

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Paulo Diego da Silva Souza

**Aprendizado de máquina para o diagnóstico da
hipersensibilidade dentinária utilizando
espectroscopia no infravermelho por
transformada de Fourier**

Uberlândia, Brasil

2023

UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Paulo Diego da Silva Souza

**Aprendizado de máquina para o diagnóstico da
hipersensibilidade dentinária utilizando espectroscopia no
infravermelho por transformada de Fourier**

Trabalho de conclusão de curso apresentado à Faculdade de Computação da Universidade Federal de Uberlândia, como parte dos requisitos exigidos para a obtenção título de Bacharel em Sistemas de Informação.

Orientador: Murillo Guimarães Carneiro

Universidade Federal de Uberlândia – UFU

Faculdade de Computação

Bacharelado em Sistemas de Informação

Uberlândia, Brasil

2023



UNIVERSIDADE FEDERAL DE UBERLÂNDIA

Faculdade de Computação

Av. João Naves de Ávila, nº 2121, Bloco 1A - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902

Telefone: (34) 3239-4144 - <http://www.portal.facom.ufu.br/> facom@ufu.br



ATA DE DEFESA - GRADUAÇÃO

Curso de Graduação em:	Bacharelado em Sistemas de Informação				
Defesa de:	Trabalho de Conclusão de Curso II				
Data:	29/11/2023	Hora de início:	17:00	Hora de encerramento:	18:30
Matrícula do Discente:	11921BSI237				
Nome do Discente:	Paulo Diego da Silva Souza				
Título do Trabalho:	Aprendizado de máquina para o diagnóstico da hipersensibilidade dentinária utilizando espectroscopia no infravermelho por transformada de Fourier				
A carga horária curricular foi cumprida integralmente?	<input checked="" type="checkbox"/> Sim <input type="checkbox"/> Não				

Reuniu-se na Sala 1B-132, Campus Santa Mônica, da Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Curso de Graduação em Sistemas de Informação, assim composta: Professores: Dr. Rodrigo Sanches Miani, Dr. Bruno Augusto Nassif Travençolo e Dr. Murillo Guimarães Carneiro, orientador(a) do(a) candidato(a).

Iniciando os trabalhos, o(a) presidente da mesa, Dr. Murillo Guimarães Carneiro, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao discente a palavra, para a exposição do seu trabalho. A duração da apresentação do discente e o tempo de arguição e resposta foram conforme as normas do curso.

A seguir o(a) senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

Aprovado Nota [100]

OU

Aprovado(a) sem nota.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Murillo Guimarães Carneiro**, **Professor(a) do Magistério Superior**, em 29/11/2023, às 19:10, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Bruno Augusto Nassif Travençolo, Professor(a) do Magistério Superior**, em 06/12/2023, às 14:19, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Rodrigo Sanches Miani, Professor(a) do Magistério Superior**, em 19/12/2023, às 10:33, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **5007364** e o código CRC **2B646A38**.

Resumo

A hipersensibilidade dentinária é um desafio que afeta significativamente a qualidade de vida dos pacientes, demandando métodos de diagnóstico não invasivos e confortáveis. Técnicas vibracionais, como a espectroscopia de infravermelho por transformada de Fourier (FTIR), oferecem uma abordagem rápida e não invasiva para investigar componentes bioquímicos em tecidos. Neste estudo, foram utilizados os dados de 40 amostras de fluido crevicular gengival (20 apresentavam diagnóstico de hipersensibilidade dentinária e 20 serviram como amostras de controle), submetidas à análise FTIR, com realização de pré-processamento, incluindo normalização pela Amida I e aplicação de *Savitzky-Golay* com derivadas. Vários modelos de aprendizado, incluindo KNN, Árvore de Decisão, Naive Bayes, XGBoost, MLP e BiLSTM, foram desenvolvidos e comparados. O XGBoost se destacou com resultados consistentes sob a normalização pela Amida I, demonstrando sua robustez. Além disso, o pré-processamento *Savitzky-Golay* teve um impacto positivo, especialmente no modelo BiLSTM, ressaltando a importância da escolha do pré-processamento. Esses resultados apoiam a viabilidade dos modelos de diagnóstico da hipersensibilidade dentinária e a importância da adequada preparação dos dados para otimizar o desempenho dos algoritmos de aprendizado de máquina.

Palavras-chave: Fluido Crevicular Gengival, Espectroscopia Infravermelho com Transformada de Fourier, Sensibilidade da Dentina, Aprendizado de Máquina, Pré-processamento de dados.

Abstract

Dentin hypersensitivity poses a significant challenge to patients' well-being, requiring the use of non-invasive and comfortable diagnostic techniques. Vibrational methods, such as Fourier-transform infrared spectroscopy (FTIR), provide a swift and non-intrusive means of examining biochemical constituents within tissues. In this study, data from 40 samples of gingival crevicular fluid (comprising 20 samples with dentin hypersensitivity diagnoses and 20 control samples) underwent FTIR analysis, with preprocessing steps encompassing normalization via Amide I and application of Savitzky-Golay with derivatives. Several machine learning models, including KNN, Decision Tree, Naive Bayes, XGBoost, MLP, and BiLSTM, were developed and compared. XGBoost stood out with consistent results under Amide I normalization, demonstrating its robustness. Additionally, Savitzky-Golay preprocessing had a positive impact, especially on the BiLSTM model, highlighting the importance of preprocessing choice. These results support the feasibility of dentin hypersensitivity diagnostic models and the importance of proper data preparation to optimize machine learning algorithm performance.

Keywords: Gingival Crevicular Fluid, Fourier Transform Infrared Spectroscopy, Dentin Sensitivity, Machine Learning, Data Preprocessing.

Lista de ilustrações

Figura 1 – Arquitetura da célula LSTM (Long short-term memory).	20
Figura 2 – Valores originais e Média representativa das amostras de ATR-FTIR para o grupo com hipersensibilidade dentinária (H.D) e grupo controle.	25
Figura 3 – Valores normalizados pelo pico da Amida I e média representativa das amostras normalizadas de ATR-FTIR para o grupo com hipersensibilidade dentinária (H.D) e grupo controle.	26
Figura 4 – Explicabilidade do modelo de XGBoost utilizando a técnica SHAP.	38
Figura 5 – Características que mais impactaram para a tomada de decisão do modelo de XGBoost.	39

Lista de tabelas

Tabela 1 – Parâmetros e valores exatos inicialmente utilizados no GridSearchCV .	28
Tabela 2 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo KNN.	30
Tabela 3 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo Árvore de Decisão.	31
Tabela 4 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo Naive Bayes.	32
Tabela 5 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo XGBoost.	32
Tabela 6 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo MLP.	33
Tabela 7 – Arquitetura do Modelo de Rede Neural	34
Tabela 8 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo BiLSTM.	35
Tabela 9 – Resultados da classificação em porcentagem (média e desvio padrão), utilizando a normalização pela Amida I.	36
Tabela 10 – Resultados dos modelos de classificação em porcentagem (média e desvio padrão), considerando a utilização do pré-processamento <i>Savitzky-Golay</i> e destacando os resultados mais promissores entre a aplicação da primeira, segunda e terceira derivadas.	37

Lista de abreviaturas e siglas

AD	Árvore de Decisão
BiLSTM	Rede neural Long Short-Term Memory bidirecional
FCG	Fluido crevicular gengival
FTIR	Espectroscopia no Infravermelho por Transformada de Fourier
HD	Hipersensibilidade dentinária
LSTM	Rede neural Long Short-Term Memory
MLP	Multilayer Perceptron
RNN	Rede neural recorrente

Sumário

1	INTRODUÇÃO	11
1.1	Objetivos	12
1.2	Justificativa	12
1.3	Organização do trabalho	12
2	REVISÃO BIBLIOGRÁFICA	14
2.1	Hipersensibilidade dentinária	14
2.2	Fluido Crevicular Gengival	14
2.3	Espectroscopia no Infravermelho	15
2.4	Algoritmos de Classificação	16
2.4.1	k-vizinhos mais próximos	16
2.4.2	Árvore de decisão	16
2.4.3	Naive Bayes	17
2.4.4	XGBoost	17
2.4.5	Redes Neurais	18
2.4.5.1	MLP - Multilayer Perceptron	18
2.4.5.2	BiLSTM - Bidirectional Long Short-Term Memory	19
2.5	Medidas de Avaliação	20
2.6	Trabalhos Correlatos	21
3	MATERIAIS E MÉTODOS	24
3.1	Base de Dados	24
3.2	Pré-processamento	24
3.3	Desenvolvimento dos modelos de classificação	27
3.4	Experimentos	28
4	RESULTADOS	30
4.1	KNN	30
4.2	Árvore de Decisão	31
4.3	Naive Bayes	31
4.4	XGBoost	32
4.5	MLP	33
4.6	BiLSTM	34
4.7	Análise comparativa	35
4.8	Explicabilidade utilizando a técnica SHAP	38

5	DISCUSSÃO	41
6	CONCLUSÃO	44
	REFERÊNCIAS	45

1 Introdução

A hipersensibilidade dentinária é uma condição clínica caracterizada por dor súbita, aguda e de curta duração, desencadeada por estímulos externos, sejam eles químicos, térmicos, táteis, osmóticos ou evaporativos, devido à perda de esmalte e cimento e consequente exposição de dentina subjacente (LIMA et al., 2021).

O fluido crevicular gengival (FCG) é constituído de uma mistura complexa de substâncias, composta de uma gama de biomarcadores celulares e moleculares, resultado da interação entre o biofilme bacteriano aderido à superfície do dente e as células do tecido periodontal. A coleta do FCG é uma técnica fácil e não invasiva, importante pelo grande potencial para servir como indicadores de doença periodontal, cicatrização pós-terapia e, mais recentemente, da hipersensibilidade dentinária com evidências de alterações químicas na composição do FCG utilizando a Espectroscopia Infravermelho com Transformada de Fourier (SOUZA, 2018).

A espectroscopia é uma técnica que fornece evidências da presença de grupos funcionais presentes na estrutura de uma substância, sendo uma ferramenta versátil para determinações qualitativas e quantitativas (SOUZA, 2018). Técnicas vibracionais, como a espectroscopia Raman e a espectroscopia de infravermelho com transformada de Fourier, são capazes de investigar de forma rápida e não invasiva componentes bioquímicos de células e tecidos com preparação mínima de amostra. Estas técnicas são consideradas uma ferramenta valiosa para a análise de processos biológicos complexos, como é o caso do FCG, frequentemente coletado na rotina clínica para diagnóstico e avaliação da progressão da doença periodontal (SAKUDO, 2016).

A área de Inteligência Artificial tem demonstrado avanços extraordinários, com uma crescente aplicação de diagnósticos automáticos, precisos e corretos, decorrentes de um período de euforia sobre os possíveis benefícios que a área pode prover. O aprendizado de máquina objetiva a construção de programas com melhoramento de desempenho através de uma gama de exemplos, com característica de aprender sem terem sido explicitamente programados. No Aprendizado de Máquina Supervisionado, o objetivo é construir um classificador capaz de determinar corretamente a classe de novos exemplos de rótulo ainda desconhecido. A utilização da Inteligência Artificial na área da saúde tem contribuído para auxiliar no processo de diagnóstico, com consequente melhoria nos resultados dos tratamentos (SICHMAN, 2021).

A grande parte dos modelos desenvolvidos por meio de Aprendizado de Máquina e aprendizado profundo têm sido caracterizados como ‘caixa-preta’, devido à complexidade, não linearidade e extrema dificuldade de interpretação e explicação de suas estruturas sub-

jacentes. Diante desse cenário, surgiu a demanda por arquiteturas de Inteligência Artificial Explicável, visando conferir maior transparência aos modelos, facilitar a interação humana e aumentar a confiabilidade em suas inferências (VILONE; LONGO, 2021).

1.1 Objetivos

O objetivo geral deste trabalho é desenvolver modelos de aprendizado supervisionado para o diagnóstico da hipersensibilidade dentinária a partir de dados de Espectroscopia no Infravermelho por Transformada de Fourier (FTIR). Os objetivos específicos são:

- Avaliar os métodos de pré-processamento e processamento existentes na literatura em face do problema em questão.
- Desenvolver e avaliar comparativamente diferentes modelos de aprendizado do estado-da-arte.
- Investigar técnicas de Inteligência Artificial Explicável (XAI) para proporcionar interpretabilidade aos modelos desenvolvidos.

1.2 Justificativa

A hipersensibilidade dentinária afeta diretamente a qualidade de vida dos portadores, enquanto há um desafio para realização de formas de diagnóstico livres de dor, desconforto, não-invasivo, de forma simples e segura (LIMA et al., 2016). O desenvolvimento de modelos de aprendizado supervisionado para o diagnóstico da hipersensibilidade dentinária pode contribuir e auxiliar no processo de diagnóstico, sendo este, de grande importância para o resultado dos tratamentos.

