminação do vírus para outras áreas [5,6].

O vírus da Zika, ou Zika vírus, é outro exemplo de um patógeno agressivo que afetou a saúde pública global. Um surto de Zika vírus foi registrado pela primeira vez no Brasil em 2015 e se espalhou rapidamente para outras partes do mundo. Embora a maioria das pessoas infectadas com o vírus não apresente sintomas ou apenas sintomas leves, o vírus pode causar microcefalia em bebês de mulheres infectadas durante a gravidez. Até o momento, não existe uma vacina ou tratamento específico para o vírus, e a prevenção se concentra em medidas de controle de mosquitos e práticas sexuais seguras [7,8].

A febre hemorrágica da Crimeia-Congo (CCHF) é outra doença viral que causa febre, dores musculares e sangramentos. O vírus é transmitido por carrapatos, e afeta principalmente pessoas que trabalham em áreas rurais ou que têm contato com animais infectados. Desde 2010, houve vários surtos de CCHF em todo o mundo, incluindo países do Oriente Médio, Europa e Ásia. A taxa de mortalidade é alta, variando de 10 % a 40 %, dependendo do surto. O tratamento se concentra no gerenciamento dos sintomas e na prevenção de complicações [9,10].

Ainda, possivelmente o mais emblemático vírus que afetou o mundo nos últimos anos, especialmente pela quantidade de casos e mortes, e a velocidade de propagação, foi a do coronavírus SARS-CoV-2, que causou a pandemia de COVID-19. A doença foi inicialmente detectada na China em dezembro de 2019 e rapidamente se espalhou para outros países, tornando-se uma pandemia global em março de 2020. Até o momento, mais de 700 milhões de casos foram confirmados em todo o mundo, com quase sete milhões de mortes. O vírus afetou quase todos os aspectos da vida humana, causando enormes danos à economia global e à saúde mental das pessoas [11,12].

Por fim, o avanço recente da dengue no Brasil tem sido uma preocupação significativa tanto para profissionais de saúde pública quanto para pesquisadores em nível global. Segundo dados do Ministério da Saúde, entre 2019 e 2022, foram registrados mais de 3 milhões de casos de dengue em todo o país, representando um aumento de mais de 20 % em relação ao período anterior. Estudos epidemiológicos têm destacado fatores como urbanização desordenada, mudanças climáticas e a presença de criadouros do mosquito *Aedes aegypti* como impulsionadores desse aumento. Além disso, a disseminação de diferentes sorotipos do vírus da dengue e a falta de imunidade em grande parte da população tem contribuído para a recorrência dos surtos [13, 14]. Segundo a OPAS, 2023 registrou o maior número de casos de infecção pelo vírus da dengue nas Américas, somando mais de 4,5 milhões de casos e mais de duas mil mortes. Esse impacto se intensificou no início de 2024, em que os números apontam para um crescimento da ordem de 157 % em relação ao mesmo período de 2023 e 225 % maior do que a média dos últimos cinco anos [15].

A mobilização global em torno dessas epidemias demonstra a necessidade de velocidade na criação de meios de detecção rápidos e precisos ou desenvolver medicamentos e vacinas eficazes. Para tanto, sensores fabricados com materiais magnetoelásticos, associados a modelos de aprendizado de máquina, que podem ser treinados com diferentes patógenos, podem ser um aliado na corrida por mecanismos de detecção rápida.

Materiais magnetoelásticos possuem a propriedade de vibrar quando submetidos a excitação magnética. Sensores feitos a partir desses materiais utilizam essa resposta vibracional, que pode ser detectada por instrumentos magnéticos, acústicos e ópticos, para monitoramento remoto [16]. Os sensores magnetoelásticos (SME) têm várias vantagens em relação a outros tipos de sensores. Eles são sensíveis, precisos e confiáveis, e podem operar em uma ampla faixa de temperaturas e frequências. Além disso, eles são menos suscetíveis a interferências eletromagnéticas do que outros sensores, o que os torna adequados para ambientes hostis ou de alta interferência. Ainda, SME têm facilidade de serem integrados à dispositivos com eletrônica embarcada, que dispensa conexão por cabos, e a podem ser produzidos em larga escala por meio do uso de monolitos do material base para o sensor [17].

Essas características fazem com que os SME, que são geralmente fabricados em formato de filmes finos de ligas de metal amorfo, encontrem aplicações no monitoramento remoto de diversos parâmetros, tais como pH, temperatura, pressão, tensão, deformação, variação de massa, umidade, variação de fluxo, entre outros [18].

A biomedicina é uma área que têm se beneficiado de sensores construídos a partir de materiais magnetoelásticos para utilização específica nesse campo, que são os Biossensores Magnetoelásticos (BSME). Eles são aplicados no monitoramento de parâmetros biológicos que vão desde a adesão e deformação de implantes ósseos [19-21] à concentração de certas substâncias, como glicose e dióxido de carbono [22, 23].

Outro fenômeno relevante na biomedicina que pode ser monitorado por um BSME é a detecção da presença de bactérias [24, 25, 26]. Por meio de engenharia de superfície, bactérias aderem a SME revestidos, de modo que a massa total do

sistema é alterada e, por consequência, o sinal de resposta emitido pelo sensor se modifica de forma quantificável [27]. Essa técnica pode ser empregada na detecção de infecções por determinados patógenos, através da deposição de uma fina camada de proteínas características do processo imunológico, como a imunoglobulina G (ou IgG), por exemplo, sobre um BSME revestido com um composto biocompatível, para diagnósticos rápidos da presença ou não desses patógenos [28]. Uma abordagem similar pode ser aplicada para a detecção do coronavírus SARS-CoV-2, objeto deste estudo, como demonstrado em trabalhos recentes [29-31].

A partir dessa perspectiva, o presente trabalho visa obter um *software* que utiliza BSME, com a aplicação de aprendizado de máquina para classificação, tendo por base os resultados de experimentos prévios para determinar a presença ou não de vírus. O emprego de métodos de inteligência artificial para a criação de um modelo otimizado de detecção visa reduzir custos com muitos testes para seleção de parâmetros. Dessa forma, é aumentar a sensibilidade e assertividade do diagnóstico, além de permitir a aplicação do modelo de detecção de forma simplificada e embarcada, em telefones móveis, por exemplo.

A relevância desta abordagem fica evidente ao se tomar como pano de fundo o cenário pandêmico que se desenvolveu entre 2020 e 2022, que ainda hoje manifesta seus efeitos, além do avanço dos casos de dengue no Brasil. Tal contexto mostrou a necessidade de instrumentos para diagnóstico rápido de determinadas doenças, como a COVID-19 e outras viroses, serem obtidos em um curto espaço de tempo, com eficiência de detecção elevada e custo reduzido.

2. OBJETIVO GERAL

Desenvolver um *software* para análise de dados de sensores magnetoelásticos usados na detecção de patógenos, utilizando análise estatística de dados e aprendizado de máquina.

2.1. OBJETIVOS ESPECÍFICOS

- Avaliar dados de sensores magnetoelásticos funcionalizados quanto à presença de vírus – SARS-CoV-2;
- Desenvolver um modelo de previsão com base no ajuste de curvas de dados de um analisador de rede;
- Propor um sistema capaz de avaliar com elevada precisão a presença ou não de vírus utilizando um algoritmo classificador de aprendizado de máquina.

3. REFERENCIAL TEÓRICO

3.1. MAGNETOELASTICIDADE

Desde 1842, quando Joule descobriu o efeito da magnetostricção, que se caracteriza pela deformação de uma amostra de material ferromagnético em razão da aplicação de campos magnéticos externos, diversas outras manifestações desse efeito foram observadas e estudadas. Na magnetostricção, uma rotação na direção da magnetização causa uma tensão mecânica. Magnetoelasticidade é um efeito muito próximo à magnetostricção, que pode ser definida como a mudança na magnetização de um material causada pela aplicação de tensão ou deformação em um material ferromagnético (Figura 1) [32].

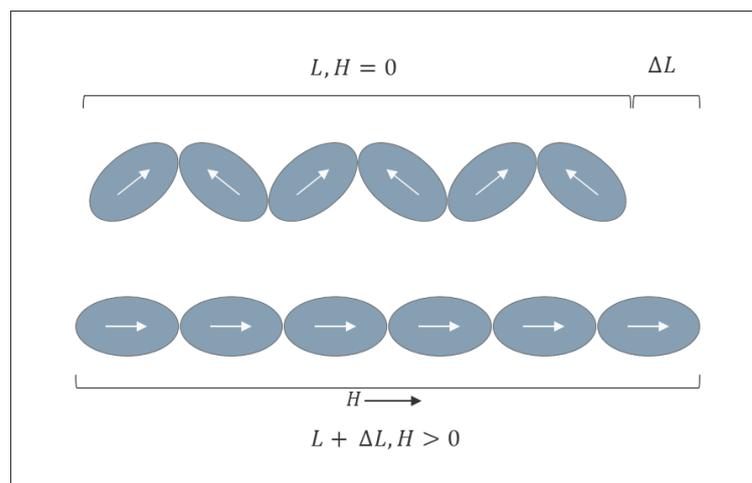


Figura 1 - Representação da influência dos domínios magnéticos na deformação de materiais ferromagnéticos.

A relação entre o efeito da magnetostricção e o da magnetoelasticidade (eventualmente, chamado de efeito inverso de magnetostricção) existe porque tensão mecânica aplicada a um material altera sua permeabilidade magnética CA (μ_{CA}). A alteração ocorre quando a tensão σ e a permeabilidade μ_{AC} são perpendiculares à direção mais favorável à magnetização do material, e pode ser descrita por

$$\mu_{CA} = \frac{\mu_0 M^2}{2K - 3\lambda\sigma} + 1 \quad (1)$$

com variação da permeabilidade CA pela tensão aplicada dada por

$$\frac{d\mu_{CA}}{d\sigma} = \frac{3\lambda\mu_0 M^2}{(2K - 3\lambda\sigma)^2}, \text{ para } 2K > 3\lambda\sigma \quad (2)$$

onde, λ é a constante de magnetostricção, K é a constante anisotrópica e M , a magnetização. Com $\lambda\sigma$ se aproximando de $2K/3$, obtém-se a variação máxima de μ_{CA} , ou a sensibilidade à tensão. Se $\lambda\sigma$ é maior que $2K/3$, $\mu_{CA} = 0$, pois a direção da magnetização M gira em 90° . Isso significa dizer, com base no entendimento de como a mudança de direção da magnetização ocorre, que é possível obter informações sobre o comportamento magnetoelástico de um material [33].

Esses materiais também respondem com vibrações mecânicas ao estímulo de um campo magnético aplicado. Quando submetidos a interferências externas, como compressão, estiramento, cisalhamento, variações de pressão e temperatura, a resposta emitida por esses sólidos varia de um modo característico, intrinsecamente relacionado ao campo magnético aplicado e a sua composição [34].

Vidros metálicos ferromagnéticos (FMG – *ferromagnetic metallic glasses*), que são ligas ferromagnéticas amorfas, por exemplo, apresentam uma combinação peculiar de propriedades magnéticas, elásticas e magnetostrictivas [35]. Embora todos os materiais sejam elásticos, apenas uma parcela manifesta alguma medida de ferromagnetismo. Ainda uma menor fração desses materiais apresentam algum tipo de comportamento magnetoelástico. De acordo com Engdahl, magnetoelasticidade pode ser definida como uma união entre as propriedades clássicas da elasticidade e deformação de sólidos, e os fenômenos quânticos e relativísticos do magnetismo [36].

3.2. SENSORES MAGNETOELÁSTICOS

Sensores magnetoelásticos são sensores que usam o efeito magnetoelástico para medir uma grandeza física, como tensão, pressão, temperatura, deformação ou corrente elétrica. O efeito magnetoelástico ocorre em materiais magnéticos quando são submetidos a um campo magnético externo, que causa uma mudança na sua

forma. Essa mudança de forma é detectada pelo sensor e convertida em um sinal elétrico que pode ser usado para medir a grandeza física [16].

SME são fabricados a partir de ligas metálicas amorfas, geralmente ricas em elementos ferromagnéticos, como ferro, níquel e cobalto, na forma de fitas. É possível, ainda, encontrar ligas ricas em manganês, zinco, gálio e alumínio, e outras com adição de elementos da família dos lantanídeos, como térbio e disprósio. Algumas ligas são bastante utilizadas pela indústria de sensores, como é o caso da $Fe_{40}Ni_{38}Mo_4B_{18}$ e da $Fe_{81}B_{13.5}Si_{3.5}C_2$, conhecidas sob o nome comercial de Metglas 2826MB e Metglas 2605SC, que fazem referência tanto ao conceito de liga amorfa (*metallic glass*, ou vidro metálico) quanto aos elementos de dopagem, como boro, silício e carbono. Eles apresentam propriedades de interesse, como alta resistência à tração, que pode variar de 1000 a 1700 MPa, coeficiente de magnetoelasticidade da ordem de 0,98 e magnetostricção de até 10^{-5} , associado ao baixo custo do material. As vibrações que estes materiais emitem quando submetidos ao impulso de campos magnéticos podem ser detectadas por meio da amplitude da modulação de feixes de laser, acusticamente, através de diferentes tipos de microfones, ou através de bobinas de detecção de fluxo magnético emitido pelo sensor [34,37-44].

3.2.1. Aplicações de SME

SME supersensíveis são empregados na quantificação de tensão e deformação. Esse tipo de sensor, baseado em materiais magnéticos de alta frequência, é utilizado para monitoramento sem fio de pneus e eixos de máquinas, e são de grande interesse pela capacidade de serem miniaturizados e acessados à distância [33].