1.3 Organização do trabalho

O restante deste trabalho está dividido nos seguintes capítulos:

- O Capítulo 2 apresenta conceitos fundamentais para o entendimento e desenvolvimento do trabalho.
- O Capítulo 3 apresenta os materiais e métodos que foram utilizados, incluindo a base de dados, o pré-processamento e o desenvolvimento dos modelos de classificação.
- O Capítulo 4 apresenta os resultados obtidos, métricas de desempenho, uma comparação detalhada dos modelos considerados e explora a interpretabilidade do modelo de melhor desempenho.

- O Capítulo 5 apresenta a discussão, procurando analisar criticamente os resultados obtidos, contextualizando-os dentro do referencial teórico adotado e identificando padrões ou tendências que emergiram durante a pesquisa.
- O Capítulo 6 apresenta a conclusão, consolidando as principais descobertas, contribuições e implicações do estudo.

2 Revisão Bibliográfica

Nesse capítulo serão apresentados os principais fundamentos teóricos para o desenvolvimento deste trabalho, além dos trabalhos que foram utilizados como referência e os trabalhos relacionados.

2.1 Hipersensibilidade dentinária

A dentina é recoberta pelo esmalte na superfície da coroa dental e por uma fina camada de cemento em sua superfície radicular. A dentina é sensível aos estímulos pela extensão da lesão do processo odontoblástico e formação do complexo dentina-polpa, que faz com que a dentina seja afetada pela polpa e a polpa seja afetada por estímulos na dentina. Isso se dá pelo fato de haver micro túbulos na dentina que são preenchidos pelos prolongamentos celulares dos odontoblastos, circundados por fluido dentinário originários dos vasos sanguíneos da polpa dental. A sensibilidade da dentina aos estímulos não causa nenhum problema enquanto estiver coberta por esmalte e cemento, que são tecidos protetores (DAVARI; ATAEI; ASSARZADEH, 2013).

A hipersensibilidade dentinária é uma condição clínica caracterizada por dor súbita, aguda e de curta duração, em resposta a estímulos químicos, térmicos, táteis ou osmóticos que não podem ser explicados como decorrentes de qualquer outro defeito dentário ou condição. É um distúrbio prevalente e os tratamentos sugeridos nem sempre são suficientes ou alcançam sucesso, levando a problemas físicos, com dificuldades na higiene oral, psicológicos e que impactam na qualidade de vida do paciente. A prevalência em pessoas do sexo feminino com faixa etária entre 30 a 40 anos pode estar relacionado à dieta e higiene dental destes indivíduos. O diagnóstico da HD pode ser realizado com através de questionários e exame clínico, avaliando-se a intensidade da hipersensibilidade (DAVARI; ATAEI; ASSARZADEH, 2013; LIMA et al., 2021).

2.2 Fluido Crevicular Gengival

O diagnóstico das doenças periodontais geralmente se baseia em métodos clínicos e radiográficos e um dos desafios é encontrar testes capazes de fornecer diagnósticos precoces e que facilitem a prevenção. Os testes mais recentes tem focado na determinação de marcadores de atividades ou de predição por doenças. O estudo da resposta inflamatória tem sido foco, e um método viável desta análise é a utilização do fluido crevicular gengival (CHIBEBE et al., 2012).

O fluido crevicular gengival (FCG) é constituído de uma mistura complexa de substâncias, composta de uma gama de biomarcadores celulares e moleculares, resultado da interação entre o biofilme bacteriano aderido à superfície do dente e as células do tecido periodontal. A coleta de amostras de fluido crevicular gengival constituem um acesso ao estado fisiopatológico do periodonto. Isso se dá pela análise bioquímica, que fornece indícios do metabolismo celular do local, permitindo antecipar riscos de doenças e também determinando sua evolução (GUPTA, 2012; SOUZA, 2018).

Durante a inflamação, o fluxo do fluido crevicular aumenta e sua composição passa a ter características semelhantes a de um exsudato inflamatório. O aumento do fluxo expulsa bactérias e metabólitos do sulco gengival e contribui para a defesa do hospedeiro (SUBBARAO et al., 2019).

A coleta do FCG também é uma técnica fácil, importante pelo grande potencial para servir como indicadores de doença periodontal, cicatrização pós-terapia e, mais recentemente, da hipersensibilidade dentinária com evidências de alterações químicas na composição do FCG utilizando a Espectroscopia Infravermelho com Transformada de Fourier (SOUZA, 2018).

2.3 Espectroscopia no Infravermelho

A radiação infravermelha é uma radiação eletromagnética com comprimentos de onda mais longos do que os da luz visível, que quando direcionada à matéria, pode estimular o movimento de moléculas e ligações atômicas. Leal et al. (2018) explicam que quando uma radiação monocromática é aplicada a algum tipo de material transparente, seja gases, líquidos ou sólidos, a maior parte dessa radiação é transmitida sem mudanças na frequência, enquanto radiação com uma frequência diferente pode surgir e pode estar associada a propriedades rotacionais e vibracionais moleculares, que caracteriza o espalhamento inelástico ou Raman. Radiação em diferentes frequências significa que as moléculas liberam mais ou menos energia do que absorvem. O efeito Raman pode ser explicado como a interação da radiação eletromagnética com a matéria que causa adição ou subtração das energias vibracionais das moléculas à radiação incidente.

A espectroscopia é uma técnica que fornece evidências da presença de grupos funcionais presentes na estrutura de uma substância, sendo uma ferramenta versátil para determinações qualitativas e quantitativas (SOUZA, 2018). As técnicas espectroscópicas têm atraído a atenção devido à velocidade e precisão da análise da amostra. Métodos espectroscópicos como fluorescência, espectroscopia de massa e espectroscopia Raman e infravermelho (FT-IR) foram desenvolvidos e adaptados para fornecer diagnóstico de várias doenças em um estágio inicial (LEAL et al., 2018).

Técnicas vibracionais, como a espectroscopia Raman e a espectroscopia de infra-

vermelho com transformada de Fourier, são capazes de investigar de forma rápida e não invasiva componentes bioquímicos de células e tecidos com preparação mínima de amostra. Estas técnicas são consideradas uma ferramenta valiosa para a análise de processos biológicos complexos, como é o caso do FCG (SAKUDO, 2016).

2.4 Algoritmos de Classificação

As técnicas de classificação são úteis para lidar com grande quantidade de dados e utilizadas para prever rótulos de classes categóricas. Classificação é um processo de encontrar um modelo que seja capaz de diferenciar classes de dados ou conceitos, enquanto os modelos de classificação são usados para classificar dados recém-disponibilizados em um rótulo de classe. A classificação prevê rótulos de classes categóricas e classifica dados com base no conjunto de treinamento. Todo classificador possui um viés específico que o diferencia dos demais. Construir classificadores rápidos e precisos é uma tarefa importante na mineração de dados e para a descoberta de conhecimento (DUKITHA, 2019).

2.4.1 k-vizinhos mais próximos

k-vizinhos mais próximos (KNN) é um método simples de aprendizado de máquina, baseado no aprendizado por amostras de treinamento, com cada amostra representando um ponto em um espaço n-dimensional. Todas as amostras de treinamento são armazenadas em um espaço padrão n-dimensional. Para classificar uma amostra desconhecida, um classificador de k-vizinhos mais próximos procura no espaço padrão as k amostras de treinamento que estão mais próximas da amostra a ser avaliada. Existem diversas métricas de distância, e o uso específico varia de acordo com cada problema, embora a mais utilizada seja a distância Euclidiana. A escolha para o valor de K pode variar de acordo com a base de dados, não existe um valor único para a constante, sendo recomendável a utilização de valores ímpares, primos ou mesmo por avaliação empírica (ZHANG, 2016). A distância euclidiana pode ser calculada com a seguinte equação:

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2}. \quad (2.1)$$

2.4.2 Árvore de decisão

Uma árvore de decisão é um algoritmo recursivo (por divisão e conquista) de aprendizado de máquina supervisionado, utilizado para classificação e para regressão. Pode ser usado para prever categorias discretas e valores numéricos. A árvore de decisão consiste em nós que formam uma árvore enraizada, direcionada por um nó raiz que não

possui arestas de entrada, enquanto os outros nós folhas possuem uma aresta de entrada. Cada nó interno divide o conjunto de instâncias em dois ou mais subespaços de acordo com uma determinada função discreta dos valores dos atributos de entrada. Cada divisão tenta reduzir um conjunto de instâncias até que todos tenham a mesma classificação. A melhor divisão é aquela que fornece o maior ganho de informação (ROKACH; MAIMON, 2005).

2.4.3 Naive Bayes

Os classificadores bayesianos são classificadores estatísticos que podem prever a associação de classe com base em probabilidades. Em sua maioria, eles assumem uma suposição muito simplista de que todos os objetos a serem classificados são completamente independentes uns dos outros nos termos usados para caracterizá-los. Apesar de simplista, a técnica de classificação do Naive Bayes é bastante adequada quando a dimensionalidade das entradas é alta e pode, muitas vezes, superar métodos de classificação mais sofisticados. O desacoplamento das distribuições de recursos condicionais de classe significa que cada distribuição pode ser estimada independentemente como uma distribuição unidimensional, aliviando os problemas decorrentes da maldição da dimensionalidade. É necessário apenas uma pequena quantidade de dados de treinamento para estimar os parâmetros de médias e variâncias das variáveis, requeridos para a classificação (VIKRAMKUMAR; B; TRILOCHAN, 2014).

O Teorema de Bayes fornece uma forma de calcular a probabilidade posterior $P(c|x)$ a partir de $P(c)$, $P(x)$ e $P(x|c)$:

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}. \quad (2.2)$$

O método de classificação Naive Bayes é amplamente utilizado em diversas áreas, incluindo a área médica e de classificação de medicamentos para determinadas doenças. Indraja e Kumarappan (2018) implementaram um sistema usando Naive Bayes para classificar medicamentos que são usados contra as doenças de febre e febre tifóide obtendo bons valores de acurácia (80%), revocação (83%) e precisão (72%).

2.4.4 XGBoost

O algoritmo XGBoost, abreviação de “eXtreme Gradient Boosting”, é um sistema avançado de aprendizado de máquina projetado para aprimorar árvores de decisão. Ele pertence à família dos algoritmos de *gradient boosting*, reconhecidos por sua habilidade em construir modelos de alta precisão. O XGBoost emprega uma técnica de impulso, que combina várias árvores de decisão fracas para criar um modelo mais robusto e preciso. Uma característica distinta do XGBoost é a otimização do algoritmo para lidar com

desafios como *overfitting*, valores ausentes e a manipulação eficaz de dados de alta dimensionalidade. Isso é alcançado por meio de estratégias de regularização e técnicas de poda de árvores, resultando em modelos que são robustos e com capacidade excepcional de generalização para uma ampla gama de aplicações (CHEN; GUESTRIN, 2016).

O XGBoost também é eficaz em pequena escala, sendo uma opção versátil que se adapta tanto a problemas de aprendizado de máquina em grande escala quanto a cenários com conjuntos de dados menores. Sua eficiência e capacidade de lidar com conjuntos de dados de diferentes tamanhos o tornam uma escolha valiosa em uma ampla variedade de aplicações de aprendizado de máquina. Um exemplo disso é sua aplicação na predição da proeminência labial, com base em medições de tecidos duros e características demográficas de uma população (XING et al., 2023).

2.4.5 Redes Neurais

As redes neurais têm sido aplicadas com sucesso em uma ampla gama de aplicações de aprendizado supervisionado e não supervisionado. É um tipo de processo de aprendizado de máquina, chamado aprendizado profundo, que usa nós ou neurônios interconectados em uma estrutura em camadas, semelhante ao cérebro humano, que de forma geral aprendem e modelam relacionamentos entre os dados de entrada e de saída complexos e de forma não linear. Os métodos de redes neurais podem ser aplicados em uma infinidade de setores, que incluem o diagnóstico médico feito pela classificação de imagens, identificação de compostos químicos, processo e controle de qualidade (CHOI et al., 2020).

As redes neurais artificiais com conexões recorrentes são chamadas de redes neurais recorrentes (RNNs), que são capazes de modelar dados sequenciais para reconhecimento e predição de sequências. A estrutura dos estados ocultos, de alta dimensão e dinâmica não linear, funciona como a memória da rede e o estado da camada oculta por vez é condicionado ao seu estado anterior. Essa estrutura permite que as RNNs armazenem, lembrem e processem sinais complexos do passado por longos períodos de tempo e podem mapear uma sequência de entrada para a sequência de saída no tempo atual e prever a sequência no próximo passo de tempo (SALEHINEJAD et al., 2017).

2.4.5.1 MLP - Multilayer Perceptron

Um perceptron é um algoritmo de aprendizado de máquina que recebe uma série de recursos e seus alvos como entrada e tenta encontrar uma linha, plano ou hiperplano que separa as classes em um espaço bi, tridimensional ou hiperdimensional, respectivamente. Esses recursos são transformados usando a função sigmóide. Quando vários perceptrons estão conectados, o modelo é referido como um algoritmo perceptron multicamadas, consistindo de uma rede neural com uma ou mais camadas ocultas com um número inde-

terminado de neurônios. Uma rede MLP consiste de uma camada de entrada, uma ou mais camadas ocultas e uma camada de saída e é um exemplo típico de uma rede neural artificial *feedforward*, onde os dados fluem para a rede através da camada de entrada, passam pela camada oculta e finalmente saem da rede pela camada de saída, sem nenhum fluxo de dados diretamente cíclico. O algoritmo de treinamento de um MLP consiste em duas etapas principais: propagação direta e retropropagação de erros (SOARES; TEIVE, 2015).

Estas redes possuem a capacidade de tratar dados que não são linearmente separáveis, ou seja, apresentam um poder computacional muito maior devido, principalmente, a presença das camadas ocultas. Tem sido amplamente utilizada para auxiliar em diagnósticos da área médica, como no trabalho de Sharifi e Alizadeh (2020), que desenvolveram um novo método de classificação baseado na técnica de rede neural artificial perceptron multicamada para diagnóstico de doença renal crônica e concluíram que o sistema proposto foi mais eficaz e rápido do que outros métodos no diagnóstico de doença renal aguda e pode ser usado como ferramenta auxiliar do médico na prática clínica e de custo-efetivo para os pacientes (SHARIFI; ALIZADEH, 2020).

2.4.5.2 BiLSTM - Bidirectional Long Short-Term Memory

O aprendizado profundo ou aprendizado estruturado profundo pode ser definido como um tipo especial de redes neurais compostas por várias camadas. Essas redes são melhores do que as redes neurais tradicionais em persistir as informações de eventos anteriores. A rede neural recorrente (RNN) é uma dessas máquinas que possui uma combinação de redes em loop e que permitem que as informações persistam. As redes neurais recorrentes comuns ficam atrasadas no aprendizado, pois a lacuna entre as informações anteriores necessárias e o ponto de exigência aumenta em grande medida (KUMAR; GOOMER; SINGH, 2018).

A rede neural LSTM é uma arquitetura especial de RNN, distinguindo-se pela célula LSTM, uma unidade recorrente em série, em vez de uma simples célula RNN, como apresentado na Figura 1. Uma célula LSTM possui memória de longo prazo, que é atribuível a três portões que modulam o fluxo de informações. O portão de entrada controla quais informações serão passadas para a célula de memória com base na saída anterior e nos dados de medição do sensor atual. O portão de esquecimento controla como a célula de memória será atualizada, decide manter ou esquecer a informação. O portão de saída controla quais informações serão transportadas para o próximo intervalo de tempo (KONG et al., 2019).

As LSTMs são boas em lembrar informações por muito tempo. Como mais informações anteriores podem afetar a precisão do modelo, elas se tornam uma escolha natural de uso. As LSTMs foram modificadas desde a sua proposta inicial e incluem diversas va-

riações (LIPTON, 2015).

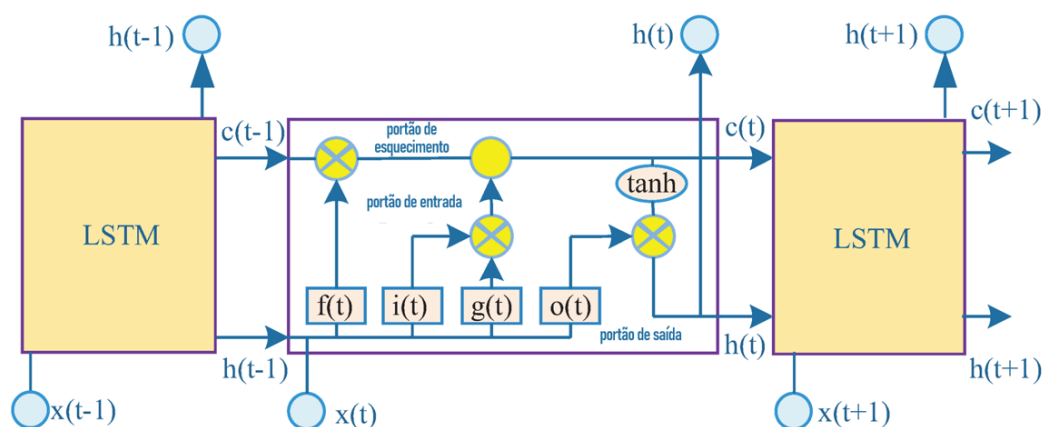


Figura 1 – Arquitetura da célula LSTM (Long short-term memory) adaptada de Kong et al. (2019).

Além disso, outra arquitetura recorrente notável é a BiLSTM (Bidirectional Long Short-Term Memory), que representa uma extensão das LSTMs. A BiLSTM é especialmente eficaz na análise de dados sequenciais, pois incorpora informações contextuais de ambas as direções da sequência. Isso é realizado por meio da execução simultânea de duas camadas de LSTM: uma lê a sequência da esquerda para a direita e a outra da direita para a esquerda. A combinação das informações de ambas as direções aprimora significativamente a capacidade do modelo em capturar dependências complexas e contextos em dados sequenciais, contribuindo para um desempenho mais robusto e preciso. São especialmente úteis nas ocasiões em que o contexto da entrada é necessário, muito útil em trabalhos de classificação de sentimentos (SHARFUDDIN; TIHAM; ISLAM, 2018).

2.5 Medidas de Avaliação

Utilizando os classificadores e as decisões tomadas por eles no conjunto de teste, constrói-se uma matriz bidimensional com as classes verdadeira e predita, denominada de matriz de confusão. Considerando um problema de classificação com somente duas classes, onde os rótulos positivos indicam hipersensibilidade dentinária e os negativos indicam a ausência da doença, obtém-se o número de amostras positivas classificadas corretamente (verdadeiros positivos - VP), positivas classificadas erroneamente (Falsos negativos - FN), negativas classificadas corretamente (Verdadeiros negativos - VN) e negativas classificadas erroneamente (Falsos positivos - FP).

A partir da matriz de confusão, várias medidas podem ser utilizadas para avaliar a performance de uma hipótese. A acurácia indica uma performance geral do modelo,

avaliando dentre todas as classificações, quantas o modelo classificou corretamente:

$$Acurácia = \frac{VP + VN}{VP + VN + FP + FN}. \quad (2.3)$$

A precisão indica quantas avaliações corretas dentre todas as classificações de classe positivo que o modelo fez:

$$Precisão = \frac{VP}{VP + FP}. \quad (2.4)$$

A sensibilidade (*recall*) indica quantas avaliações corretas dentre todas as situações de classe positivo como valor esperado que o modelo realizou:

$$Sensibilidade = \frac{VP}{VP + FN}. \quad (2.5)$$

A especificidade indica quantas avaliações corretas foram obtidas dentre todas as situações de classe negativo como valor esperado:

$$Especificidade = \frac{VN}{VN + FP}. \quad (2.6)$$

A F-score é a média harmônica entre precisão e sensibilidade, embora possa ser adaptada como a média harmônica entre sensibilidade e especificidade, com bastante uso na área médica (SAMPAIO et al., 2019):

$$Fscore = \frac{2}{\frac{1}{sensibilidade} + \frac{1}{especificidade}}. \quad (2.7)$$

2.6 Trabalhos Correlatos

A utilização da Inteligência Artificial na área da saúde tem sido ampla e contribuído para auxiliar nos processos de diagnóstico e prevenção de doenças.

Este trabalho tem como referência o estudo de Souza (2018), que avaliou a composição química do fluido crevicular gengival em dentes com hipersensibilidade dentinária comparados a dentes controle através da espectroscopia de infravermelho com transformada de Fourier e da espectroscopia Raman. Através de seu estudo, comprovou-se a presença de modificações químicas no fluido crevicular gengival de dentes com hipersensibilidade dentinária quando comparado ao fluido de dentes controle.

Há estudos odontológicos utilizando inúmeros classificadores, em bases de dados diversas e com a obtenção de resultados positivos e significativos. Alguns pesquisadores utilizaram uma rede neural artificial MLP para diagnóstico de cárie dentária proximal. A rede foi treinada para fazer diagnóstico radiográfico e as comparações foram feitas por meio da análise da característica operacional do receptor (ROC) da precisão diagnóstica obtida com e sem o uso da rede. Confirmaram a aplicação potencial da rede neural em

odontologia, com a conclusão de que o uso de uma rede neural adequada pode melhorar o desempenho do diagnóstico de cárie proximal (DEVITO; BARBOSA; FILHO, 2008).

Métodos bayesianos estão sendo cada vez mais usados em pesquisas clínicas. A abordagem bayesiana é ideal para adaptar informações sempre acumuladas, encontradas na prática odontológica. Chattopadhyay et al. (2010) aplicaram o classificador Bayesiano para o diagnóstico de dor dentária e obtiveram resultados promissores. Os autores também propuseram que o classificador possa ser testado em clínicas odontológicas, com o intuito de auxiliar dentistas iniciantes no diagnóstico de doenças odontológicas com base nos sintomas de dor. Wijaya (2018) realizou a análise de implementação dos Métodos GLCM (Matriz de Coocorrência em Nível de Cinza) e Naive Bayes na realização de extrações em imagens odontológicas, concluindo que o Naive Bayes pode realizar a classificação de imagens com base nas anormalidades de textura dos dentes extraídos com o método de matriz de coocorrência, mesmo com a acurácia não ideal obtida, provavelmente pelos intervalos espaçados entre as classes.

Wang, Tian e Qiu (2019) apresentaram um método de previsão de risco de múltiplas doenças com análise sistemática com base nos registros médicos longitudinais de pacientes, utilizando as redes neurais recorrentes do tipo LSTM. Obtiveram bons resultados, concluindo que as LSTMs funcionam bem em diferentes níveis de agregação de diagnóstico, e o modelo desenvolvido pode ser utilizado para ajudar na prevenção de riscos futuros das doenças, sendo que a abordagem pode ser desenvolvida como uma ferramenta de referência para os sistemas de informação hospitalares.

Um grupo de pesquisadores tentou determinar os espectros de infravermelho com transformada de Fourier (FTIR) de exossomos salivares de pacientes com câncer oral e indivíduos saudáveis para avaliar seu potencial diagnóstico usando modelos auxiliados por computador. Conseguiram detectar mudanças sutis nas conformações de proteínas, lipídios e ácidos nucleicos usando redes neurais artificiais otimizadas com pequenos conjuntos de dados (ZLOTOGORSKI-HURVITZ et al., 2019). Filho e Carneiro (2023) investigaram técnicas de classificação de alto nível baseadas em propriedades e medidas de redes complexas para a detecção salivar de câncer de boca a partir da ATR-FTIR, evidenciando a importância das medidas de proximidade na contribuição de redes mais efetivas, além de terem desenvolvido uma técnica eficiente e adaptável, de enorme potencial para classificação de alto nível para o problema em questão. A ATR-FTIR tem contribuído para a detecção do câncer oral, com o desenvolvimento de modelos preditivos de bom desempenho, combinando-se métodos de correção de linha de base e técnicas de classificação, como realizado no trabalho de Filho et al. (2023).

A combinação de espectroscopia ATR-FTIR e técnicas de aprendizado de máquina em diversas amostras biológicas tem se destacado como uma ferramenta alternativa, mesmo em comparação com métodos altamente precisos, devido às suas características de

ser não invasiva, rápida e de baixo custo. [Caixeta et al. \(2023\)](#) utilizaram a ATR-FTIR associada à análise discriminante linear (LDA) e um classificador de máquina de vetores de suporte (SVM) para identificar alterações nos componentes salivares a serem utilizados como biomarcadores alternativos para o diagnóstico do Diabetes Mellitus tipo 2 e destacaram o potencial das plataformas ATR-FTIR aliadas ao aprendizado de máquina como ferramenta sustentável, de possível aplicação para triagem e monitoramento de pacientes diabéticos.

Sharifi e Alizadeh desenvolveram um novo método de classificação baseado na MLP para diagnóstico de doença renal crônica, com a conclusão de que o sistema proposto foi mais eficaz e rápido do que outros métodos no diagnóstico de doença renal aguda, além de poder ser utilizado como ferramenta auxiliar ([SHARIFI; ALIZADEH, 2020](#)). [Brito e Oliveira \(2020\)](#) realizaram uma análise de características baseada na rede neural Multi-Layer Perceptron (MLP) utilizando técnicas de *feature selection* para o pré-diagnóstico da doença COVID-19, com a obtenção de resultados expressivos.

3 Materiais e Métodos

Este capítulo descreve os materiais e métodos que foram utilizados para o desenvolvimento do trabalho.