Na construção civil, sensores feitos a partir de filmes de materiais ferromagnéticos são capazes de detectar vibrações de forma muito mais precisa do que medidores convencionais, como extensômetros resistivos, por exemplo. A amplitude das ondas de ressonância magnetoelásticas, depende de fatores como permeabilidade magnética e coeficientes elásticos do material de que é feito o sensor, de tal modo que é possível detectar a interferência de efeitos físicos externos, como a vibração de uma construção, por exemplo, medindo a variação dessas propriedades com grande sensibilidade [18].

Na área biológica, uma vez que o corpo humano possui muitos processos enzimáticos extremamente sensíveis à variação de pH, e grande parte desses processos operam de forma ótima na faixa entre 7,38 e 7,42, é importante monitorar a variação do caráter ácido/alcalino do plasma no organismo. O desequilíbrio nesses níveis pode levar ao aumento da concentração de CO₂ no corpo, bem como o metabolismo ineficiente de determinadas gorduras e proteínas. Em pacientes obesos, o excesso de CO₂ pode levar a perda de capacidade respiratória em razão da redução do pH do plasma [45]. Khan *et al.* apresentam uma proposta de sensor formado por uma fita de material magnetoelástico que vibra a uma frequência conhecida quando submetido a um campo magnético. A fita é coberta por uma camada de um material sensível ao pH do meio, de forma que a variação do pH faz com que a massa desse material também varie. Se a massa aumenta, a frequência de vibração diminui, e vice-versa [47].

Sob princípios similares, há sensores sendo utilizados para detecção simultânea de pH, temperatura e pressão, como demonstrado por Jain *et al.* Em seu trabalho, os autores desenvolveram um sensor composto de quatro partes, contendo três sensores semelhantes, tensionados, feitos de diferentes ligas de Metglas® (MG2826MB e MG2605SC), cuja comparação permite medir a temperatura absoluta em um ambiente em que a pressão varia, e um sensor não tensionado fabricado com liga MG2826MB, com deposição de um polímero sensível à variação de pH. Utilizando um sistema em que uma bobina estimula o sensor magneticamente, e outra captura o fluxo magnético emitido por ele, foi possível obter leituras precisas de temperatura, pressão e pH [44,48].

3.2.2. Princípio de funcionamento de SME

Os SME apresentam uma frequência fundamental, que leva em conta o comprimento L do sensor, seu módulo de elasticidade E , coeficiente de Poisson ν e densidade ρ da liga da qual é feito, definido pela Equação (3). A variação da frequência fundamental ocorre com a mudança da massa.

$$f_0 = \frac{1}{2L} \sqrt{\frac{E}{\rho(1-\nu)}} \Rightarrow \Delta f = \frac{f_0 \Delta m}{2\rho Lwh} \quad (3)$$

A frequência fundamental do sensor, contudo, pode ser afetada por forças de amortecimento, entre as quais aquela causada pela variação de massa na superfície do sensor. A depender do meio em que o SME é submetido para análise, como no caso de sensores funcionalizados para detecção de patógenos, a variação da massa causada pela adesão de vírus ou bactérias faz com que a frequência de ressonância do sensor diminua, como efeito da diferença nas oscilações vibracionais [16].

Para a detecção dessas variações, é necessário um sistema composto por um excitador, como uma bobina geradora de campo AC, e um coletor, como um microfone ou uma bobina receptora, que recebe o sinal de vibração (Figura 2). A comparação entre sensores funcionalizados com e sem a presença do patógeno é capaz de fornecer os dados da amplitude da variação da frequência [46].

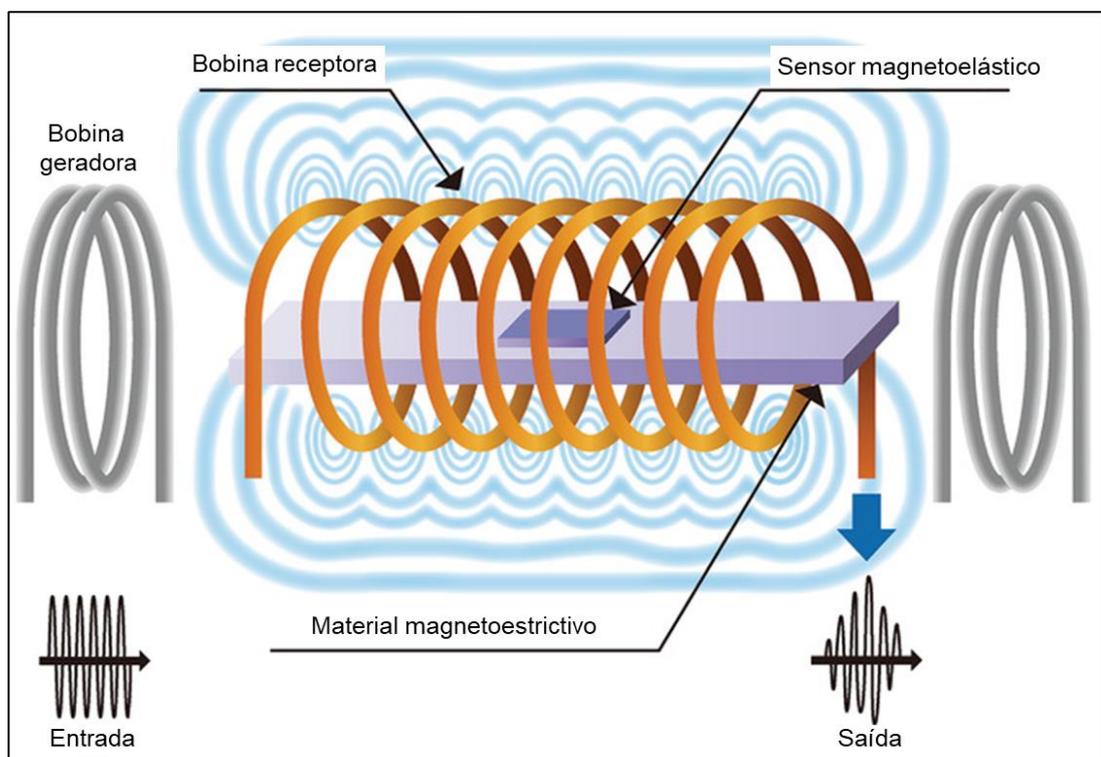


Figura 2 - Representação de um sistema de detecção de frequência de biossensores. O sinal vibracional de entrada se propaga longitudinalmente. Adaptado da Ref. [46].

Para a funcionalização, uma fina camada de cromo é adicionada à superfície da fita do sensor, seguida por uma camada de ouro, ambas aplicadas por meio de *magnetron sputtering*, uma técnica de deposição física de vapor (*PVD*). O cromo tem a função de aumentar a aderência do ouro sobre o substrato do material magnetoelástico e o ouro atua como promotor da biocompatibilidade com as

camadas biológicas do biossensor, assim como melhora a resistência à corrosão do BSME [49].

Posteriormente, anticorpos do tipo IgG, referentes aos antígenos do patógeno que se deseja detectar, são adicionados ao recobrimento de ouro, de modo que o sítio de ligação do antígeno fique exposto. Finalmente, um bloqueio é aplicado para evitar ligações inespecíficas indesejadas de outros compostos à superfície. Quando na presença do antígeno, este se conecta ao sítio exposto do anticorpo, alterando a massa do sensor e, conseqüentemente, diminuindo sua frequência fundamental de vibração [16,46], de tal modo que é possível realizar o diagnóstico.

Nesse sentido, como a ordem de magnitude da variação da frequência fundamental é pequena (menos de 100 Hz), a aplicação de aprendizado de máquina pode ser um recurso importante na detecção de pequenas variações no sinal do sensor.

3.3. INTELIGÊNCIA ARTIFICIAL E APRENDIZADO DE MÁQUINA

Inteligência artificial é um termo cunhado por J. McCarthy em 1956, que o definiu como a “engenharia e ciência de construir e tornar máquinas inteligentes, especialmente programas de computador”. É um ramo da ciência da computação que busca desenvolver sistemas capazes de realizar tarefas que requerem inteligência humana, como reconhecimento de padrões, tomada de decisão e resolução de problemas. A inteligência artificial se baseia em técnicas de aprendizado de máquina, processamento de linguagem natural, visão computacional e outros métodos para criar sistemas que podem analisar dados, aprender com eles e tomar decisões com base nas informações disponíveis [51-53].

Os sistemas de inteligência artificial podem ser divididos em duas categorias principais: sistemas baseados em regras e sistemas baseados em aprendizado de máquina. Os sistemas baseados em regras usam uma série de regras lógicas e heurísticas para tomar decisões, enquanto os sistemas baseados em aprendizado de máquina usam algoritmos para aprender a partir de dados e tomar decisões com base nessa aprendizagem. O aprendizado de máquina inclui várias técnicas, como redes neurais, árvores de decisão, *k-means* e outros métodos. Nas últimas décadas

muito do foco colocado nesse campo de estudo tem sido concentrado especificamente no estudo de aprendizado de máquina (AM) [54-56].

Aprendizado de máquina é uma subárea da inteligência artificial que se concentra no desenvolvimento de algoritmos e modelos que permitem que um sistema aprenda a partir de dados, sem que haja programação explícita para executar uma tarefa específica. Os modelos de aprendizado de máquina são alimentados com dados, e o algoritmo é capaz de identificar padrões nesses dados e usar esses padrões para fazer previsões ou tomar decisões [57].

Existem várias técnicas de aprendizado de máquina, incluindo aprendizado supervisionado, aprendizado não supervisionado e aprendizado por reforço. O aprendizado supervisionado envolve o uso de um conjunto de dados de treinamento que inclui entradas e saídas desejadas para o modelo aprender a fazer previsões. O aprendizado não supervisionado envolve o uso de um conjunto de dados sem saídas desejadas, e o modelo é capaz de identificar padrões nos dados por conta própria. O aprendizado por reforço envolve o treinamento de um modelo por meio de recompensas e punições, permitindo que o modelo aprenda a tomar ações para maximizar a recompensa [58-63].

Os estudos na área de AM se difundiram, principalmente, devido às pesquisas relacionadas ao reconhecimento de padrões e da importância da automação destes processos, através da concepção de máquinas computadorizadas que detivessem a capacidade de aprender novas atividades, sem a necessidade de serem pré-programadas, com base em dados retroativos [52]. Desta maneira, a técnica de AM pode ser caracterizada como a habilidade de proporcionar conhecimento a computadores utilizando-se de dados de eventos passados. Por meio desta técnica é possível realizar a identificação de padrões e características que dificilmente seriam perceptíveis de maneira clara para um humano [58,60].

Para problemas de classificação, quando é necessário determinar para qual conjunto de saídas esperadas um determinado conjunto de dados aponta, um classificador pode ser treinado com uma parcela do conjunto total de entradas. Usualmente, utiliza-se de 70 a 80 % dos dados para treinamento do modelo, e o restante para testá-lo. Para avaliar se o modelo é suficientemente otimizado para

classificar, lança-se mão de um conjunto de métricas, tais como acurácia, precisão e sensibilidade (*recall* ou revocação), especificidade e F-score [64,65].

Outra forma de abordar as métricas de avaliação dos modelos é por meio de suas definições. A Tabela 1 detalha as principais métricas de avaliação de modelos de aprendizado de máquina, em que VP = valores positivos, VN = valores negativos, FP = falsos positivos, FN = falsos negativos, N = número total de elementos, P = precisão e S = sensibilidade. A acurácia mede o grau de acerto do modelo, a quantidade de previsões corretas dentro do universo de elementos avaliados. Sensibilidade avalia a capacidade de o modelo prever corretamente resultados classificados como positivos, enquanto a especificidade atua da mesma forma sobre as previsões classificadas como negativas. Precisão mede o quão relevantes são os resultados obtidos a partir da quantidade de verdadeiros positivos previstos sobre o total de valores positivos. F-score é a média harmônica da precisão e da sensibilidade. Quanto mais o resultado da métrica F-score se aproxima de 1, melhor é o desempenho do modelo na previsão correta de valores positivos e negativos [66, 72,73].

Tabela 1 - Definição de métricas utilizadas em modelos de classificação. Adaptado de [67].

Métrica	Fórmula
Sensibilidade	$S = \frac{VP}{VP + FN}$
Especificidade	$E = \frac{VN}{(FP + VN)}$
Acurácia	$A = \frac{(VP + VN)}{N}$
Precisão	$P = \frac{VP}{(VP + FP)}$
F-score (F ₁)	$F_1 = 2 \times \frac{P \times S}{(P + S)}$

Essas métricas serão importantes para avaliação dos resultados obtidos nesse trabalho, pois como os resultados de frequência dos BSME podem variar de acordo com o ambiente de teste de cada amostra, é necessário que seja possível diferenciar amostras positivas de amostras negativas. Assim, métricas como

precisão e sensibilidade, podem ser utilizadas para avaliar essa diferenciação, seja pelo valor de deslocamento de frequência ou pelo padrão da curva.

3.4. EXTREME GRADIENT BOOSTING

O algoritmo *Extreme Gradient Boosting (XGBoost)*, ou Impulso Gradiente Extremo, é um algoritmo muito usado em problemas de classificação que envolvem aprendizado supervisionado e é capaz de trabalhar com grandes conjuntos de dados de forma mais eficiente do que modelos mais simples. É classificado como um método de *ensemble*, em virtude de combinar modelos fracos para criar um modelo forte de previsão. O *XGBoost* utiliza um método de *boosting*, ou impulso, em que múltiplos algoritmos mais fracos, como árvores de decisão simples, são treinados em sequência. Cada modelo é treinado para corrigir os erros de predição do modelo anterior, por meio do ajuste das previsões com base no gradiente da função de perda, que é a derivada dessa função [67].

A função perda mede a diferença entre valores previstos pelos modelos e os valores reais. Para minimizá-la, o algoritmo adiciona novas árvores de decisão para prever o erro em relação às previsões anteriores. A função de perda mais comumente utilizada é o erro quadrático médio, demonstrado na Equação (4), em que y_i é o valor real (ou observado) e \hat{y}_i é o valor predito pelo modelo [67,68].

$$L(\theta) = \sum_i (y_i - \hat{y}_i)^2 \quad (4)$$

Para esse modelo, a função objetivo é composta por dois termos, a própria função de perda e um componente chamado de regularização, que penaliza a complexidade do modelo para reduzir o sobre ajuste (*overfitting*). Desse modo, o algoritmo minimiza a função objetivo em cada iteração para encontrar a melhor árvore de decisão, com base nos parâmetros de entrada, calculando os pesos ótimos para cada folha da árvore a partir de técnicas de descida gradiente (*gradient descent*). A função objetivo é apresentada nas Equações (5) e (6), em que T é o número de folhas da árvore, w_j é o peso da j -ésima folha e γ e λ são hiper parâmetros de controle [67,68].

$$O(\theta) = L(\theta) + \Omega(\theta) \quad (5)$$

$$O(\theta) = \sum_i (y_i - \hat{y}_i)^2 + \left(\gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \right) \quad (6)$$

3.5. MÉTODO *BOOTSTRAP*

Bootstrap é um método estatístico de reamostragem usado para estimar estatísticas de uma população a partir de uma amostra de dados. O método *bootstrap* funciona por meio da reamostragem aleatória com reposição dos dados da amostra para criar uma quantidade grande número de amostras replicadas e, em seguida, calcular a estatística de interesse em cada uma dessas amostras. Esse método pode ser aplicado quando a amostra é pequena ou a distribuição dos dados é desconhecida [69]. Além disso, é usado para estimar intervalos de confiança, testes de hipóteses e outras estatísticas, sem fazer suposições sobre a distribuição subjacente dos dados. O método *bootstrap* é baseado em princípios simples, mas pode ser aplicado a uma ampla variedade de problemas estatísticos e tem sido amplamente utilizado em muitas áreas, incluindo finanças, economia, biologia, engenharia e ciência social [70].

4. MATERIAIS E MÉTODOS

4.1. MATERIAIS

4.1.1. Biossensores

Os sensores¹ foram obtidos a partir de uma fita magnetoelástica comercial, amorfa, que é comercializada com o nome de Metglas 2826MB3 da empresa *Metglas Company*. Para a fabricação dos sensores, as fitas foram cortadas no tamanho de 5 mm x 1 mm x 28 μm , uma vez que estas proporções estão associadas aos modos de vibração preferíveis para obtenção de medidas precisas da frequência de ressonância fundamental [16].

Posteriormente, os sensores foram recobertos por camadas de cromo e ouro (40 e 100 nm, respectivamente), por meio da técnica de *magnetron sputtering*, em ambos os lados dos sensores, como os que podem ser vistos na Figura 3.

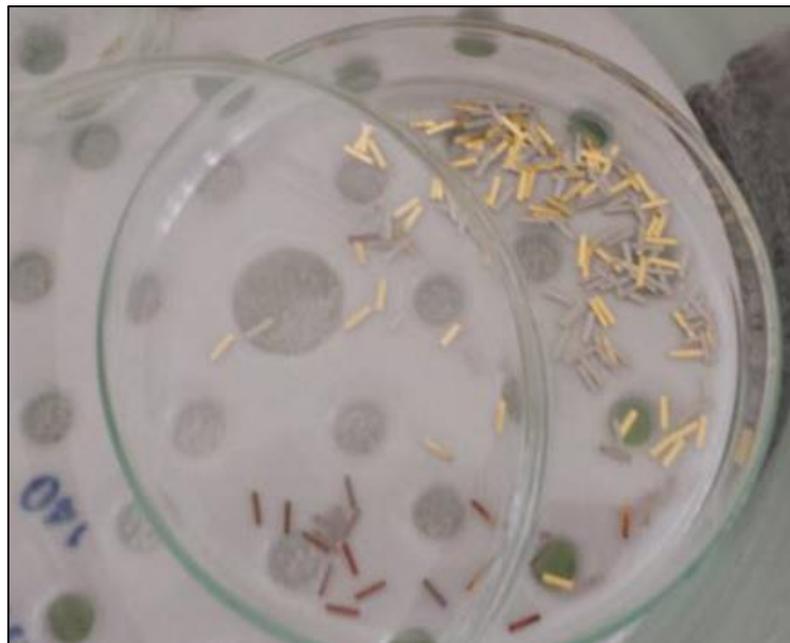


Figura 3 - Sensores magnetoelásticos similares aos usados neste trabalho para diagnóstico de SARS-CoV-2. As dimensões são 5 mm x 1 mm. Adaptado da Ref. [47].

¹ Este trabalho avaliou os testes realizados com sensores magnetoelásticos preparados no Instituto de Biotecnologia da UCS, pela aluna de doutorado Luíza Felippi de Lima.

O sistema de medição dos SME é composto por um conjunto de bobinas eletromagnéticas (AC e DC) conectadas a um analisador de rede (Keysight E5061B) (Figura 4) e a uma fonte DC, a partir do qual, o sinal de ressonância é coletado a cada 5 minutos, por 30 minutos (Figura 5).



Figura 4 - Analizador de rede Keysight E5061B.

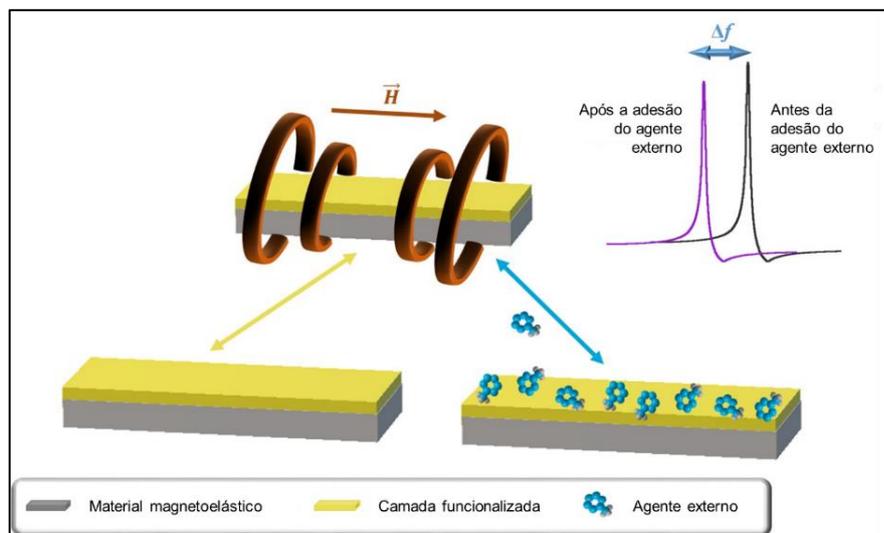


Figura 5 - Modelo esquemático do sistema de detecção utilizando sensores magnetoelásticos e analisador de rede. Adaptado da Ref. [74].

4.1.2. Ambiente

Um ambiente foi montado utilizando *Jupyter Notebook* como plataforma de desenvolvimento, em um computador com Windows 10, rodando Python na versão 3.9.7. O código fonte do *software*, juntamente com todos os recursos necessários para sua utilização, foi disponibilizado num repositório público disponível em <https://github.com/andreatta-ale/external-agents-detector>.

4.2. MÉTODOS

Um conjunto de dados foi montado contendo dois grupos, teste e controle, contendo os dados de três sensores cada, totalizando 6 sensores (Figura 6). Para cada sensor, o processo de medição foi executado pelo analisador de rede em ciclos com duração de 30 minutos, em que a coleta dos dados de frequência e sinal foram obtidos a cada 5 minutos, incluindo $t = 0$, gerando um total de 7 subconjuntos de dados para cada sensor, com 201 leituras de sinal e frequência, para cada grupo.

A faixa de frequência utilizada no ensaio foi entre 430000 Hz e 460000 Hz (201 leituras entre 0,43 e 0,46 MHz), intervalo no qual está a frequência de ressonância característica do material utilizado nos sensores. O conjunto de dados final reuniu os dados de frequência e sinal de todos os ciclos dos sensores de cada grupo, totalizando 8442 linhas.

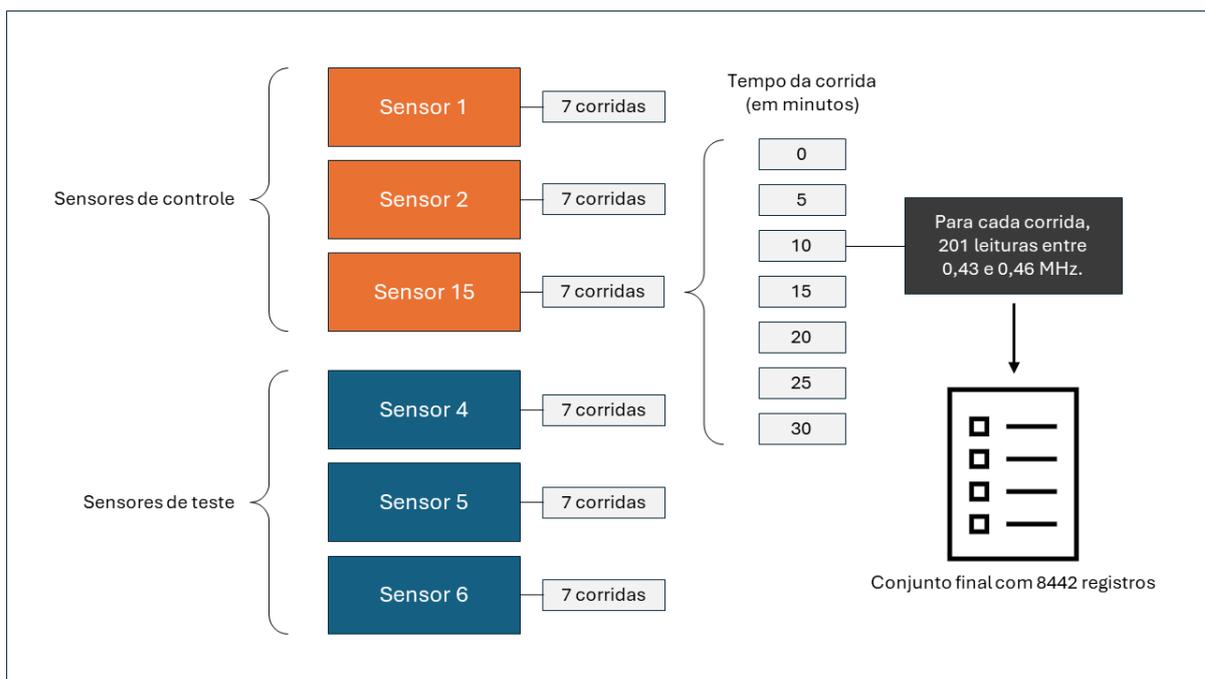


Figura 6 - Modelo esquemático do processo de aquisição de dados.

Uma coluna de *hash* foi criada para atuar como campo de referência para análise, que consiste na união entre grupo (teste ou controle), o número do sensor e o tempo em que a obtenção do sinal foi executada (corrida), conforme a Figura 7.

O *software* desenvolvido lê os dados de frequência e sinal, captura os valores em torno do pico estimado e ajusta uma curva lorentziana assimétrica. O ponto

mínimo dessa curva é o ponto de interesse da análise, porque a partir dele se mede o desvio da frequência usado para identificar a alteração de massa do sensor, denunciando a presença de novos materiais, como vírus, por exemplo.

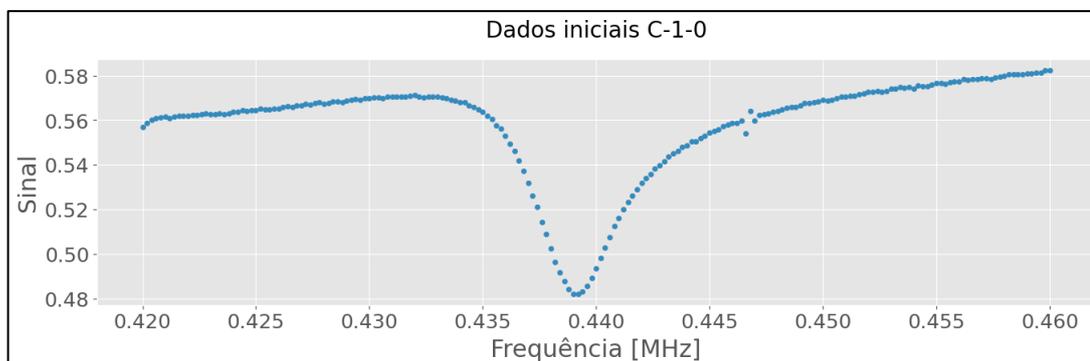


Figura 7 - Dados iniciais do sinal versus frequência, identificados pelo hash.

	Frequência	Sinal	Grupo	Sensor	Tempo	Hash
0	0.4200	0.557114	C	1	0	C-1-0
1	0.4202	0.558667	C	1	0	C-1-0
2	0.4204	0.560101	C	1	0	C-1-0
3	0.4206	0.560883	C	1	0	C-1-0
4	0.4208	0.561223	C	1	0	C-1-0

Figura 8 - Representação do conjunto de dados utilizado para análise, mostrando a coluna de hash usada para correlacionar diferentes conjuntos de dados.

A Figura 8 mostra os dados iniciais importados para o ambiente de análise; a Tabela 2 apresenta a distribuição dos dados no conjunto de dados analisado.