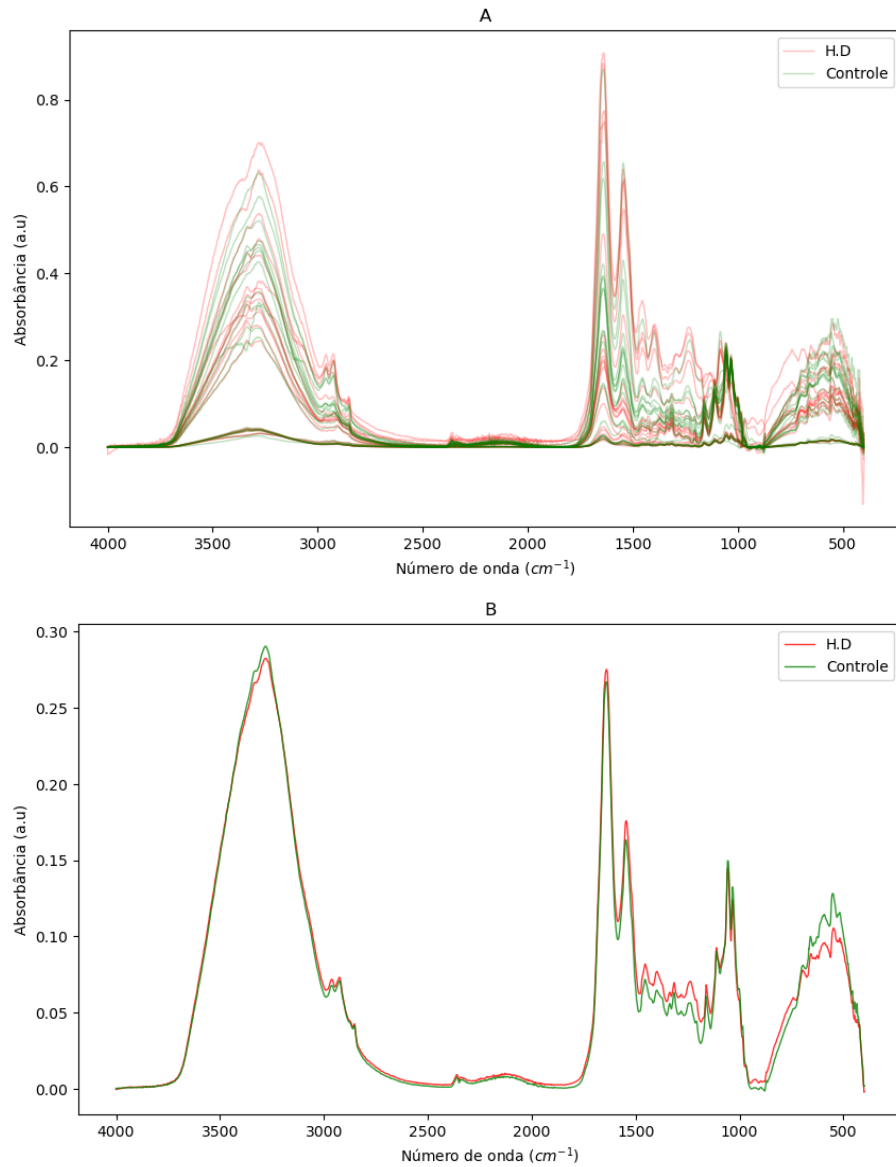
3.1 Base de Dados

A base de dados é proveniente do trabalho de Souza (SOUZA, 2018), composta de 40 amostras de 20 pacientes (dois dentes posteriores da mesma arcada de cada paciente) com faixa etária de 20 a 50 anos, dentição completa, perfil gengival semelhante, sem periodontite e boas condições de higiene oral. O diagnóstico da hipersensibilidade dentinária foi realizado através da estimulação pelo teste do jato de ar por 2 segundos na região cervical e mensuração da dor pela escala visual analógica, totalizando 20 amostras de controle e 20 amostras com diagnóstico positivo.

Cada uma das 40 amostras possui, como atributos, valores da região do infravermelho vibracional resultantes da análise dos espécimes contendo fluido crevicular, os quais foram submetidos ao espectrômetro de infravermelho por transformada de Fourier de refletância total. A radiação na região do infravermelho vibracional do espectro eletromagnético é frequentemente avaliada em termos de número de onda, expressos em centímetros recíprocos (cm^{-1}), calculados pelo recíproco do comprimento de onda expresso em centímetros e que podem ser facilmente convertidos em frequência. Um número de onda maior implica em maior energia, portanto diretamente proporcional. Levando em conta o número de onda, o infravermelho vibracional está entre 4000 e 400 cm^{-1} , correspondentes aos comprimentos de onda de 2,5 a 25 μm (LEAL et al., 2018).

3.2 Pré-processamento

Os dados foram pré-processados, a fim de torná-los mais adequados e em conformidade com as técnicas de classificação selecionadas e desenvolvimento de modelos melhores. Foi realizada a avaliação das características e qualidade do conjunto de dados, verificação de ruídos, artefatos, valores ausentes, além da definição de estratégias para lidar com estas ocorrências. Foi realizada a normalização dos espectros pelo pico da amida I (região entre 1630 e 1660 cm^{-1}). Os dados originais e normalizados são apresentados nas Figuras 2 e 3, respectivamente. O filtro digital de *Savitzky-Golay* é utilizado para filtragem e suavização de dados, incluindo as derivadas do sinal, com a utilização de coeficientes provenientes da realização de um ajuste linear não-ponderado de mínimos quadrados, com um polinômio de ordem específica no centro de uma janela de análise (ZHAO et al., 2014). Considerando

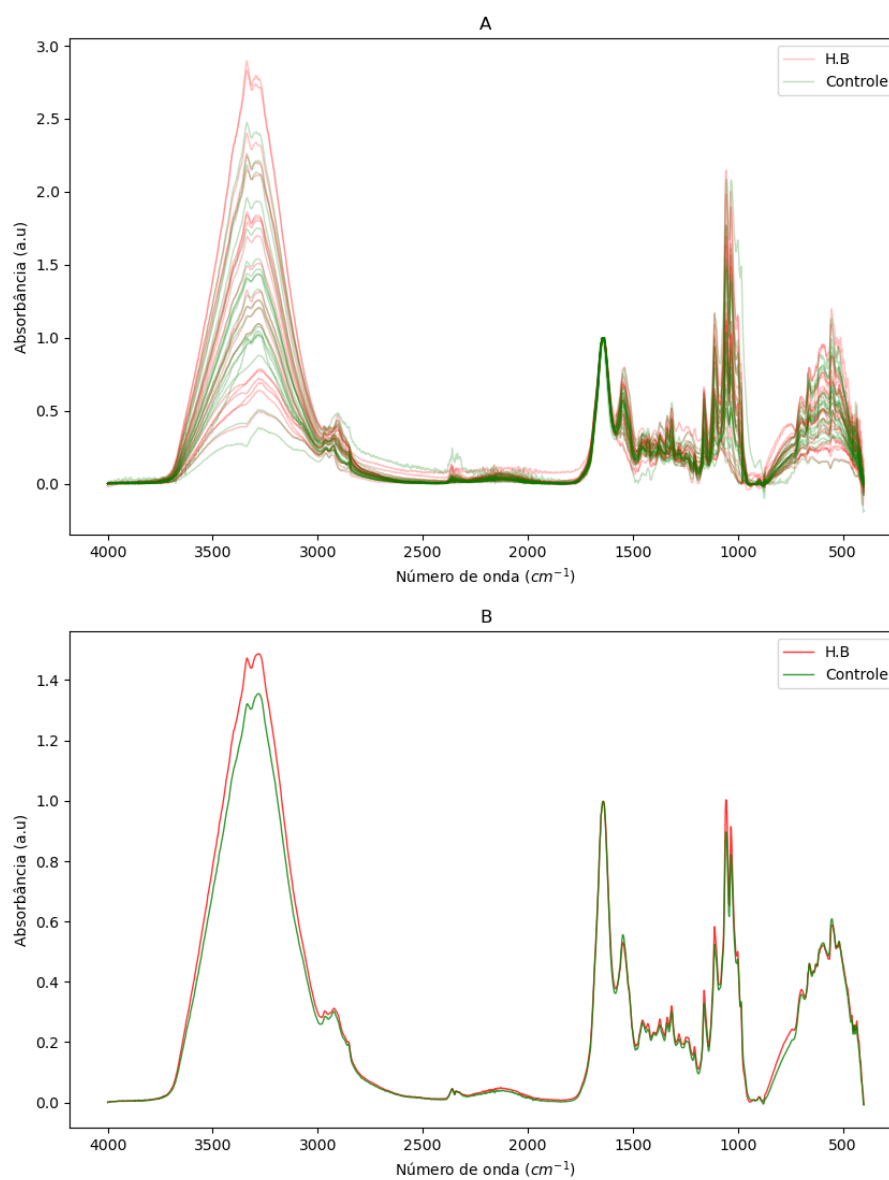


Fonte: Autoria própria.

Figura 2 – Valores originais (A) e Média representativa (B) das amostras de ATR-FTIR para o grupo com hipersensibilidade dentinária (H.D) e grupo controle.

a otimização de parâmetros e sua aplicação em espectros de FTIR, o filtro Savitzky–Golay foi empregado com uma janela de tamanho de 20, um polinômio de grau 4 e ordens de derivação de 1, 2 e 3.

Por fim, foi realizado o truncamento dos espectros para a região entre 1800 cm^{-1} e 900 cm^{-1} , a fim de evitar ruídos e outliers.



Fonte: Autoria própria.

Figura 3 – Valores normalizados pelo pico da Amida I (A) e média representativa das amostras normalizadas (B) de ATR-FTIR para o grupo com hipersensibilidade dentinária (H.D) e grupo controle.

3.3 Desenvolvimento dos modelos de classificação

Python versão 3.9.13 foi empregado no processo de desenvolvimento dos modelos. A biblioteca scikit-learn, versão 1.2.2, foi utilizada para os algoritmos KNN, árvore de decisão e Naive Bayes. Para os modelos baseados em XGBoost, foi utilizada a versão 2.0.0 e a integração com scikit-learn foi adotada, ao passo que a biblioteca TensorFlow, versão 2.14.0, foi utilizada na criação dos modelos de redes neurais.

A rede neural MLP foi utilizada de forma similar a (BRITO; OLIVEIRA, 2020), com treinamento supervisionado através de etapas:

- Apresentar um padrão às unidades da camada de entrada, e estimativa do erro através dos cálculos de resposta produzidos na camada de saída.
- Propagação do erro a partir da camada de saída até a camada de entrada, e modificação dos pesos das conexões das unidades das camadas internas de acordo com a regra delta generalizada.
- Após o treinamento, a rede será utilizada para a classificação do novo conjunto de dados.

O desempenho de um modelo está fortemente vinculado à configuração dos hiperparâmetros, cuja otimização é uma tarefa desafiadora, frequentemente exigindo experimentação iterativa. O teste manual pode consumir tempo e recursos consideráveis, e para automatizar o ajuste dos hiperparâmetros, recorreu-se à biblioteca GridSearchCV.

Os hiperparâmetros iniciais considerados para cada algoritmo são apresentados na tabela 1. Alguns valores de parâmetros foram posteriormente inseridos (além dos valores padrões e iniciais) na etapa de experimentação, tendo em vista a melhoria dos modelos gerados, como ajustes nos parâmetros de ‘beta_1’ e ‘beta_2’ para a MLP.

Para os modelos baseados em redes neurais BiLSTM (Bidirectional Long Short-Term Memory), foram analisadas diversas arquiteturas, que incluíram a variação do número de camadas LSTM (1, 2 e 3), número de unidades (20 a 200), aplicação de camadas de dropout (valores entre 0 e 1), aplicação de camadas densamente conectadas e diferentes tipos de regularização. Os modelos foram compilados utilizando a biblioteca Keras, que é uma parte integrante do TensorFlow. Uma semente específica (7) foi configurada para o gerador de números aleatórios do TensorFlow, o que resulta em números aleatórios determinísticos e previsíveis, possibilitando a replicação dos resultados.

A compilação de um modelo é um passo crucial no processo de treinamento, onde são definidos três elementos-chave. Primeiro, o otimizador foi configurado como o Adam, um algoritmo amplamente adotado para otimização de redes neurais. A taxa de aprendizado (learning_rate) para o otimizador foi testada com diversos valores entre 0 e 1.