Tabela 2 - Constituição do conjunto de dados analisado.

GRUPO	QTD SENSORES	QTD CICLOS POR SENSOR	QTD SUBCONJUNTOS
Teste	3	7	21
Controle	3	7	21
Totais	6	7	42

Para que fosse possível avaliar a estatística do ponto de interesse, uma SMC foi executada sobre os parâmetros de saída do ajuste inicial, utilizando a definição da função da curva lorentziana assimétrica (Equações (7) e (8)) como função de entrada, levando em conta a parcela linear da linha de base $g(x) = mx + b$, com um total de 1000 iterações, para cada sensor de ambos os grupos, em cada intervalo de tempo.

$$f(x; A, \mu, \sigma, \sigma_r) = \frac{2A}{\pi(\sigma + \sigma_r)} \beta + g(x) \quad (7)$$

com

$$\beta = \frac{\sigma^2}{(x - \mu)^2 + \sigma^2} H(\mu - x) + \frac{\sigma_r^2}{(x - \mu)^2 + \sigma_r^2} H(\mu - x) \quad (8)$$

onde A = amplitude da curva, σ e σ_r equivalem às larguras assimétricas em relação ao centro da curva μ , e $H(x)$ corresponde à função degrau de *Heaviside* (Equação (9)).

$$H(x) = \begin{cases} 0, & x < 0 \\ \frac{1}{2}, & x = 0 \\ 1, & x > 0 \end{cases} \quad (9)$$

O conjunto de dados resultante desse processo contém os valores da frequência minimizada e suas respectivas incertezas, a partir do qual é possível plotar o gráfico da frequência média em cada instante (Figura 9).

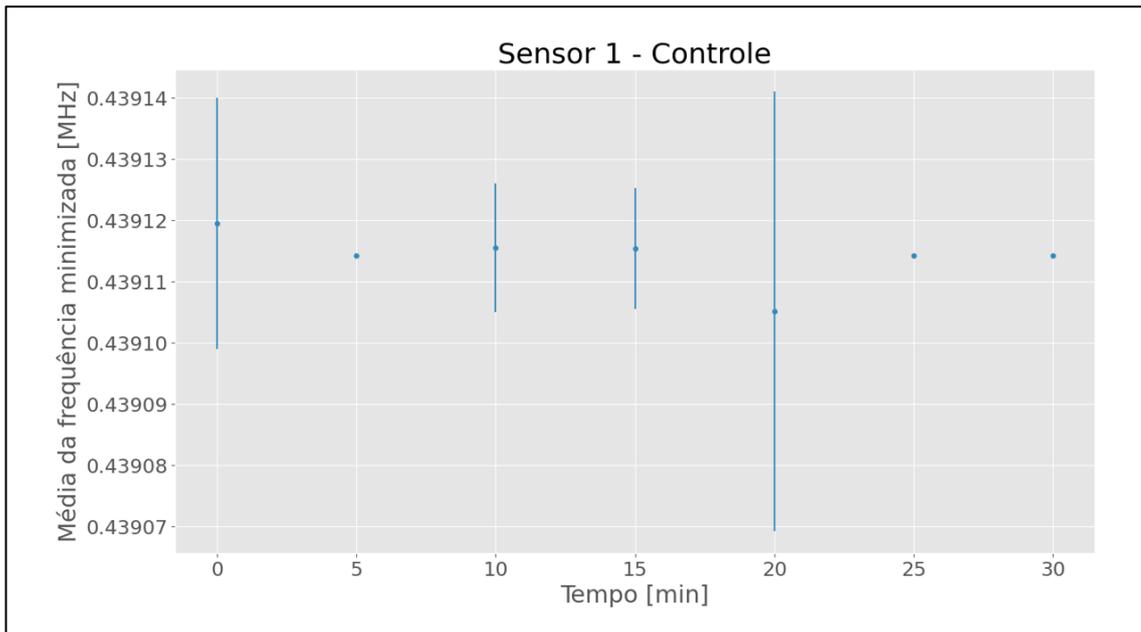


Figura 9 - Variação da frequência no tempo.

Foi utilizada a equação de Lagergren [71], demonstrada a seguir, que descreve a cinética de *pseudo* primeira ordem de adsorção, para ajustar os dados de frequência pelo tempo, viabilizando uma comparação mais precisa entre os deslocamentos das frequências para os sensores de teste e controle.

$$\frac{dm(t)}{dt} = k[m_{sat} - m(t)] \quad (10)$$

Integrando para $m(t) = m(0)$ e $m(t)$ no tempo t :

$$\int_{m(0)}^{m(t)} \frac{dm}{m_{sat} - m(t)} = k \int_0^t dt \quad (11)$$

$$-\ln \left[\frac{m_{sat} - m(t)}{m_{sat} - m(0)} \right] = kt \quad (12)$$

$$m(t) = m_{sat}(1 - ae^{-kt}) \quad (13)$$

Para a frequência do sensor no tempo:

$$\Delta f = f(t) - f(0) = -\frac{f(0)\Delta m}{2m_s} = -f(0) \frac{[m(t) - m(0)]}{2m_s} \quad (14)$$

$$f(t) = f(0) \left[1 - \frac{m(t) - m(0)}{2m_s} \right] \quad (15)$$

$$f(t) = f(0) \left\{ 1 - \left[\frac{m_{sat}(1 - ae^{-kt}) - m(0)}{2m_s} \right] \right\} \quad (16)$$

$$f(t) = f(0) \left[1 - \frac{m_{sat}}{2m_s} (1 - ae^{-kt}) \right] \quad (17)$$

O conjunto de resultados do ajuste das equações de Lagergren aos dados, tempo e frequência, contém as informações sobre a variação da frequência – Δf – em cada diferença entre o tempo da corrida e o tempo inicial $t_n - t_0$. Os dados referentes à $t_n = 30$ foram, então, submetidos ao método de *bootstrap* para obtenção de dados aleatoriamente escolhidos como entrada para o classificador.

Para tanto, uma lista contendo os dados de cada grupo – controle e teste – foi adicionada a um laço de 2000 iterações, e em cada iteração uma função de *bootstrapping* escolheu, aleatoriamente, um ponto do conjunto contendo o valor da variação da frequência e suas incertezas, além do rótulo (C ou T). A lista final, com 4000 registros, foi aleatoriamente misturada, para contornar vieses, e dividida em conjuntos de treino e teste, considerando 70 % dos dados para o primeiro e os 30 % restantes dos dados para o segundo.

Para o treinamento do algoritmo de classificação, foi implementada uma função que utiliza o algoritmo *Extreme Gradient Boosting*, ou Impulso Gradiente Extremo (*XGBoost*) como classificador, a partir da biblioteca *Scikit-learn*.

Também foi utilizada uma validação cruzada de cinco repetições para cada classificador. As métricas relacionadas à avaliação do conjunto de dados por cada classificador foram compiladas e serão demonstradas nos resultados. O fluxograma do processo completo de análise é apresentado na Figura 10.

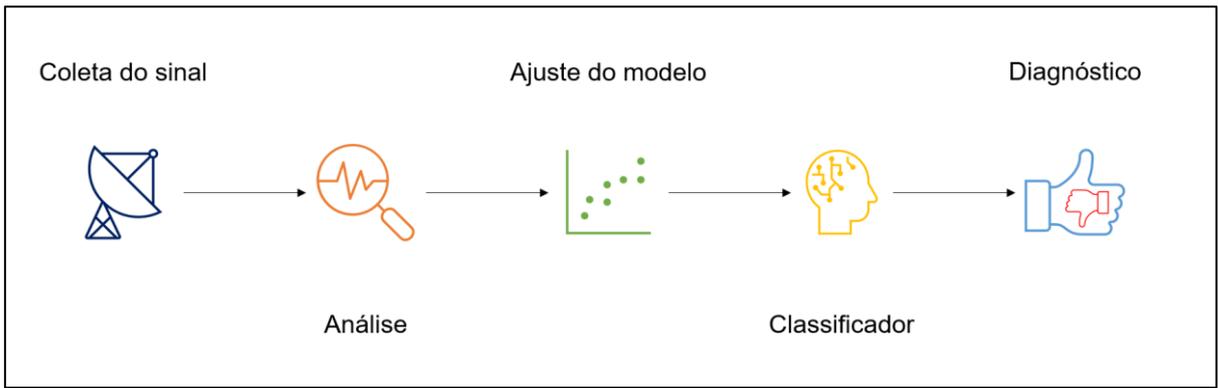


Figura 10 - Fluxograma do funcionamento do programa desenvolvido em Python.

5. RESULTADOS E DISCUSSÃO

5.1. AJUSTES INICIAIS

Para o conjunto de dados de cada um dos sensores foram selecionados 26 pontos adjacentes ao menor valor do sinal dos dados originais (13 de cada lado), que corresponde ao pico estimado, como demonstrado na Figura 11.

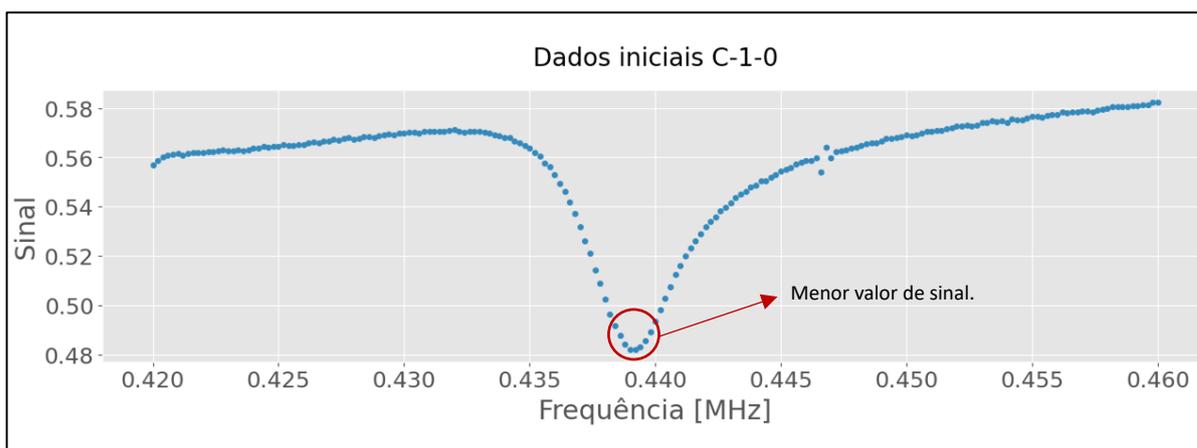


Figura 11 - Posição do ponto mínimo de frequência e sinal.

Os dados em torno do pico foram ajustados à curva do modelo lorentziano assimétrico utilizando o pacote LMFIT, em cuja saída se encontra a lista de parâmetros ótimos encontrados pelo método de ajuste do pacote (por padrão, método dos mínimos quadrados) e suas respectivas incertezas. A simulação de Monte Carlo foi implementada em um laço de 1000 iterações, utilizando os valores e as incertezas dos parâmetros A , σ , σ_r e μ , que correspondem aos parâmetros da função lorentziana assimétrica e intercepto e coeficiente angular (m e b), parcela linear da modelagem da função completa. O resultado da simulação convergiu, então, para um valor mínimo, que contempla o valor otimizado da frequência e seu desvio padrão, para cada sensor e em cada minuto em que a coleta de dados foi feita.

A Figura 12 mostra o resultado do ajuste nos pontos em torno do valor de sinal mínimo (μ dos dados originais), representado pelo ponto em amarelo. O ponto em verde se refere ao ponto de mínimo estimado pelo otimizador do LMFIT. Em preto, a média da frequência da SMC equivale ao ponto estimado a partir da média

dos valores da simulação de Monte Carlo (1000 iterações). A curva em vermelho representa os pontos ajustados aos dados pelo modelo do LMFIT.

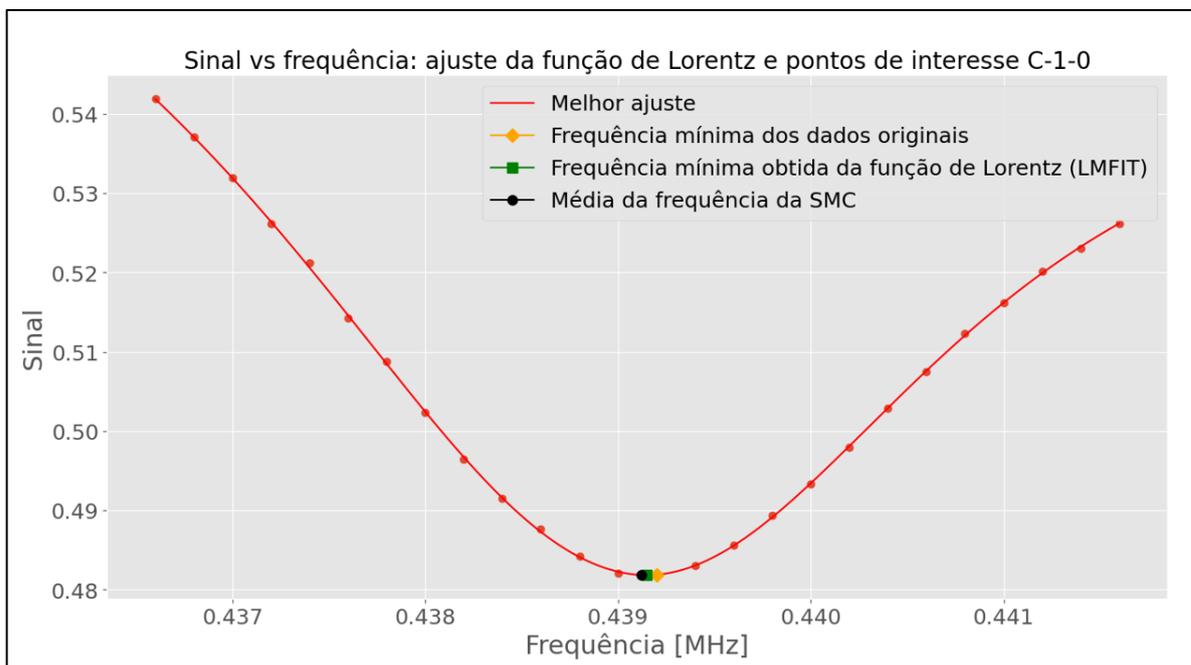


Figura 12 - Comparação entre dados reais e ajustados com LMFIT e SMC.