Tabela 1 – Parâmetros e valores exatos inicialmente utilizados no GridSearchCV

<i>Algoritmo</i>	<i>Parâmetros</i>	<i>Valores</i>
KNN	n_neighbors metrics	1, 3, 5, 7, 9 cosine, euclidean
Árvore de Decisão	criterion max_depth splitter min_samples_split min_samples_leaf	gini, entropy 5, 10, 15, 20 best, random 2, 5, 10 1, 5, 10
Naive Bayes	var_smoothing	$1e^{-9}$, $1e^{-8}$, $1e^{-7}$
XGBoost	objective eta subsample reg_lambda n_estimators early_stopping_rounds	binary:logistic 0.1, 0.3, 0.5, 0.7, 1 0.1, 0.3, 0.5, 0.7, 1 0.1, 0.3, 0.5, 0.7, 0.9, 1 10, 25, 50, 75, 100 10, 20, 30
MLP	solver activation learning_rate_init max_iter hidden_layer_sizes	adam relu 0.009, 0.005, 0.01, 0.05, 0.1 100, 500, 1000 (10,),(50,),(100,),(10,10), (10,50),(50,10),(50,50)

Fonte: Autoria própria.

Este parâmetro é essencial para controlar o tamanho das atualizações de peso durante o treinamento. Em seguida, a função de perda foi definida como ‘binary_crossentropy’, indicando que o modelo é destinado a tarefas de classificação binária. Essa função de perda quantifica a discrepância entre as previsões do modelo e os rótulos reais. Por fim, a métrica de acurácia foi especificada para avaliar o desempenho do modelo, medindo a proporção de previsões corretas em relação ao total de exemplos. Foi empregado o método EarlyStopping como uma técnica de treinamento de modelos de rede neural. A função EarlyStopping é utilizada para monitorar o processo de treinamento e interrompê-lo quando determinados critérios são atendidos.

3.4 Experimentos

Para a condução dos experimentos, foi utilizada a técnica de k-fold cross-validation, com valor de k igual a 10. O conjunto de dados, com 40 amostras, foi dividido em 10 partições mutualmente exclusivas, a cada iteração, 9 partições foram utilizadas para treinar o modelo, enquanto uma delas foi utilizada para teste. Para se obter uma estimativa mais

precisa do desempenho dos modelos no conjunto de dados, foram realizadas 10 repetições utilizando a função ‘repeatedkfold’, considerando o resultado médio e o desvio padrão de todas as execuções. Para que se pudesse replicar os experimentos, também foi considerado o valor do parâmetro ‘random_state’ para cada algoritmo.

A análise e comparação de desempenho dos modelos foi realizada com base nas medidas de sensibilidade, especificidade, acurácia, precisão, taxa de FP, análise da matriz de confusão e principalmente pela média harmônica entre sensibilidade e especificidade (F-score).

4 Resultados

Com base na abordagem metodológica adotada, os resultados são apresentados neste capítulo, proporcionando uma visão abrangente das capacidades e limitações de cada modelo, fornecendo esclarecimentos valiosos para a seleção e otimização de modelos em contextos de aprendizado de máquina.

Para aprimorar os modelos e assegurar a replicação dos experimentos, uma série de novos parâmetros foi considerada e integrada à lista de avaliação pelo GridSearchCV. Dentre esses parâmetros, destaca-se o amplamente empregado ‘random_state’, com uma função crítica no contexto de algoritmos de machine learning. Sua principal finalidade reside em controlar a aleatoriedade, garantindo a consistência e reprodutibilidade dos resultados em diferentes execuções. Ao conferir estabilidade à inicialização aleatória, o ‘random_state’ desempenha um papel fundamental em cenários de validação cruzada, mitigando problemas de variabilidade e, simultaneamente, simplificando o processo de ajuste fino de hiperparâmetros.

Uma análise individual dos resultados para cada algoritmo e uma análise comparativa são apresentadas a seguir.

4.1 KNN

Considerando a normalização pela Amida I, o algoritmo KNN (Tabela 2) apresentou 48,7% de acurácia, 48,8% de precisão, 42,5% de sensibilidade, 55% de especificidade e 47,7% de F1-score. A especificidade apresentou a maior variabilidade, com um desvio padrão de 7,7%, enquanto as outras métricas mantiveram desvios padrão entre 5,1% e 6,5%.

Tabela 2 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo KNN.

Pré-processamento	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Amida I	48,7 ±5,6%	48,8 ±6,5%	42,5 ±5,1%	55,0 ±7,7%	47,7 ±5,5%
Savitzky-Golay	54,7 ±4,8%	54,9 ±5,0%	53,5 ±6,3%	56,0 ±6,6%	54,4 ±4,7%

Fonte: Autoria própria.

A aplicação do método de pré-processamento *Savitzky-Golay* resultou em melhorias significativas, especialmente ao empregar a primeira derivada. O modelo obteve 54,7% de

acurácia, 54,9% de precisão, 53,5% de sensibilidade, 56% de especificidade e 54,4% de F1-score. Todas as métricas apresentaram desvios padrão entre 4,7% e 6,6%, sendo a especificidade a com maior variabilidade.

4.2 Árvore de Decisão

Ao considerar a normalização pela Amida I, o modelo baseado no algoritmo de Árvore de Decisão (Tabela 3) demonstrou uma acurácia de 57%, uma precisão de 56,5%, uma sensibilidade de 61,5%, uma especificidade de 52,5% e um F1-score de 56%. Foi observado que tanto a sensibilidade quanto a especificidade apresentaram as maiores variações, com desvios padrão de 9,5% e 10,1%, respectivamente. Enquanto isso, as demais métricas exibiram desvios padrão variando de 7,4% a 8%.

Tabela 3 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo Árvore de Decisão.

Pré-processamento	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Amida I	57,0 ±7,7%	56,5 ±7,4%	61,5 ±9,5%	52,5 ±10,1%	56,0 ±8,0%
Savitzky-Golay	55,5 ±6,0%	55,4 ±5,5%	54,0 ±10,9%	57,0 ±5,6%	54,8 ±5,7%

Fonte: Autoria própria.

A aplicação do método de pré-processamento *Savitzky-Golay* produziu uma melhoria notável na especificidade, mas resultou em um desempenho ligeiramente inferior nas demais métricas, embora tenha sido observada uma menor variação nos resultados. Os resultados mais promissores foram alcançados com a utilização da primeira derivada. O modelo registrou uma acurácia de 55,5%, uma precisão de 55,4%, uma sensibilidade de 54%, uma especificidade de 57% e um F1-score de 54,8%. As métricas em geral apresentaram desvios padrão situados na faixa de 5,5% a 6%, com a exceção da sensibilidade, que apresentou um desvio padrão de 10,9%.

4.3 Naive Bayes

Quando a normalização pela Amida I foi levada em consideração, o modelo fundamentado no algoritmo Naive Bayes (Tabela 4) exibiu uma acurácia de 53%, uma precisão de 53,6%, uma sensibilidade de 46%, uma especificidade de 60%, e um F1-score de 51,8%. Observou-se que a acurácia, a precisão e o F1-score apresentaram desvios padrão de 3,1%, 3,6%, e 3,3%, respectivamente. Entretanto, a sensibilidade e a especificidade demonstraram uma maior variabilidade, com desvios padrão mais significativos de 4,9% e 5,5%, respectivamente.

Tabela 4 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo Naive Bayes.

Pré-processamento	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Amida I	53,0 ±3,1%	53,6 ±3,6%	46,0 ±4,9%	60,0 ±5,5%	51,8 ±3,3%
Savitzky-Golay	50,8 ±4,3%	50,3 ±4,1%	43,5 ±4,8%	58,0 ±4,6%	49,1 ±3,9%

Fonte: Autoria própria.

O pré-processamento *Savitzky-Golay* produziu resultados mais favoráveis quando aplicado à terceira derivada; no entanto, houve uma piora significativa nos resultados em comparação com a técnica anterior. O modelo obteve uma acurácia de 50,8%, uma precisão de 50,3%, uma sensibilidade de 43,5%, uma especificidade de 58%, e um F1-score de 49,1%. O F1-score mostrou menor variabilidade, com um desvio padrão de 3,9%, enquanto a sensibilidade apresentou a maior variabilidade, com 4,8%.

4.4 XGBoost

Os modelos que fazem uso do algoritmo XGBoost foram submetidos a ajustes destinados a serem avaliados pelo GridSearchCV, com o objetivo de mitigar o risco de overfitting ou underfitting. Essas modificações envolveram a inclusão dos valores 0,53 e 0,6 para o parâmetro ‘subsample’, o valor 10 para o ‘max_depth’ e 20 para ‘early_stopping_rounds’.

Tabela 5 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo XGBoost.

Pré-processamento	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Amida I	74,7 ±6,4%	74,3 ±7,9%	77,0 ±5,1%	72,5 ±10,8%	74,2 ±7,9%
Savitzky-Golay	68,0 ±6,4%	66,3 ±6,1%	74,5 ±7,2%	61,5 ±10,3%	66,7 ±7,5%

Fonte: Autoria própria.

Quando aplicamos a normalização pela Amida I (Tabela 5), o modelo demonstrou um desempenho com uma acurácia de 74,7%, uma precisão de 74,3%, uma sensibilidade de 77%, uma especificidade de 72,5%, e um F1-score de 74,2%. Destaca-se que a sensibilidade demonstrou a menor flutuação, com um desvio padrão de 5,1%, enquanto a especificidade apresentou a maior variabilidade, atingindo 10,8%.

O pré-processamento *Savitzky-Golay* proporcionou resultados superiores ao ser aplicado na terceira derivada. O modelo alcançou 68% de acurácia, 66,3% de precisão,

74,5% de sensibilidade, 61,5% de especificidade e um F1-score de 66,7%. É importante notar que a precisão e a acurácia tiveram os menores desvios padrão, 6,1% e 6,4%, respectivamente, enquanto a sensibilidade e o F1-score apresentaram desvios intermediários, em torno de 7,2% e 7,5%. A especificidade, por outro lado, mostrou a maior variabilidade, atingindo 10,3%.

4.5 MLP

Os parâmetros foram ajustados com o propósito de avaliação através do Grid-SearchCV, resultando em melhorias notáveis nos modelos. Essas adaptações envolveram a configuração dos seguintes valores: 0,84 para o parâmetro ‘beta_1’, que representa a taxa de decaimento exponencial para estimativas do primeiro vetor de momento usado no otimizador de peso Adam, e 0,94 para ‘beta_2’, a taxa de decaimento exponencial para estimativas do segundo vetor de momento.

Tabela 6 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo MLP.

Pré-processamento	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Amida I	58,5 ±3,6%	59,7 ±4,7%	52,5 ±5,6%	64,5 ±4,7%	57,6 ±3,6%
Savitzky-Golay	60,7 ±2,8%	60,3 ±2,6%	63,0 ±4,6%	58,5 ±3,9%	60,5 ±2,7%

Fonte: Autoria própria.

Ao empregarmos a normalização pela Amida I (Tabela 6), o modelo final revelou um desempenho com os seguintes resultados: uma acurácia de 58,5%, uma precisão de 59,7%, uma sensibilidade de 52,5%, uma especificidade de 64,5% e um F1-score de 57,6%. Vale ressaltar que a acurácia manteve-se com a menor flutuação, apresentando um desvio padrão de 3,6%, enquanto a sensibilidade demonstrou a maior variabilidade, com um desvio padrão de 5,6%.

Por outro lado, o pré-processamento *Savitzky-Golay* com aplicação da terceira derivada resultou em melhorias notáveis nos resultados e proporcionou uma boa estabilidade do modelo, que atingiu 60,7% ($\pm 2,8\%$) de acurácia, 60,3% ($\pm 2,6\%$) de precisão, 63% ($\pm 4,6\%$) de sensibilidade, 58,5% ($\pm 3,9\%$) de especificidade e um F1-score de 60,5% ($\pm 2,7\%$). Destaca-se que a média dos desvios padrão foi a menor em comparação com todos os outros modelos analisados neste trabalho.

4.6 BiLSTM

A arquitetura da rede neural do modelo final consiste em três camadas distintas (Tabela 7). A primeira camada é uma Bidirectional LSTM, composta por 155 unidades. Uma taxa de dropout de 0,1 é aplicada para prevenir o overfitting, e a função de ativação ‘relu’ é utilizada para calcular as saídas. Além disso, a camada não retorna sequências de saída. A segunda camada é uma camada de Dropout com uma taxa de dropout de 0,64, projetada para desativar aleatoriamente unidades durante o treinamento e evitar overfitting. Por fim, a terceira camada é uma camada Dense com uma única unidade e uma função de ativação ‘sigmoid’, adequada para tarefas de classificação binária. Essa arquitetura é caracterizada pela sua capacidade de processar sequências bidirecionais, incorporando mecanismos de regularização para melhorar o desempenho do modelo em tarefas de classificação.

Quando se utiliza uma taxa de aprendizado mais alta, o treinamento pode ocorrer mais rapidamente, e o modelo pode convergir mais rapidamente para um mínimo global da função de perda. No entanto, taxas de aprendizado mais altas também podem tornar o treinamento instável, levando a oscilações na perda (YU et al., 2020). A utilização da técnica de EarlyStopping pode tornar a taxa de aprendizado mais alta uma opção viável, monitorando a métrica de perda e interrompendo o treinamento assim que a perda começa a aumentar no conjunto de validação, indicando que o modelo pode estar sobreajustando aos dados.

Tabela 7 – Arquitetura do Modelo de Rede Neural

<i>Camada</i>	<i>Parâmetros</i>
1 - Bidirectional LSTM	Número de unidades: 155 Tamanho da entrada: (1, 901) Taxa de dropout: 0,1 Função de ativação: relu Retornar sequências: False
2 - Dropout	Taxa de dropout: 0,64
3 - Dense	Número de unidades: 1 Função de ativação: sigmoid

Fonte: Autoria própria.

Foram realizados ajustes nos hiperparâmetros do modelo, incluindo uma taxa de aprendizado (`learning_rate`) de 0,16 no otimizador, um `batch_size` de 1 para processamento individual de exemplos de treinamento, um treinamento de 400 épocas para permitir que o modelo se ajustasse aos dados, e um parâmetro `patience` de 17 no EarlyStopping para controlar o treinamento com base na melhoria da perda no conjunto de

validação. Essas configurações otimizaram o desempenho do modelo (Tabela 8), levando em consideração a natureza do problema e o conjunto de dados.

Tabela 8 – Resultados da classificação em porcentagem (média e desvio padrão) para o algoritmo BiLSTM.

Pré-processamento	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
Amida I	70,0 \pm 6,3%	70,1 \pm 5,4%	69,5 \pm 10,5%	70,5 \pm 7,6%	68,8 \pm 6,9%
Savitzky-Golay	71,5 \pm 5,0%	73,5 \pm 4,9%	67,5 \pm 11,9%	75,5 \pm 6,1%	70,4 \pm 5,4%

Fonte: Autoria própria.

Considerando a aplicação da normalização pela Amida I, o modelo registrou uma acurácia média de 70% (com um desvio padrão de \pm 6,3%), uma precisão média de 70,1% (\pm 5,4%), uma sensibilidade média de 69,5% (\pm 10,5%), uma especificidade média de 70,5% (\pm 7,6%), e um F1-score médio de 68,8% (\pm 6,9%). Notavelmente, a sensibilidade foi a métrica com a maior variabilidade, enquanto o F1-score mostrou o desempenho mais desafiador entre todas as métricas avaliadas.

Por outro lado, a aplicação do pré-processamento *Savitzky-Golay* com a utilização da primeira derivada resultou em melhorias significativas nos resultados, que atingiu 71,5% (\pm 5%) de acurácia, 73,5% (\pm 4,9%) de precisão, 67,5% (\pm 11,9%) de sensibilidade, 75,5% (\pm 6,1%) de especificidade e um F1-score de 70,4% (\pm 5,4%). É importante destacar que sob essa técnica de pré-processamento, o modelo apresentou o melhor desempenho em comparação com todas as outras abordagens avaliadas.

4.7 Análise comparativa

Os dados na Tabela 9 fornecem uma visão geral do desempenho de vários algoritmos de classificação sob o cenário de normalização pela Amida I. Cada linha da tabela representa um algoritmo específico, e as colunas apresentam métricas de avaliação de desempenho com médias e desvios padrão associados.

O algoritmo XGBoost se destacou notavelmente em relação aos outros algoritmos em várias métricas. Possui a maior acurácia, precisão e F1-Score, todos superando 74%, indicando sua capacidade de fazer previsões precisas. Além disso, tem uma alta sensibilidade de 77%, o que significa que é eficaz em identificar a hipersensibilidade dentinária. A especificidade de 72,5% mostra sua capacidade de distinguir a ausência da condição. No geral, apresentou um desempenho consistente, como evidenciado pelos desvios padrão em todas as métricas.

Tabela 9 – Resultados da classificação em porcentagem (média e desvio padrão), utilizando a normalização pela Amida I.

	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
KNN	48,7 \pm 5,6%	48,8 \pm 6,5%	42,5 \pm 5,1%	55,0 \pm 7,7%	47,7 \pm 5,5%
AD	57,0 \pm 7,7%	56,5 \pm 7,4%	61,5 \pm 9,5%	52,5 \pm 10,1%	56,0 \pm 8,0%
Naive Bayes	53,0 \pm 3,1%	53,6 \pm 3,6%	46,0 \pm 4,9%	60,0 \pm 5,5%	51,8 \pm 3,3%
XGBoost	74,7 \pm 6,4%	74,3 \pm 7,9%	77,0 \pm 5,1%	72,5 \pm 10,8%	74,2 \pm 7,9%
MLP	58,5 \pm 3,6%	59,7 \pm 4,7%	52,5 \pm 5,6%	64,5 \pm 4,7%	57,6 \pm 3,6%
BiLSTM	70,0 \pm 6,3%	70,1 \pm 5,4%	69,5 \pm 10,5%	70,5 \pm 7,6%	68,8 \pm 6,9%

Fonte: Autoria própria.

Por outro lado, algoritmos como KNN, Árvore de Decisão (AD), e Naive Bayes tiveram desempenho inferior em todas as métricas em comparação ao XGBoost. O KNN apresentou a menor acurácia, precisão, sensibilidade e F1-Score, indicando que é o menos eficaz na classificação. A maior variabilidade nas métricas de AD sugere que possa ser mais sensível às variações dos dados. No entanto, o Naive Bayes apresentou um desempenho intermediário, com destaque para sua alta especificidade.

Os algoritmos de redes neurais, MLP (Multi-Layer Perceptron) e BiLSTM (Bidirectional Long Short-Term Memory), tiveram um desempenho intermediário em relação aos outros algoritmos, considerando a normalização pela Amida I.

O MLP demonstrou resultados moderados em várias métricas. Possui uma acurácia e precisão razoáveis, superando os 58%. No entanto, seu desempenho em relação à sensibilidade foi relativamente baixo, com apenas 52,5%, o que implica que pode não ser tão eficaz na identificação de verdadeiros positivos. O desvio padrão nas métricas do MLP foi relativamente baixo, sugerindo que é consistente, mas seus resultados não são tão promissores quanto os do algoritmo XGBoost.

Por outro lado, o BiLSTM apresentou um desempenho relativamente melhor do que o MLP em todas as métricas. No entanto, o desvio padrão em suas métricas foi mais alto do que o MLP. Embora essa variabilidade seja mais pronunciada, o BiLSTM manteve uma consistência notável em seu desempenho.

No geral, tanto o MLP quanto o BiLSTM tiveram desempenho intermediário em comparação com o XGBoost. Eles podem ser opções viáveis, dependendo dos requisitos específicos do problema, mas podem exigir ajustes adicionais de hiperparâmetros ou considerações sobre o pré-processamento de dados para melhorar seu desempenho.

Em resumo, a tabela destaca o XGBoost como o algoritmo mais promissor sob

as condições de normalização pela Amida I, devido ao seu desempenho geral robusto e consistente em todas as métricas.

Tabela 10 – Resultados dos modelos de classificação em porcentagem (média e desvio padrão), considerando a utilização do pré-processamento *Savitzky-Golay* e destacando os resultados mais promissores entre a aplicação da primeira, segunda e terceira derivadas.

	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
KNN	54,7 ±4,8%	54,9 ±5,0%	53,5 ±6,3%	56,0 ±6,6%	54,4 ±4,7%
AD	55,5 ±6,0%	55,4 ±5,5%	54,0 ±10,9%	57,0 ±5,6%	54,8 ±5,7%
Naive Bayes	50,8 ±4,3%	50,3 ±4,1%	43,5 ±4,8%	58,0 ±4,6%	49,1 ±3,9%
XGBoost	68,0 ±6,4%	66,3 ±6,1%	74,5 ±7,2%	61,5 ±10,3%	66,7 ±7,5%
MLP	60,7 ±2,8%	60,3 ±2,6%	63,0 ±4,6%	58,5 ±3,9%	60,5 ±2,7%
BiLSTM	71,5 ±5,0%	73,5 ±4,9%	67,5 ±11,9%	75,5 ±6,1%	70,4 ±5,4%

Fonte: Autoria própria.

Ao comparar os resultados dos modelos de classificação com o pré-processamento *Savitzky-Golay* (Tabela 10) em relação aos resultados anteriores com a normalização pela Amida I, observamos algumas tendências interessantes.

Houve um aumento geral no desempenho do KNN após a aplicação do método *Savitzky-Golay*. O desempenho da Árvore de Decisão permaneceu relativamente estável, com algumas variações. A sensibilidade diminuiu, mas a especificidade aumentou, tornando-o um algoritmo mais equilibrado em termos de *trade-off* entre sensibilidade e especificidade.

Embora os resultados do Naive Bayes não sejam particularmente altos em nenhum dos dois métodos de pré-processamento, quando o pré-processamento *Savitzky-Golay* foi aplicado, a acurácia caiu para 50,8%, e a sensibilidade diminuiu ainda mais, chegando a 43,5%. A queda na sensibilidade após o pré-processamento *Savitzky-Golay* sugere que essa técnica pode ter tido um impacto negativo nas previsões do Naive Bayes para este conjunto de dados específico.

O XGBoost, que já era uma opção promissora, manteve seu desempenho sólido. A sensibilidade permaneceu alta e consistente, tornando-o uma escolha robusta.

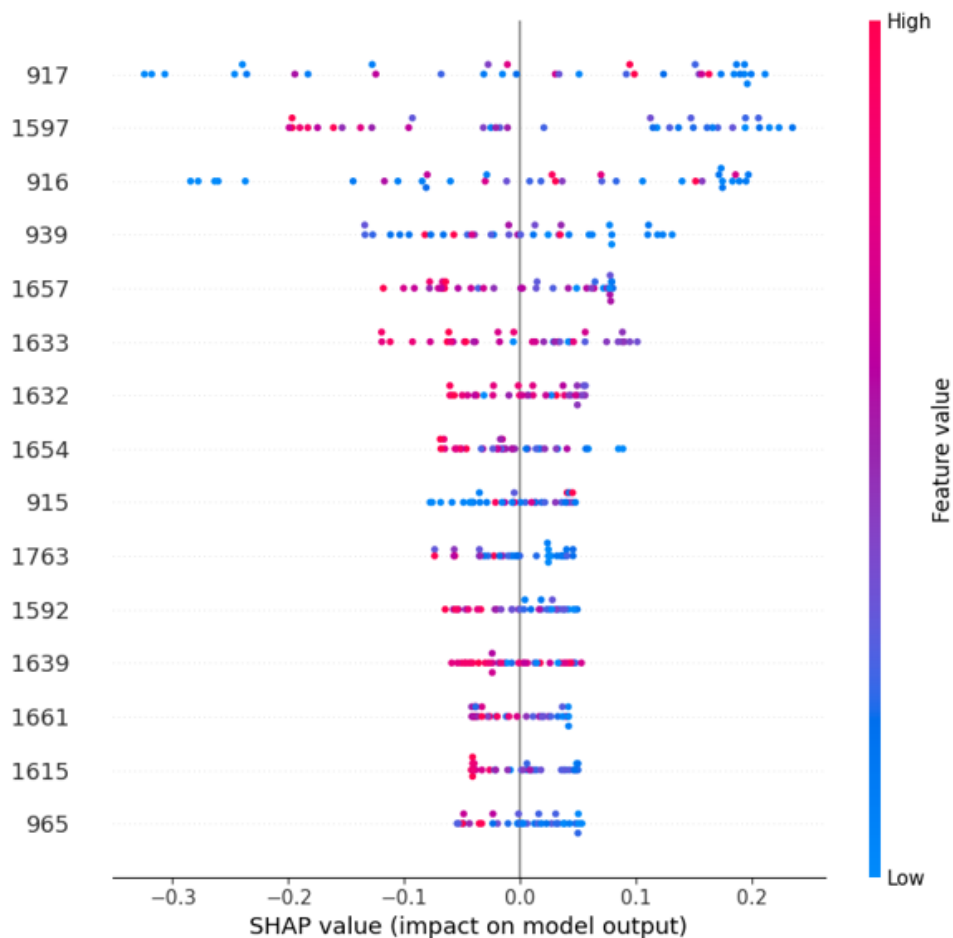
O MLP apresentou melhorias em quase todas as métricas, tornando-o mais equilibrado em termos de acurácia, precisão, sensibilidade e especificidade. A sensibilidade melhorou significativamente em relação ao desempenho anterior.

O BiLSTM foi o grande destaque após a aplicação do *Savitzky-Golay*. Ele alcan-

çou os melhores resultados em termos de acurácia, precisão, especificidade e F1-Score, tornando-o o algoritmo mais promissor e indicando sua capacidade de identificar verdadeiros positivos e verdadeiros negativos.