A Figura 13 apresenta o índice de acurácia entre a frequência média obtida do primeiro ajuste, com o pacote LMFIT, e a média obtida a partir da simulação de Monte Carlo e minimização da frequência. Esse índice indica o quão perto as médias obtidas pelas diferentes abordagens estão. Um índice muito acima ou muito abaixo de 99 % poderia sugerir a inadequação do uso do modelo inicial – LMFIT – como fonte dos parâmetros para a simulação de Monte Carlo.

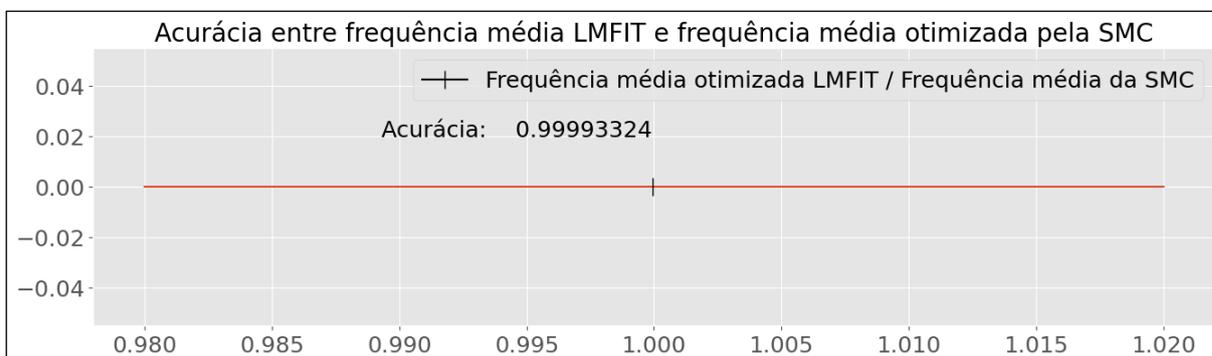


Figura 13 - Índice de acurácia entre LMFIT e a simulação de Monte Carlo.

Como resultado desse processo, foi possível obter o conjunto de dados correspondente às médias das frequências minimizadas ao longo do tempo de coleta de sinal de todos os sensores, a partir do qual se pode modelar uma função de deslocamento de frequência. A Figura 14 mostra as primeiras linhas desse conjunto.

hash	Média da frequência minimizada em MHz	Desvio padrão	Erro padrão	Frequência minimizada com incertezas em Hz
C-1-0	0.439119	2.050176e-05	6.483227e-07	439119+/-21
C-1-5	0.439114	5.551115e-17	1.755417e-18	439114.21875000006+/-0.00000000006
C-1-10	0.439116	1.050217e-05	3.321079e-07	439116+/-11
C-1-15	0.439115	9.845387e-06	3.113385e-07	439115+/-10
C-1-20	0.439105	3.589229e-05	1.135014e-06	(4.3911+/-0.0004)e+05
C-1-25	0.439114	1.110223e-16	3.510833e-18	439114.21875000012+/-0.00000000011
C-1-30	0.439114	5.551115e-17	1.755417e-18	439114.21875000006+/-0.00000000006
C-15-0	0.439489	1.561351e-05	4.937426e-07	439489+/-16
C-15-5	0.439484	1.293440e-05	4.090216e-07	439484+/-13
C-15-10	0.439478	2.421575e-05	7.657692e-07	439478+/-24

Figura 14 - Amostra do conjunto de dados de médias minimizadas.

5.2. ANÁLISE DOS DESLOCAMENTOS DE FREQUÊNCIAS

Os dados das frequências minimizadas podem ser observados na Figura 15, que apresenta o deslocamento das frequências para os conjuntos de controle e de teste ao longo do tempo.

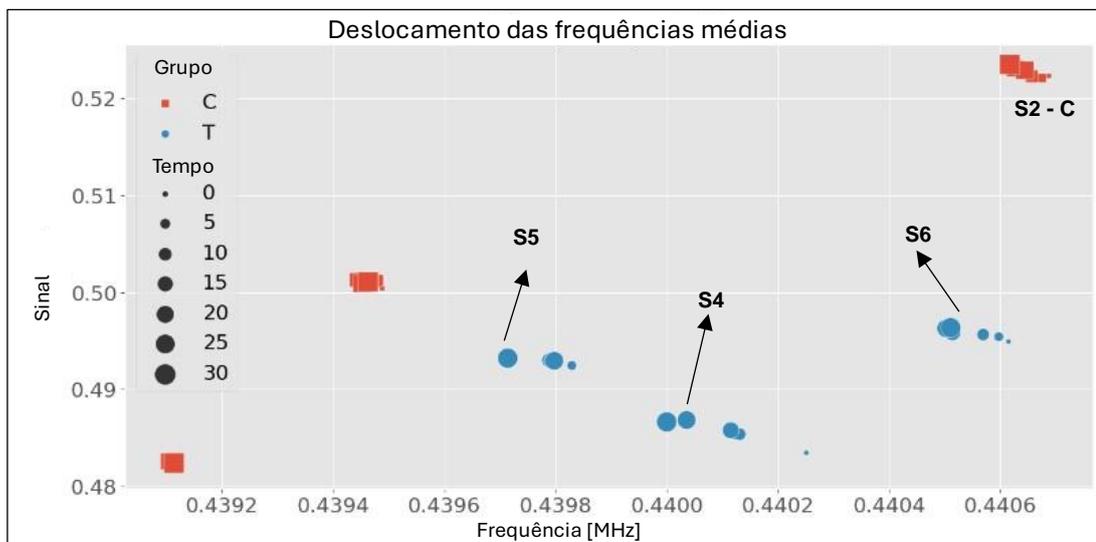


Figura 15 - Deslocamento das frequências. O tempo do ensaio é demonstrado pelo tamanho do marcador. S4, S5 e S6 se referem aos sensores 4, 5 e 6, do grupo de teste. Sensor 2 no detalhe mostrando alguma variação de frequência.

Nessa visualização é possível observar que à medida que o tempo avança, a frequência diminui. Esse comportamento é esperado quando há aumento de massa no sensor, já que a frequência característica do sensor diminui de forma inversamente proporcional ao aumento de sua massa, como mostrado na Equação (3).

Colocando-se os dados das frequências minimizadas pelo tempo em gráfico, obtém-se uma visão do comportamento dos sensores ao longo do ensaio. Também é esperado que a variação das frequências seja pequena para os sensores de controle. A Figura 16 apresenta as frequências pelo tempo em um sensor de controle e em um sensor de teste.

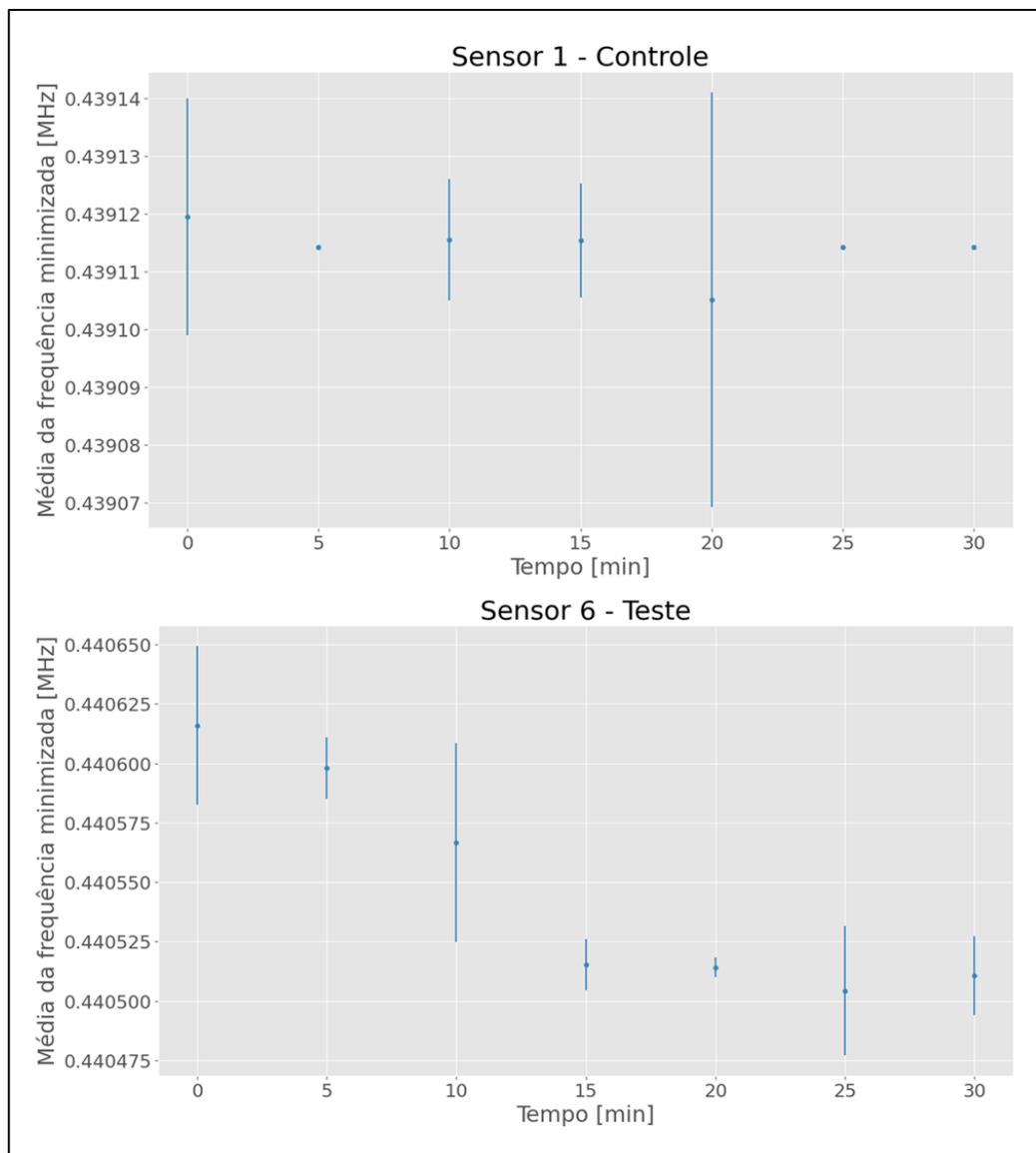


Figura 16 - Comparação entre a variação de frequência entre dois sensores (controle e teste).

Esses dados foram ajustados à equação de Lagergren, o que permitiu obter $\Delta f = f(30) - f(0)$, que corresponde à variação de frequência entre o início do ensaio e o final da coleta de dados, em $t = 30$ min, como apresentado na Figura 17, considerando os mesmos sensores da Figura 16.

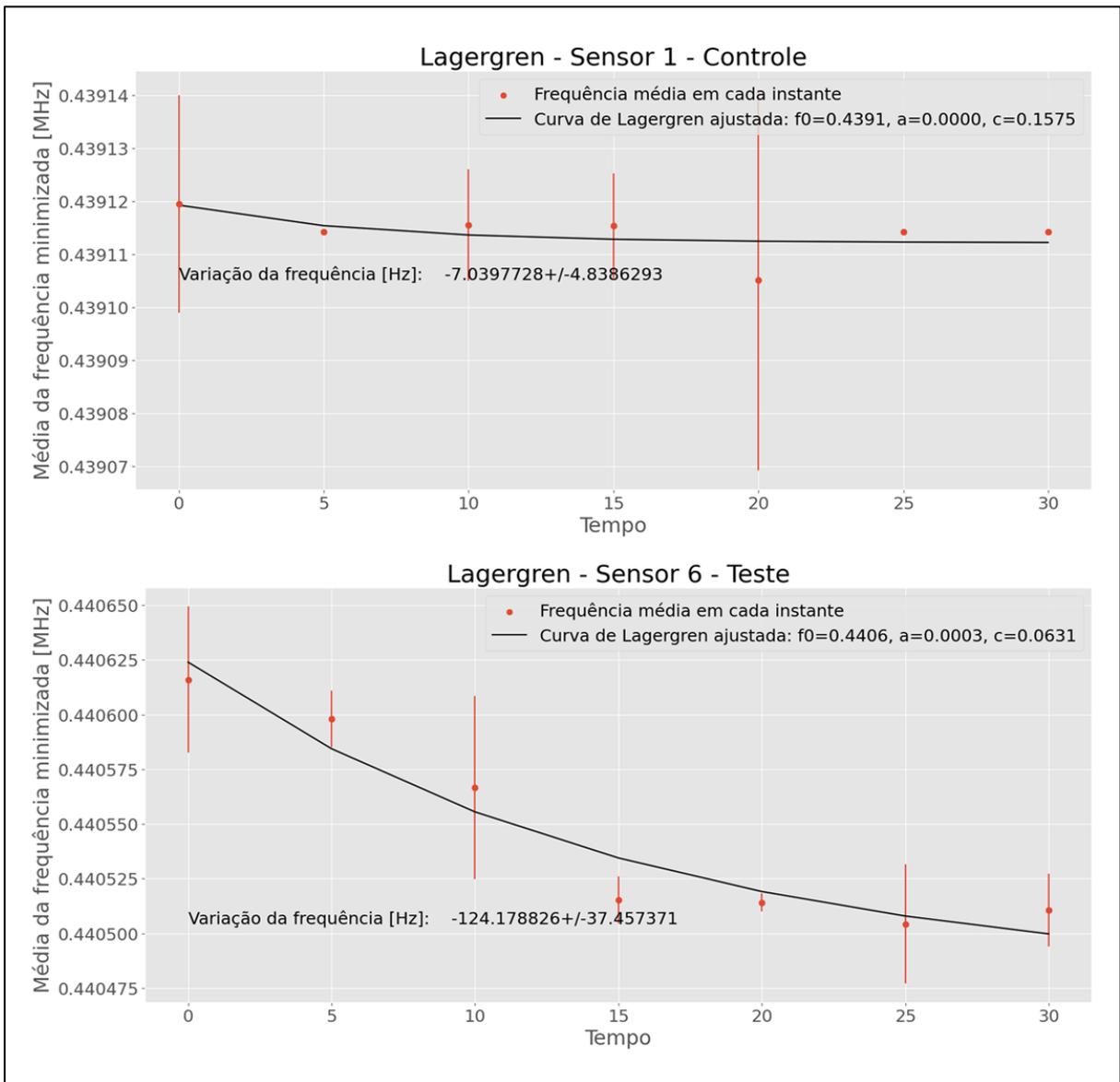


Figura 17 - Ajuste dos dados de frequência pelo tempo utilizando a equação de Lagergren.

Os resultados dos ajustes são exibidos na Tabela 3.

Tabela 3 - Deslocamentos de frequência total para os sensores de teste e controle.

GRUPO	SENSOR	DESLOCAMENTO DE FREQUÊNCIA [Hz]	ERRO PADRÃO [Hz]
Controle	1	7,2814	4,9159
	2	65,6469	34,7618
	15	35,6223	2,2862
Teste	4	202,1629	793,7374
	5	87,2731	15402780,0000
	6	123,9632	42,5311
	5*	122,2627	24,1407

Nota-se que o erro padrão para a frequência do sensor de teste número 5 é consideravelmente maior que os erros dos demais sensores. Ao se verificar o comportamento das frequências pelo tempo para esse sensor, percebe-se que a variação de frequência em $t = 30$ min foi tão expressiva que o ajuste por meio da equação de Lagergren pode tê-lo considerado como um *outlier* (Figuras 18 e 19).

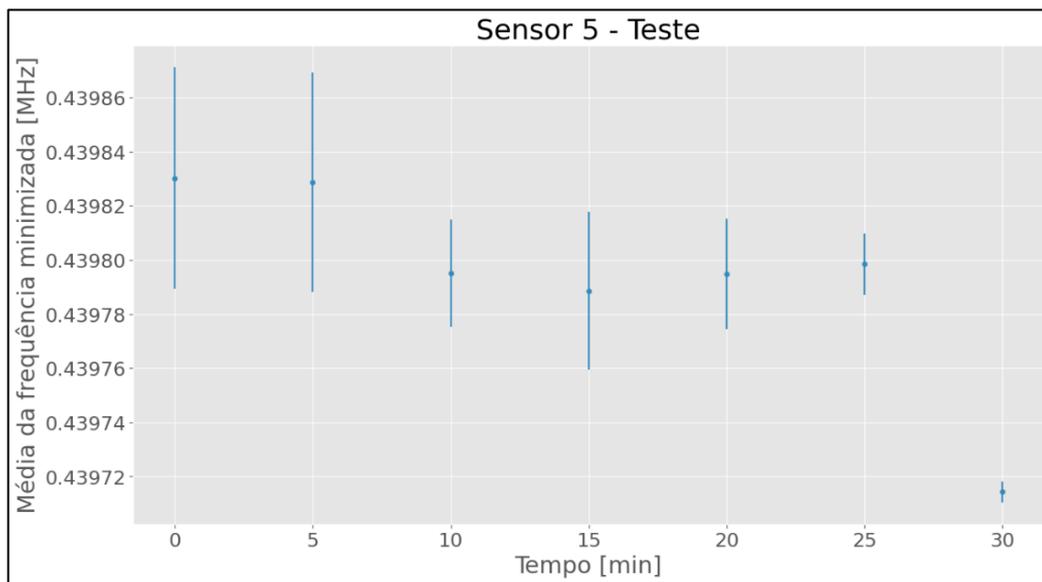


Figura 18 - Deslocamentos de frequência pelo tempo para o sensor 5 (teste).

Apenas para efeito de comparação, foi adicionada uma linha à Tabela 3 contendo uma estimativa fora do ajuste de Lagergren obtida subtraindo a frequência em $t = 30$ min de $f(0)$, e referenciada como 5*.

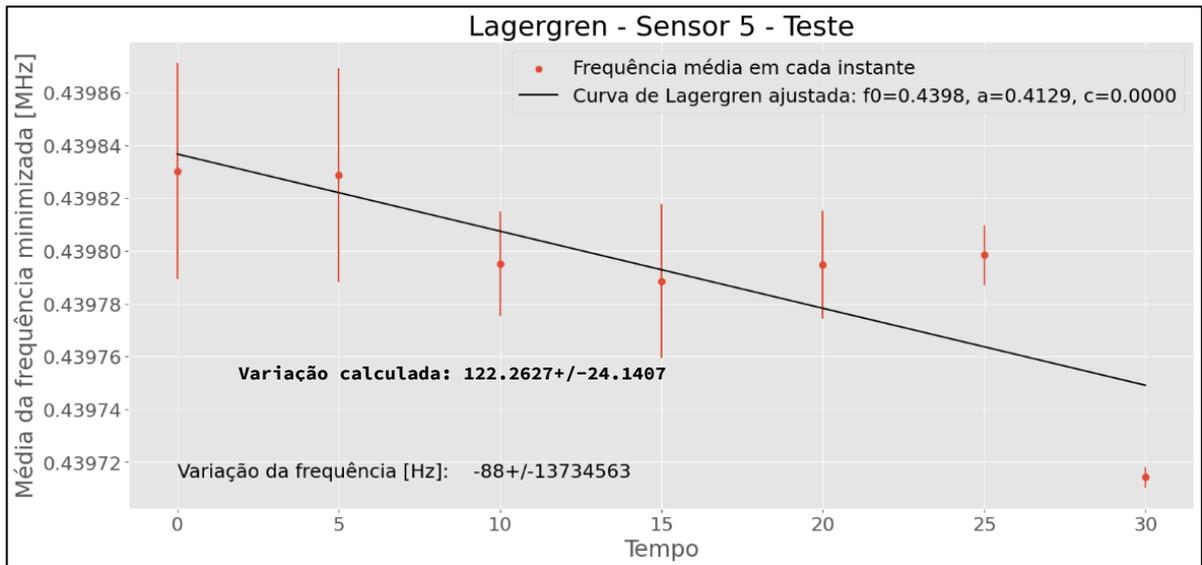


Figura 19 - Ajuste das frequências à equação de Lagergren para o sensor 5. O comentário adicional faz referência a $f(t=0) - f(t=30)$, como meio de se observar uma aproximação fora do ajuste.

As curvas de deslocamentos foram agrupadas para observar a distinção entre os grupos. No segundo quadro da Figura 20, a distinção fica evidente a partir da corrida 15 ($t = 15$ min).

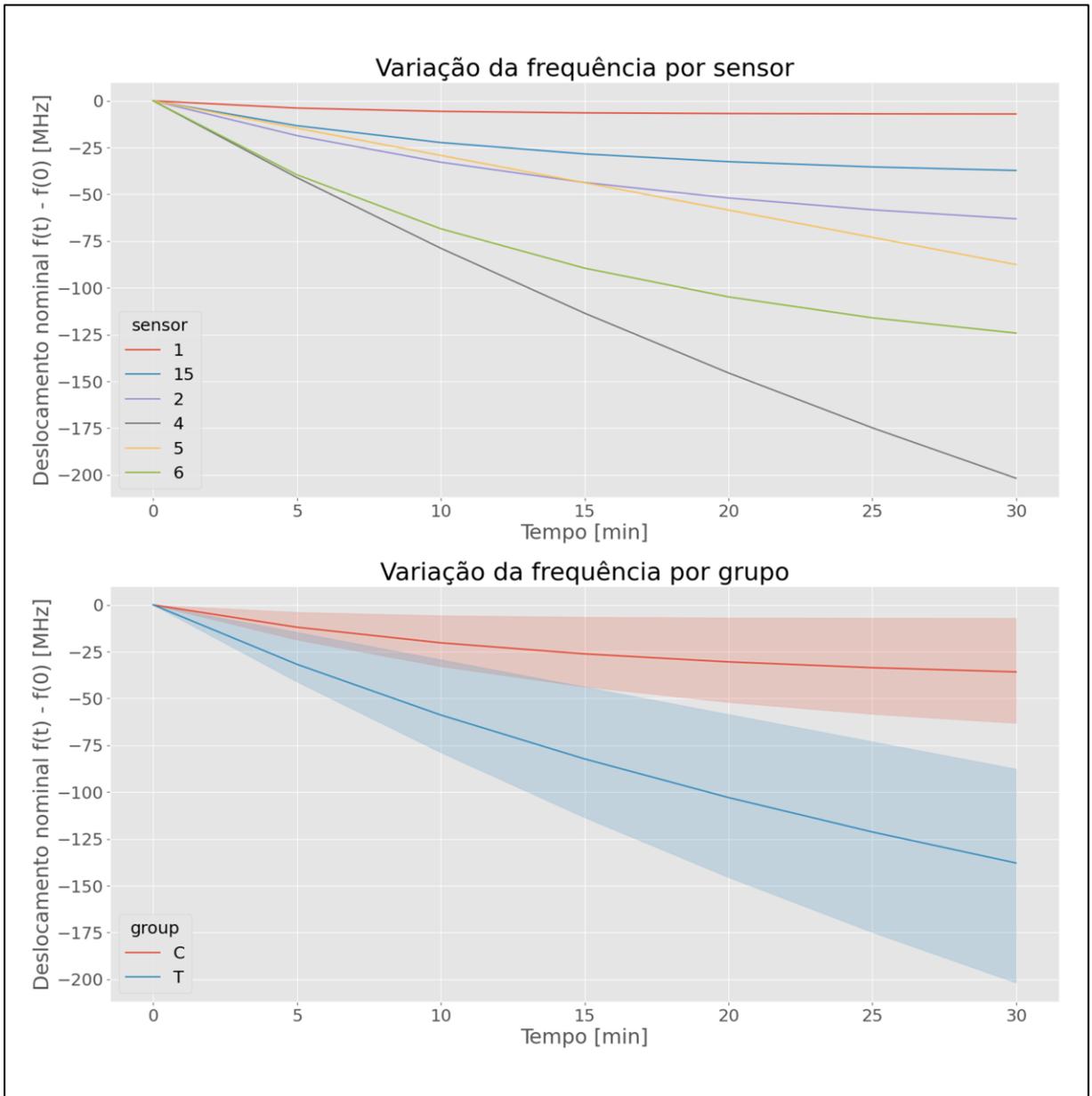


Figura 20 - Agrupamento das curvas de deslocamento por grupo.

O gráfico de caixa da Figura 21 apresenta a zona indistinta antes dos 15 minutos como *outliers* dos dados de teste. O limite superior dos primeiros 25 % dos dados de controle aparece muito próximo ao limite inferior dos últimos 25 % dos dados de teste (linha amarela). Contudo, a distribuição compreendida entre 25 % e 75 % ainda comporta a maior parte dos dados distinguíveis.

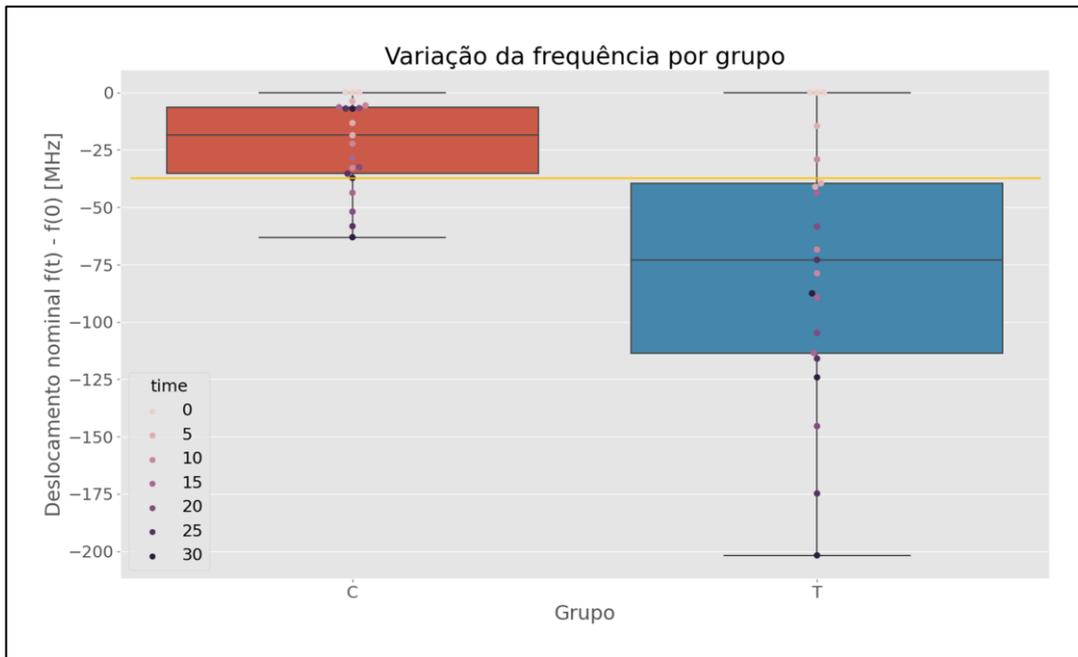


Figura 21 - Agrupamento das curvas de deslocamento por grupo em gráfico de caixa.