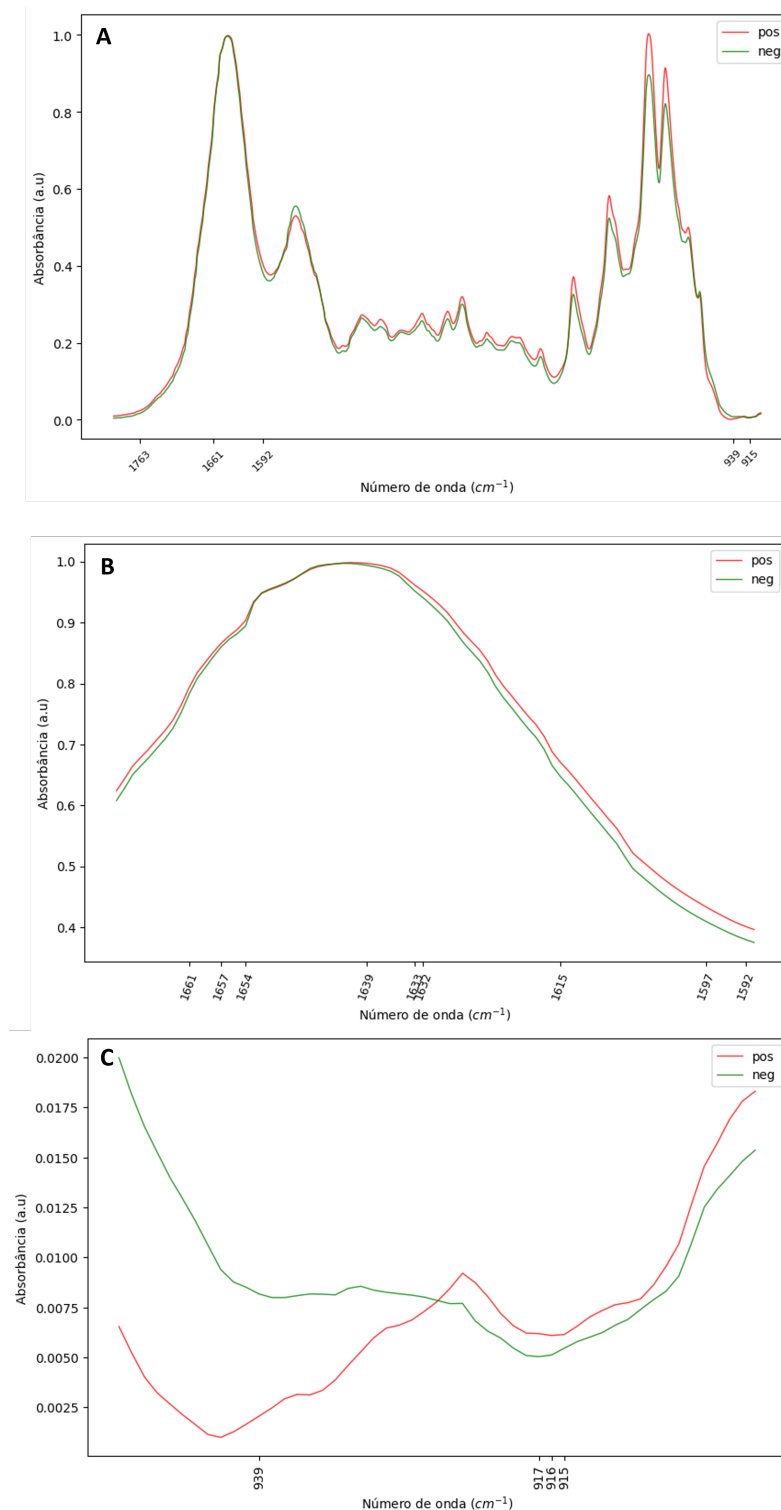
4.8 Explicabilidade utilizando a técnica SHAP

Considerando que o XGBoost apresentou melhores resultados com a normalização de Amida I e o BiLSTM com o pré-processamento *Savitzky-Golay*, a aplicação do SHAP (SHapley Additive exPlanations) se mostrou crucial para entender a contribuição de cada número de onda nos modelos de previsão. O SHAP é uma técnica de explicabilidade amplamente utilizada em aprendizado de máquina e interpretação de modelos, sobretudo em modelos complexos de inteligência artificial, com o propósito de fornecer explicações de maneira mais acessível e intuitiva. O SHAP faz parte do campo de pesquisa conhecido como ‘Inteligência Artificial Explicável’ ou XAI (*Explainable Artificial Intelligence*), que engloba várias técnicas e métodos destinados a tornar os modelos de inteligência artificial mais compreensíveis e interpretáveis para os seres humanos (ALI et al., 2023).



Fonte: Autoria própria.

Figura 4 – Explicabilidade do modelo de XGBoost utilizando a técnica SHAP.



Fonte: Autoria própria.

Figura 5 – Características que mais impactaram para a tomada de decisão do modelo de XGBoost.

A Figura 4 oferece uma visualização das contribuições das características em um modelo de XGBoost, onde as *features* (números de onda entre 1800 cm^{-1} e 900 cm^{-1}) são representadas no eixo vertical e as contribuições SHAP no eixo horizontal. Conforme

as características se movem para a direita no gráfico, suas contribuições às previsões são mais positivas, enquanto movimentos para a esquerda indicam contribuições negativas. A relação entre os valores das características e suas contribuições é codificada por cores, facilitando a compreensão dessa interação. Além disso, o gráfico fornece percepções sobre a distribuição das contribuições, permitindo uma análise da variabilidade presente nas previsões do modelo.

As características que desempenharam um papel mais significativo são destacadas na Figura 5. Os números de onda mais significativos, em ordem de importância, foram os seguintes: 917, 1597, 916, 939, 1657, 1633, 1632, 1654, 915, 1763, 1592, 1639, 1661, 1615, 965. Eles representam as frequências nas quais o espectro FTIR demonstrou contribuições mais relevantes para as previsões do modelo. O modelo considerou importante a faixa de frequência em que ocorre uma vibração de estiramento associada às ligações químicas C=O (carbono-oxigênio), típica em grupos de pectina. Esses valores são corroborados pelo trabalho de Souza (2018), que sugere uma relação com o mecanismo de proteção da degradação da dentina e sua sintomatologia.

5 Discussão

O uso da inteligência artificial em conjunto com a espectroscopia no infravermelho com transformada de Fourier (FTIR) na área da saúde é uma abordagem altamente eficaz para a análise e diagnóstico de diversas condições médicas. Essa poderosa combinação, quando aplicada à análise de amostras biológicas, como saliva, oferece uma abordagem não invasiva e sensível para a detecção precoce de uma ampla variedade de doenças, incluindo câncer, diabetes, doenças autoimunes e muito mais (SOUZA, 2018). Além disso, possibilita a personalização de tratamentos com base nas características moleculares individuais dos pacientes (ZLOTOGORSKI-HURVITZ et al., 2019). Os resultados deste trabalho corroboram a importância do uso da tecnologia para resolver problemas complexos e auxiliar no diagnóstico, destacando-se o bom desempenho dos algoritmos XGBoost e BiLSTM.

O algoritmo de XGBoost superou a BiLSTM na classificação da hipersensibilidade dentinária com base na normalização pela Amida I, sendo uma técnica de aprendizado de máquina altamente adequada para problemas de classificação binária e podendo produzir resultados sólidos com conjuntos de dados menores. Enquanto a BiLSTM geralmente requer volumes significativos de dados para atingir seu potencial máximo. Além disso, a configuração de hiperparâmetros na BiLSTM pode ser mais complexa, e erros na configuração podem afetar negativamente seu desempenho. O algoritmo de XGBoost foi capaz de produzir resultados sólidos com recursos computacionais mais modestos, como já explicitado no trabalho de Chen e Guestrin (2016).

A normalização *Savitzky-Golay* pode ajudar a realçar informações importantes em dados espectrais, tornando-os mais adequados para a análise por redes neurais recorrentes, como citado por Zhao et al. (2014) e que pode explicar a melhoria nos resultados dos modelos de KNN, MLP e BiLSTM.

A capacidade da BiLSTM de considerar o contexto bidirecional a torna apta a capturar dependências complexas e relações entre as características dos espectros FTIR, resultando em uma melhoria na precisão da classificação, como afirmado por (SALEHI-NEJAD et al., 2017). Isso é particularmente valioso em aplicações de diagnóstico e classificação na área da saúde, onde a identificação precisa de padrões nos espectros FTIR pode ter implicações significativas para o tratamento e a detecção precoce de doenças (WANG; TIAN; QIU, 2019). Com a aplicação do pré-processamento *Savitzky-Golay*, o modelo de BiLSTM obteve os melhores resultados em termos de acurácia, precisão, sensibilidade, especificidade e F1-Score, destacando-o como o algoritmo mais promissor.

Através dos resultados obtidos, foi possível constatar que alguns algoritmos po-

dem ser mais sensíveis a variações nos dados do que outros. Portanto, uma técnica de pré-processamento que funcione bem para um algoritmo pode não ser a melhor escolha para outro. Além da técnica de pré-processamento em si, a escolha dos hiperparâmetros do modelo (por exemplo, os parâmetros do GridSearchCV) também pode influenciar os resultados (YU et al., 2020). A combinação de uma técnica de pré-processamento com hiperparâmetros otimizados pode resultar em um desempenho aprimorado.

Uma taxa de aprendizado alta pode acelerar o processo de treinamento, mas também pode tornar o modelo instável e propenso a convergir rapidamente para mínimos locais em vez de alcançar o mínimo global da função de perda (YU et al., 2020). Isso pode resultar em overfitting e desempenho insatisfatório. A taxa de aprendizado fixa de 0,16, utilizada no algoritmo BiLSTM, é relativamente alta em comparação com valores típicos, que geralmente estão na faixa de 0,1 ou menores (ZLOTOGORSKI-HURVITZ et al., 2019; KISER et al., 2023). Embora a taxa de aprendizado ideal possa variar dependendo do conjunto de dados e da arquitetura da rede, como no trabalho de (WANG et al., 2022), que utilizou o valor de 0,3 com bons resultados.

A utilização da técnica de ‘Early Stopping’, neste trabalho, permitiu que o treinamento fosse interrompido antes que essas oscilações provocassem *overfitting*. Portanto, a combinação de uma taxa de aprendizado mais alta com o ‘Early Stopping’ pode ser uma estratégia eficaz para treinar modelos de forma mais eficiente. As taxas de aprendizagem adaptativas poderão ser consideradas em trabalhos futuros, visto que é razoável definir taxas de aprendizagem diferentes para cada parâmetro e a adaptação automática ao longo do processo de aprendizagem (YU et al., 2020).

Em seu estudo, Zlotogorski-Hurvitz et al. (2019) conseguiram detectar mudanças sutis nas conformações de proteínas, lipídios e ácidos nucleicos utilizando redes neurais artificiais otimizadas com pequenos conjuntos de dados de FTIR. Essa pesquisa foi conduzida com base em 40 amostras coletadas do estudo de Souza (2018) e apresentou bons resultados. Através da técnica de explicabilidade SHAP, foi possível avaliar que o modelo mais eficiente destacou a importância da faixa de frequência associada à vibração de estiramento dos grupos de pectina. Essa descoberta corroborou a sugestão de uma possível relação entre essa faixa de frequência e o mecanismo de proteção contra a degradação da dentina, bem como sua sintomatologia, demonstrando ser uma região de impacto no desempenho dos modelos.

A hipersensibilidade dentinária tem um impacto significativo na qualidade de vida dos indivíduos, e seu diagnóstico é frequentemente um desafio devido à necessidade de métodos não invasivos e livres de dor (LIMA et al., 2016). Os resultados obtidos neste estudo demonstram a viabilidade de modelos de diagnóstico da hipersensibilidade dentinária, oferecendo uma valiosa contribuição ao processo diagnóstico. Esses modelos se revelam aliados promissores nos esforços para melhorar a precisão e a acessibilidade dos

diagnósticos, proporcionando uma abordagem mais simples e segura para identificar e tratar essa condição.

6 Conclusão

Neste estudo, foram desenvolvidos modelos de aprendizado supervisionado para o diagnóstico da hipersensibilidade dentinária utilizando dados de Espectroscopia no Infravermelho por Transformada de Fourier (FTIR). A técnica de normalização pela Amida I e o pré-processamento *Savitzky-Golay* foram escolhidos por serem amplamente utilizados na literatura, considerando o problema abordado, e impactaram nos resultados. Enquanto a normalização pela Amida I se mostrou mais eficiente para o algoritmo XGBoost, o pré-processamento *Savitzky-Golay* proporcionou melhorias expressivas em outros modelos, como KNN, MLP, e, sobretudo, no BiLSTM.

Os modelos que mais se destacaram foram aqueles implementados por meio das técnicas XGBoost e BiLSTM. O algoritmo XGBoost produziu resultados sólidos, com uma configuração de hiperparâmetros mais simplificada, enquanto a técnica SHAP possibilitou o entendimento das contribuições mais relevantes para as previsões do modelo. O número de amostras e configurações de hiperparâmetros representam limitações importantes que podem ter restringido o potencial dos modelos, principalmente os de redes neurais. No entanto, mesmo diante dessas limitações, o modelo BiLSTM conseguiu um bom desempenho.