Na visualização apresentada na Figura 22 para os dados após o *bootstrapping*, nota-se a clara variação das frequências com o passar do tempo de ensaio.

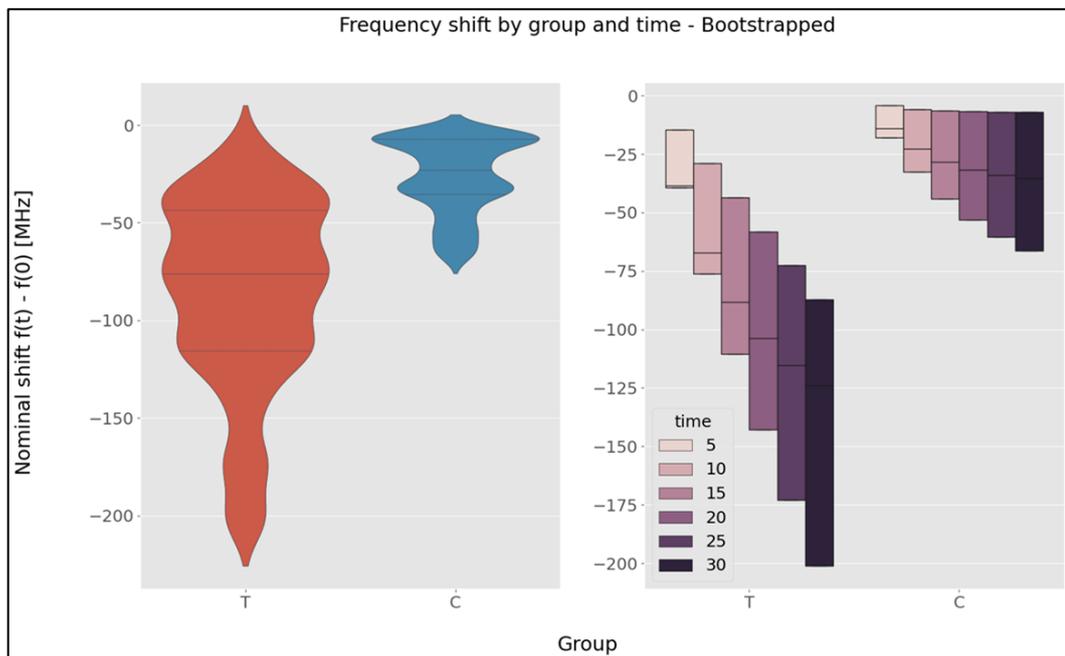


Figura 22 - Distribuição do deslocamento de frequências por grupo em gráfico de violino e caixa para os dados gerados pelo *Bootstrap*.

5.3. CLASSIFICADOR E ESTIMATIVA DE DIAGNÓSTICO

A função contendo o classificador foi executada sobre o conjunto de dados aleatoriamente escolhidos, depois da aplicação do método de *Bootstrap*, a fim de estimar os parâmetros mais adequados para classificar os dados de sensores quanto à presença ou não de vírus (classe C ou classe T). Para tanto, os dados do ajuste contendo a classe, deslocamento da frequência e desvio padrão, foram divididos em treino (70 %) e teste (30 %).

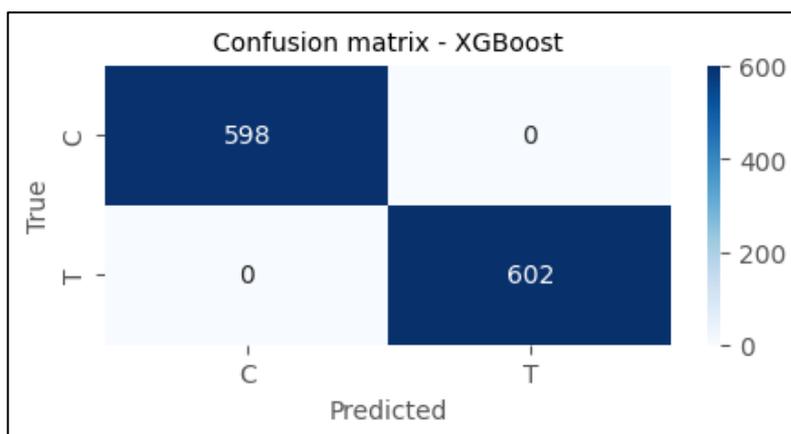


Figura 23 - Matriz de confusão: XGBoost.

O resumo das métricas relacionadas ao desempenho do modelo pode ser observado na Tabela 4.

Tabela 4 - Métricas de avaliação do classificador.

MÉTRICA	VALOR
Acurácia	100.00%
Precisão	100.00%
Sensibilidade	100.00%
F-score	100.00%

O alto nível de acerto do modelo é esperado, porquanto os dados, ainda que randomizados, refletem os dados originais, cuja quantidade é consideravelmente menor do que os utilizados para o treinamento do modelo.

O modelo treinado foi exportado como *XBoost.joblib*, que pode ser invocado dentro de aplicações, locais ou *web*, para classificar dados externos (Figura 23). Para efeito de classificação, o modelo distingue entre dados de teste (T) e controle

(C), que são as classes. Para uso do modelo, é necessário submeter os dados de deslocamento de frequência e desvio padrão para $t = 30$, de um sensor analisado, de modo que o resultado seja T ou C.

Com a aplicação do modelo em um pequeno conjunto de dados externos ao teste, contendo a variação de frequência após 30 minutos, foi possível verificar que o modelo classificou corretamente a presença de vírus em virtude do aumento de massa no sensor.

6. CONSIDERAÇÕES FINAIS E CONCLUSÕES

Com o ajuste dos dados do analisador de rede ao modelo Lorentziano assimétrico ao conjunto de dados de todos os sensores, em todos os tempos do ensaio, gerou-se um agrupamento dos parâmetros de cada iteração da modelagem, com o valor das incertezas. Esse foi o ponto de partida para a simulação de Monte Carlo, na qual foram aplicados os parâmetros e suas incertezas, produzindo um conjunto de dados otimizados para um valor mínimo de sinal, para uma determinada frequência, no intervalo de 13 pontos para cada lado do pico da curva dos originais (Figura 7). Esse valor de frequência foi usado como referência para comparação entre o ponto real de mínimo sinal (que equivale ao pico dos dados), o ponto otimizado pelo ajuste do pacote LMFIT sobre a curva ajustada e o ponto obtido da simulação de Monte Carlo.

O resultado desse processo foi uma tabela contendo o *hash* de identificação da linha dos dados, o grupo a que o dado pertence (controle ou teste), o identificador do sensor, o tempo da coleta de dados, o sinal (*y_eval*), média da frequência minimizada, desvio padrão dessa média, erro padrão e a representação em Hz das frequências e suas incertezas. A partir desse conjunto de dados otimizados, foi modelada a função de deslocamento de frequências.

Para a análise dos deslocamentos de frequência, inicialmente se avaliou a ocorrência de deslocamentos em um gráfico de dispersão de sinal por frequência, avaliando o tempo como função do tamanho do marcador (Figura 14). Nota-se deslocamento nas frequências de um sensor de controle, sensor 2, no qual não se espera aumento de massa. Esse fenômeno pode indicar contaminação do ambiente da amostra, mas não representa uma variação da ordem de grandeza esperada para a presença de vírus em 30 min (deslocamentos de 200 Hz ou mais). Também, observou-se o comportamento das frequências médias no tempo de ensaio com as incertezas, o que reforçou a hipótese de que há diferença no comportamento das frequências entre os dois grupos de sensores, perceptíveis a classificadores.

Em se constatando deslocamentos nos dados de frequências dos sensores de teste, o modelo da equação de Lagergren foi ajustado ao conjunto de dados referentes à frequência e tempo, para obtenção do valor de Δf , observados no gráfico dos dados com a curva ajustada aos dados e a indicação do deslocamento

(Figura 17). A distinção entre os dois grupos de sensores pode ser observada no gráfico das curvas do modelo de Lagergren ajustado aos dados, a partir de 15 min de ensaio (Figura 20). O método *Bootstrap* foi utilizado para se obter um conjunto de dados maior que o original, mas espelhando sua distribuição. É possível notar a separação significativa entre as médias dos dados de controle e de teste, já vistas no gráfico de caixa dos dados originais ajustados. Ainda é possível notar no gráfico de caixa dos dados do *Bootstrap* a presença da zona indistinta, ainda que as médias sejam visivelmente distintas.

Os dados obtidos pelo método de *Bootstrap* foram usados para treinar um modelo de classificação utilizando o algoritmo *XGBoost – Extreme Gradient Boosting*. Para avaliar o desempenho do modelo, gerou-se a matriz de confusão ao fim do treinamento, e as métricas foram compiladas em uma tabela única (Tabela 4). Considerando que a amostra original de dados era pequena, era esperado um grau considerável de sobre ajuste, caracterizado pelo alto nível de acurácia do modelo, além de sugerir que processos mais simples são capazes de determinar o grupo dos dados inseridos, a exemplo da classificação a partir das médias, tratado neste trabalho.

Considerando o cenário mundial cada vez mais globalizado, em que doenças virais têm o potencial de infecção na velocidade das pontes aéreas, os resultados encontrados demonstram que, não apenas é importante o uso de tecnologias de predição, a exemplo de modelos de aprendizado de máquina, como são necessários para a obtenção de diagnósticos rápidos. Com o avanço da computação em nuvem, e da democratização de aplicações como serviço que demandam pouco, ou nenhum, conhecimento de programação (*low coding*), a implementação dessas formas de detecção e diagnóstico se tornam ainda mais facilitadas. Contudo, por mais que a tecnologia esteja cada vez mais disponível, ainda é necessário preparar amostras para obter os dados que serão utilizados na modelagem. Assim, a ausência de um extenso conjunto de dados sob as mesmas condições foi, sem dúvidas, limitante para a obtenção de um modelo de inteligência mais preciso, treinado com ainda mais variações de informação. Entretanto, a falta desse conjunto de dados não impede de se obter um diagnóstico como abordado nesse trabalho, uma vez que a estratégia proposta, cujo caráter é generalista, se mostrou funcional

na diferenciação entre grupos de sensores em contextos em SME são utilizados para esse tipo de classificação.

Uma evolução da proposta desse estudo, além da implementação de outros modelos de classificação, a partir de um volume de dados maior, é extrapolar o universo dos vírus para além do analisado, SARS-CoV-2, construindo conjuntos de dados de análises de outros vírus, como alguns dos mencionados na introdução, e, para além disso, a implementação desses recursos na detecção de bactérias nocivas, tanto em contexto da medicina humana como veterinária. Além disso, ensaios podem ser conduzidos em sensores de diferentes formatos, para avaliar a abordagem computacional discutida em relação à diferença da geometria na sensibilidade dos SME, como sugerido por Aransay *et al.* [74].

REFERÊNCIAS

- [1] FRIEDEN, Thomas R. *et al.* Ebola 2014 - new challenges, new global response and responsibility. *New England Journal of Medicine*, v. 371, n. 13, p. 1177-1180, 2014.
- [2] CENTER FOR DISEASE CONTROL AND PREVENTION. History of Ebola Disease Outbreaks. Cases and Outbreaks of Ebola Disease by Year. Estados Unidos: CDC, 2023. Disponível em: https://www.cdc.gov/vhf/ebola/history/chronology.html?CDC_AA_refVal=https%3A%2F%2Fwww.cdc.gov%2Fvhf%2Febola%2Foutbreaks%2Fhistory%2Fchronology.html. Acesso em: 15 mar. 2023.
- [3] ROUGERON, V. *et al.* Ebola and Marburg haemorrhagic fever. *Journal of Clinical Virology*, v. 64, p. 111-119, 2015.
- [4] CENTER FOR DISEASE CONTROL AND PREVENTION. Marburg (Marburg Virus Disease). Estados Unidos: CDC, 2023. Disponível em: <https://www.cdc.gov/vhf/marburg/index.html>. Acesso em: 15 mar. 2023.
- [5] YADAV, Pragya D. *et al.* Nipah virus sequences from humans and bats during Nipah outbreak, Kerala, India, 2018. *Emerging Infectious Diseases*, v. 25, n. 5, p. 1003, 2019.
- [6] WORLD HEALTH ORGANIZATION. Nipah virus. [S.l.]. WHO, 2018. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/nipah-virus>. Acesso em: 15 mar. 2023.
- [7] LEAL, Mariana C. *et al.* Characteristics of dysphagia in infants with microcephaly caused by congenital Zika virus infection, Brazil, 2015. *Emerging Infectious Diseases*, v. 23, n. 8, p. 1253, 2017.
- [8] CENTER FOR DISEASE CONTROL AND PREVENTION. Zika Virus. Estados Unidos: CDC, 2023. Disponível em: <https://www.cdc.gov/zika/index.html>. Acesso em: 15 mar. 2023.
- [9] BENTE, Dennis A. *et al.* Crimean-Congo hemorrhagic fever: history, epidemiology, pathogenesis, clinical syndrome and genetic diversity. *Antiviral Research*, v. 100, n. 1, p. 159-189, 2013.
- [10] WORLD HEALTH ORGANIZATION. Crimean-Congo haemorrhagic fever. [S.l.]. WHO, 2022. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/crimean-congo-haemorrhagic-fever>. Acesso em: 15 mar. 2023.
- [11] CENTER FOR DISEASE CONTROL AND PREVENTION. COVID-19. Estados Unidos: CDC, 2023. Disponível em: <https://www.cdc.gov/coronavirus/2019-ncov/>. Acesso em: 15 mar. 2023.
- [12] WORLD HEALTH ORGANIZATION. WHO Coronavirus (COVID-19) Dashboard. [S.l.]. WHO, 2023. Disponível em: <https://covid19.who.int/>. Acesso em: 16 mai. 2023.
- [13] MINISTÉRIO DA SAÚDE (Brasil). Boletim Epidemiológico Vol. 52. 48. ed. [S. l.]: Ministério da Saúde, 14 nov. 2023. Disponível em: <https://www.gov.br/saude/pt->

br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/edicoes/2021/boletim-epidemiologico-vol-52-no-48.pdf/view. Acesso em: 27 dez. 2023.

[14] BARRETO, Maurício L.; TEIXEIRA, Maria Glória. Dengue no Brasil: situação epidemiológica e contribuições para uma agenda de pesquisa. *Estudos avançados*, v. 22, p. 53-72, 2008.

[15] ORGANIZAÇÃO PAN-AMERICANA DE SAÚDE. Dengue. [S. l.]: Organização Pan-americana de Saúde, 7 mar. 2024. Disponível em: <https://www.paho.org/pt/topicos/dengue#:~:text=O%20n%C3%BAmero%20de%20casos%20de,na%20d%C3%A9cada%20de%202010%2D2019.&text=n%C3%BAmero%20representa%20um%20aumento%20de,m%C3%A9dia%20dos%20%C3%BAltimos%205%20anos>. Acesso em: 9 mar. 2024.

[16] GRIMES, Craig A. *et al.* Theory, instrumentation and applications of magnetoelastic resonance sensors: a review. *Sensors*, v. 11, n. 3, p. 2809–2844, 2011.

[17] JACKSON, Roger G. *Novel sensors and sensing*. CRC Press, 2019.

[18] AUSANIO, Giovanni *et al.* Magnetoelastic sensor application in civil buildings monitoring. *Sensors and Actuators A: Physical*, v. 123, p. 290-295, 2005.

[19] HOLMES, Hal R. *et al.* Fabrication of biocompatible, vibrational magnetoelastic materials for controlling cellular adhesion. *Biosensors*, v. 2, n. 1, p. 57-69, 2012.

[20] VLAISAVLJEVICH, Eli *et al.* Magnetoelastic materials as novel bioactive coatings for the control of cell adhesion. *IEEE Transactions on Biomedical Engineering*, v. 58, n. 3, p. 698-704, 2010.

[21] TAN, Yisong *et al.* A passive and wireless sensor for bone plate strain monitoring. *Sensors*, v. 17, n. 11, p. 2635, 2017.

[22] CAI, Qingyun *et al.* A wireless, remote query glucose biosensor based on a pH-sensitive polymer. *Analytical Chemistry*, v. 76, n. 14, p. 4038-4043, 2004.

[23] CAI, Qing Y.; CAMMERS-GOODWIN, Arthur; GRIMES, Craig A. A wireless, remote query magnetoelastic CO₂ sensor. *Journal of Environmental Monitoring*, v. 2, n. 6, p. 556-560, 2000.

[24] RUAN, Chuanmin *et al.* Magnetoelastic immunosensors: amplified mass immunosorbent assay for detection of *Escherichia coli* O157: H7. *Analytical Chemistry*, v. 75, n. 23, p. 6494-6498, 2003.

[25] RUAN, Chuanmin *et al.* A staphylococcal enterotoxin B magnetoelastic immunosensor. *Biosensors and Bioelectronics*, v. 20, n. 3, p. 585-591, 2004.

[26] POSSAN, André L. Biosensor magnetoelástico para a detecção de *Escherichia coli*. Dissertação de Mestrado. 105 p. Programa de Pós-Graduação em Engenharia de Processos e Tecnologias. Universidade de Caxias do Sul. Caxias do Sul, 2015.

[27] REN, Limin; YU, Kun; TAN, Yisong. Applications and advances of magnetoelastic sensors in biomedical engineering: A review. *Materials*, v. 12, n. 7, p. 1135, 2019.

- [28] MENTI, Caroline *et al.* Antibody-based magneto-elastic biosensors: potential devices for detection of pathogens and associated toxins. *Applied Microbiology and Biotechnology*, v. 100, n. 14, p. 6149-6163, 2016.
- [29] CHEN, Guorui *et al.* Wearable respiratory sensors for COVID-19 monitoring. *View*, v. 3, n. 5, p. 20220024, 2022.
- [30] NARITA, Fumio *et al.* A review of piezoelectric and magnetostrictive biosensor materials for detection of COVID-19 and other viruses. *Advanced Materials*, v. 33, n. 1, p. 2005448, 2021.
- [31] SINGH, Bishal *et al.* A comprehensive review on current COVID-19 detection methods: From lab care to point of care diagnosis. *Sensors International*, v. 2, p. 100119, 2021.
- [32] VANDER HEYDEN, Yvan; PERRIN, Catherine; MASSART, Désiré L. Optimization strategies for HPLC and CZE. *In: Handbook of analytical separations*. Elsevier Science BV, 2000. p. 163-212.
- [33] LUDWIG, Alfred *et al.* High-frequency magnetoelastic materials for remote-interrogated stress sensors. *Journal of Magnetism and Magnetic Materials*, v. 242, p. 1126-1131, 2002.
- [34] CAI, Qing Y.; GRIMES, Craig A. A remote query magnetoelastic pH sensor. *Sens. Actuators B*, v. 71, p. 112-117, 2000.
- [35] BARANDIARÁN, J. M.; GUTIÉRREZ, J.; GARCÍA-ARRIBAS, A. Magnetoelasticity in amorphous ferromagnets: Basic principles and applications. *Physica Status Solidi (a)*, v. 208, n. 10, p. 2258-2264, 2011.
- [36] ENGDAHL, Göran; MAYERGOYZ, Isaak D. *Handbook of giant magnetostrictive materials*. San Diego: Academic press, 2000.
- [37] SABLİK, Martin J.; JILES, David C. Coupled magnetoelastic theory of magnetic and magnetostrictive hysteresis. *IEEE Transactions on Magnetics*, v. 29, n. 4, p. 2113-2123, 1993.
- [38] SZEWCZYK, Roman; BIEŃKOWSKI, Adam; SALACH, Jacek. Extended Jiles–Atherton model for modelling the magnetic characteristics of isotropic materials. *Journal of Magnetism and Magnetic Materials*, v. 320, n. 20, p. e1049-e1052, 2008.
- [39] CULLEN, James R. *et al.* Magnetoelasticity of Fe–Ga and Fe–Al alloys. *Journal of Magnetism and Magnetic Materials*, v. 226, p. 948-949, 2001.
- [40] DESIMONE, Antonio; JAMES, Richard D. A constrained theory of magnetoelasticity. *Journal of the Mechanics and Physics of Solids*, v. 50, n. 2, p. 283-320, 2002.
- [41] GRIMES, Craig A. *et al.* Wireless magnetoelastic resonance sensors: A critical review. *Sensors*, v. 2, n. 7, p. 294-313, 2002.
- [42] HERNANDO, A.; VAZQUEZ, M.; BARANDIARAN, M. Metallic glasses and sensing applications. *Journal of Physics. E: Scientific Instruments*, v. 21, p. 1129-1139, 1988.

- [43] MODZELEWSKI, C.; SAVAGE, H. T.; KABACOFF, L. T.; CLARK, A. E. Magnetomechanical coupling and permeability in transversely annealed Metglas 2605 alloys. *IEEE Transactions on Magnetics*, v. 17, n. 6, p. 2837-2839, 1981.
- [44] JAIN, Mahaveer K.; CAI, Qingyun; GRIMES, Craig A. A wireless micro-sensor for simultaneous measurement of pH, temperature, and pressure. *Smart Materials and Structures*, v. 10, n. 2, p. 347, 2001.
- [45] CROCE, Robert A. *et al.* A highly miniaturized low-power CMOS-based pH monitoring platform. *IEEE Sensors Journal*, v. 15, n. 2, p. 895-901, 2014.
- [46] MORI, Kotaro *et al.* On the possibility of developing magnetostrictive Fe-Co/Ni clad plate with both vibration energy harvesting and mass sensing elements. *Materials*, v. 14, n. 16, p. 4486, 2021.
- [47] KHAN, Muhammad Imran *et al.* A review on pH sensitive materials for sensors and detection methods. *Microsystem Technologies*, v. 23, n. 10, p. 4391-4404, 2017.
- [48] KOUZOUDIS, Dimitris; GRIMES, Craig A. The frequency response of magnetoelastic sensors to stress and atmospheric pressure. *Smart Materials and Structures*, v. 9, n. 6, p. 885, 2000.
- [49] NARITA, Fumio *et al.* A review of piezoelectric and magnetostrictive biosensor materials for detection of COVID-19 and other viruses. *Advanced Materials*, v. 33, n. 1, p. 2005448, 2021.
- [50] SISNIEGA, Beatriz *et al.* Improved determination of Q quality factor and resonance frequency in sensors based on the magnetoelastic resonance through the fitting to analytical expressions. *Materials*, v. 13, n. 21, p. 4708, 2020.
- [51] MELLIT, Adel; KALOGIROU, Soteris A. Artificial intelligence techniques for photovoltaic applications: A review. *Progress in Energy and Combustion Science*, v. 34, n. 5, p. 574-632, 2008.
- [52] MOOR, James. The Dartmouth College artificial intelligence conference: The next fifty years. *AI Magazine*, v. 27, n. 4, p. 87-87, 2006.
- [53] MCCARTHY, J. *et al.* Dartmouth summer research conference on artificial intelligence. Dartmouth College, 1956.
- [54] DOMINGOS, Pedro. A few useful things to know about machine learning. *Communications of the ACM*, v. 55, n. 10, p. 78-87, 2012.
- [55] VALDATI, Aline de Brittos. *Inteligência artificial–IA*. Curitiba: Contentus, 2020.
- [56] NILSSON, Nils J. *The quest for artificial intelligence*. Cambridge University Press, 2009.
- [57] BISHOP, Christopher M.; NASRABADI, Nasser M. *Pattern recognition and machine learning*. New York: Springer, 2006.
- [58] MACARIO FILHO, Valmir. Um novo algoritmo de agrupamento semi-supervisionado baseado no fuzzy c-means. *Dissertação de Mestrado*. 109 p. Universidade Federal de Pernambuco. Recife, 2009.
- [59] SAS - STATISTICAL ANALYSIS SYSTEM. *Machine Learning: what it is and why it matters*. What it is and why it matters. Disponível em:

https://www.sas.com/en_us/insights/analytics/machine-learning.html. Acesso em: 30 nov. 2020.

[60] AMARAL, Fernando. Aprenda mineração de dados: teoria e prática. São Paulo: Alta Books, 2016.

[61] SHALEV-SHWARTZ, Shai; BEN-DAVID, Shai. Understanding machine learning: from theory to algorithms. Cambridge: Cambridge University Press, 2014.

[62] BROWNLEE, Jason. Supervised and unsupervised machine learning algorithms. 2016. Disponível em: <https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/>. Acesso em: 30 nov. 2020.

[63] MONARD, M. C.; BARANAUSKAS, J. A. Conceitos sobre aprendizado de máquina. *In*: REZENDE, S. O. Sistemas inteligentes: fundamentos e aplicações. [S.l.]: Editora Manole Ltda, 2003. Cap. 4, p. 39-56.

[64] GERON, A. Hands-on Machine Learning with Scikit-Learn and Tensor Flow. O'Reilly Media Inc., USA, 2018.

[65] GRUS, Joel. Data science from scratch: first principles with python. O'Reilly Media, 2019.

[66] MARIANO, Diego. Métricas de avaliação em machine learning: acurácia, sensibilidade, precisão, especificidade e F-score. 2021. Disponível em: <https://doi.org/10.51780/978-6-599-275326-15>. Acesso em: 15 jun. 2024.

[67] CHEN, Tianqi; GUESTRIN, Carlos. XGBoost: A scalable tree boosting system. *In*: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. 2016. p. 785-794.

[68] INTRODUCTION to Boosted Trees. XGBoost Documentation. Disponível em: <https://xgboost.readthedocs.io/en/stable/tutorials/model.html>. Acesso em: 15 jun. 2024.

[69] EFRON, Bradley. The bootstrap and modern statistics. Journal of the American Statistical Association, v. 95, n. 452, p. 1293-1296, 2000.

[70] DAVISON, A. C.; HINKLEY, D. V. Bootstrap methods and their application, Cambridge University Press, Cambridge, 1997.

[71] REVELLAME, Emmanuel D. *et al.* Adsorption kinetic modeling using pseudo-first order and pseudo-second order rate laws: A review. Cleaner Engineering and Technology, v. 1, p. 100032, 2020.

[72] BISHOP, Christopher M. *et al.* Neural networks for pattern recognition. Oxford University Press, 1995.

[73] NIELSEN, Michael A. Neural networks and deep learning. San Francisco, CA, USA: Determination Press, 2015.

[74] ARANSAY, Andoni Lasheras *et al.* Magnetoelastic sensors: detection method of external agents. [S.l.]: Elhuyar, 14 maio 2020. Disponível em: <https://aldizkaria.elhuyar.eus/gai-librean/sentsore-magnetoelastikoak-kanpo-agenteak-detektat/en>. Acesso em: 27 dez. 2023.