O desempenho positivo alcançado pelos modelos, apesar das limitações, sugere um potencial significativo para a aplicação dessas abordagens no diagnóstico de hipersensibilidade dentinária, destacando-se a efetividade das técnicas XGBoost e BiLSTM. Para futuras investigações, é importante estendermos o número de amostras da base de dados, além de considerar métodos de correção de *baseline*, como *rubberband* e *polynomial*, bem como explorar novas técnicas de aprendizado profundo tais como redes neurais convolucionais.

Referências

- ALI, S.; ABUHMED, T.; EL-SAPPAGH, S.; MUHAMMAD, K.; ALONSO-MORAL, J. M.; CONFALONIERI, R.; GUIDOTTI, R.; Del Ser, J.; DÍAZ-RODRÍGUEZ, N.; HERRERA, F. Explainable artificial intelligence (xai): What we know and what is left to attain trustworthy artificial intelligence. **Information Fusion**, v. 99, p. 101805, 2023. ISSN 1566-2535. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1566253523001148>>. Citado na página 38.
- BRITO, R. de; OLIVEIRA, A. de. Análise de características a partir do classificador mlp para auxílio no diagnóstico da covid-19. SBC, Porto Alegre, RS, Brasil, p. 89–95, 2020. ISSN 0000-0000. Disponível em: <<https://sol.sbc.org.br/index.php/ercemapi/article/view/11472>>. Citado 2 vezes nas páginas 23 e 27.
- CAIXETA, D. C.; CARNEIRO, M. G.; RODRIGUES, R.; ALVES, D. C. T.; GOULART, L. R.; CUNHA, T. M.; ESPINDOLA, F. S.; VITORINO, R.; SABINO-SILVA, R. Salivary atr-ftir spectroscopy coupled with support vector machine classification for screening of type 2 diabetes mellitus. **Diagnostics**, v. 13, n. 8, 2023. ISSN 2075-4418. Disponível em: <<https://www.mdpi.com/2075-4418/13/8/1396>>. Citado na página 23.
- CHATTOPADHYAY, S.; DAVIS, R. M.; MENEZES, D. D.; SINGH, G.; ACHARYA, R. U.; TAMURA, T. Application of bayesian classifier for the diagnosis of dental pain. **J Med Syst**, United States, v. 36, n. 3, p. 1425–1439, out. 2010. Citado na página 22.
- CHEN, T.; GUESTRIN, C. Xgboost: A scalable tree boosting system. In: . [S.l.: s.n.], 2016. p. 785–794. Citado 2 vezes nas páginas 18 e 41.
- CHIBEBE, P. C.; TERRERI, M.; RICARDO, L. H.; PALLOS, D. Uma visão atual do fluido gengival crevicular como método de diagnóstico periodontal. **Revista de Ciências Médicas**, v. 17, n. 3/6, ago. 2012. Disponível em: <<https://seer.sis.puc-campinas.edu.br/cienciasmedicas/article/view/754>>. Citado na página 14.
- CHOI, R. Y.; COYNER, A. S.; KALPATHY-CRAMER, J.; CHIANG, M. F.; CAMPBELL, J. P. Introduction to Machine Learning, Neural Networks, and Deep Learning. **Translational Vision Science Technology**, v. 9, n. 2, p. 14–14, 02 2020. ISSN 2164-2591. Disponível em: <<https://doi.org/10.1167/tvst.9.2.14>>. Citado na página 18.
- DAVARI, A.; ATAELI, E.; ASSARZADEH, H. Dentin hypersensitivity: etiology, diagnosis and treatment; a literature review. **J Dent (Shiraz)**, v. 14(3), p. 136–145, 2013. Citado na página 14.
- DEVITO, K. L.; BARBOSA, F. de S.; FILHO, W. N. F. An artificial multilayer perceptron neural network for diagnosis of proximal dental caries. **Oral Surg Oral Med Oral Pathol Oral Radiol Endod**, United States, v. 106, n. 6, p. 879–884, ago. 2008. Citado na página 22.

- DUKITHA, R. S. M. Classification techniques in data mining. **International Journal of Computer Sciences and Engineering**, IJCSE, Indore, INDIA, v. 07, p. 25–28, 5 2019. ISSN 2347-2693. Disponível em: <https://www.ijcseonline.org/full_spl_paper_view.php?paper_id=1303>. Citado na página 16.
- FILHO, A. M.; FERNANDES, J.; SABINO-SILVA, R.; CARNEIRO, M. Oceanspectra: an oral cancer detection system from salivary atr-ftir spectroscopy. In: **Anais do XX Encontro Nacional de Inteligência Artificial e Computacional**. Porto Alegre, RS, Brasil: SBC, 2023. p. 984–996. ISSN 2763-9061. Disponível em: <<https://sol.sbc.org.br/index.php/eniac/article/view/25758>>. Citado na página 22.
- FILHO, R. L.; CARNEIRO, M. Diagnóstico do câncer oral através da classificação de alto nível. In: **Anais Estendidos do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde**. Porto Alegre, RS, Brasil: SBC, 2023. p. 54–59. ISSN 2763-8987. Disponível em: <https://sol.sbc.org.br/index.php/sbcas_estendido/article/view/25331>. Citado na página 22.
- GUPTA, G. Gingival crevicular fluid as a periodontal diagnostic indicator–i: Host derived enzymes and tissue breakdown products. **J Med Life**, Romania, v. 5, n. 4, p. 390–397, dez. 2012. Citado na página 15.
- INDRAJA, B.; KUMARAPPAN, A. Classification of medicines using naive bayes classifier. **Research Journal of Pharmacy and Technology**, v. 11, p. 1940–1944, 05 2018. Citado na página 17.
- KISER, A. C. et al. Developing an lstm model to identify surgical site infections using electronic healthcare records. **AMIA Joint Summits on Translational Science proceedings**, AMIA Joint Summits on Translational Science, v. 2023, p. 330–339, June 16 2023. Citado na página 42.
- KONG, Z.; CUI, Y.; XIA, Z.; LV, H. Convolution and long short-term memory hybrid deep neural networks for remaining useful life prognostics. **Applied Sciences**, v. 9, n. 19, 2019. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/9/19/4156>>. Citado 2 vezes nas páginas 19 e 20.
- KUMAR, J.; GOOMER, R.; SINGH, A. K. Long short term memory recurrent neural network (lstm-rnn) based workload forecasting model for cloud datacenters. **Procedia Computer Science**, v. 125, p. 676–682, 2018. ISSN 1877-0509. The 6th International Conference on Smart Computing and Communications. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050917328557>>. Citado na página 19.
- LEAL, L.; NOGUEIRA, M.; CANEVARI, R.; CARVALHO, L. Vibration spectroscopy and body biofluids: Literature review for clinical applications. **Photodiagnosis and Photodynamic Therapy**, v. 24, p. 237–244, 2018. ISSN 1572-1000. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S157210001830036X>>. Citado 2 vezes nas páginas 15 e 24.
- LIMA, J. B.; NASCIMENTO, A. R. F.; NASCIMENTO, V. R. P.; OLIVEIRA, I. L. M. Hipersensibilidade dentinária: etiologia, diagnóstico e tratamento. **Odontol. Clín.-Cient**, v. 20, n. 2, p. 46–51, 2021. Citado 2 vezes nas páginas 11 e 14.

- LIMA, T. C.; VIEIRA-BARBOSA, N. M.; AZEVEDO, C. Grasielle de S.; MATOS, F. R. de; OLIVEIRA, D. W. Douglas de; OLIVEIRA, E. S. de; RAMOS-JORGE, M. L.; GONÇALVES, P. F.; FLECHA, O. D. Oral Health-Related quality of life before and after treatment of dentin hypersensitivity with cyanoacrylate and laser. **J Periodontol**, United States, v. 88, n. 2, p. 166–172, out. 2016. Citado 2 vezes nas páginas 12 e 42.
- LIPTON, Z. C. A critical review of recurrent neural networks for sequence learning. **ArXiv**, abs/1506.00019, 2015. Citado na página 20.
- ROKACH, L.; MAIMON, O. Decision trees. In: _____. **Data Mining and Knowledge Discovery Handbook**. Boston, MA: Springer US, 2005. p. 165–192. ISBN 978-0-387-25465-4. Disponível em: <https://doi.org/10.1007/0-387-25465-X_9>. Citado na página 17.
- SAKUDO, A. Near-infrared spectroscopy for medical applications: Current status and future perspectives. **Clinica Chimica Acta**, v. 455, p. 181–188, 2016. ISSN 0009-8981. Citado 2 vezes nas páginas 11 e 16.
- SALEHINEJAD, H.; SANKAR, S.; BARFETT, J.; COLAK, E.; VALAEE, S. Recent advances in recurrent neural networks. 12 2017. Citado 2 vezes nas páginas 18 e 41.
- SAMPAIO, I.; BERNARDINI, F.; PAES, A.; ANDRADE, E.; VITERBO, J. Avaliação de modelos de predição e previsão construídos por algoritmos de aprendizado de máquina em problemas de cidades inteligentes. In: _____. [S.l.: s.n.], 2019. p. 81–113. ISBN 9788576694809. Citado na página 21.
- SHARFUDDIN, A.; TIHAM, N.; ISLAM, M. S. A deep recurrent neural network with bilstm model for sentiment classification. In: . [S.l.: s.n.], 2018. p. 1–4. Citado na página 20.
- SHARIFI, A.; ALIZADEH, K. A novel classification method based on multilayer perceptron-artificial neural network technique for diagnosis of chronic kidney disease. **Annals of Military and Health Sciences Research**, v. 18, n. 1, p. e101585, 2020. Disponível em: <<https://brieflands.com/articles/amhsr-101585.html>>. Citado 2 vezes nas páginas 19 e 23.
- SICHMAN, J. S. Inteligência artificial e sociedade: avanços e riscos. **Estudos Avançados**, v. 35, n. 101, p. 37–50, 2021. Citado na página 11.
- SOARES, D. G.; TEIVE, R. Estudo comparativo entre as redes neurais artificiais mlp e rbf para previsão de cheias em curto prazo. **Revista de Informática Teórica e Aplicada**, v. 22, 08 2015. Citado na página 19.
- SOUZA, P. G. **Análise da composição química do fluido crevicular gengival em dentes com hipersensibilidade dentinária através das espectroscopias ATR-FTIR e Raman**. Dissertação (Dissertação de Mestrado) — Universidade Federal de Uberlândia, Uberlândia, 2018. Citado 7 vezes nas páginas 11, 15, 21, 24, 40, 41 e 42.
- SUBBARAO, K. C.; NATTUTHURAI, G. S.; SUNDARARAJAN, S. K.; SUJITH, I.; JOSEPH, J.; SYEDSHAH, Y. P. Gingival crevicular fluid: An overview. **J Pharm Bioallied Sci**, India, v. 11, n. Suppl 2, p. S135–S139, maio 2019. Citado na página 15.

VIKRAMKUMAR; B, V.; TRILOCHAN. Bayes and naive bayes classifier. **CoRR**, abs/1404.0933, 2014. Disponível em: <<http://arxiv.org/abs/1404.0933>>. Citado na página 17.

VILONE, G.; LONGO, L. Notions of explainability and evaluation approaches for explainable artificial intelligence. **Information Fusion**, v. 76, 05 2021. Citado na página 12.

WANG, H.; REHMETULLA, A.; GUO, S.; KONG, X.; Lü, Z.; GUAN, Y.; XU, C.; SULAIMAN, K.; WEI, G.; LIU, H. Machine learning based on structural and ftir spectroscopic datasets for seed autotclassification. **RSC Adv**, v. 12, n. 18, p. 11413–11419, Apr 12 2022. Citado na página 42.

WANG, T.; TIAN, Y.; QIU, R. G. Long Short-Term memory recurrent neural networks for multiple diseases risk prediction by leveraging longitudinal medical records. **IEEE J Biomed Health Inform**, United States, v. 24, n. 8, p. 2337–2346, dez. 2019. Citado 2 vezes nas páginas 22 e 41.

WIJAYA, E. Implementation analysis of glm and naive bayes methods in conducting extractions on dental image. **IOP Conference Series: Materials Science and Engineering**, v. 407, p. 012146, 09 2018. Citado na página 22.

XING, L.; ZHANG, X.; GUO, Y.; BAI, D.; XU, H. Xgboost-aided prediction of lip prominence based on hard-tissue measurements and demographic characteristics in an asian population. **Am J Orthod Dentofacial Orthop**, v. 164, n. 3, p. 357–367, 2023. Citado na página 18.

YU, C.; QI, X.; MA, H.; HE, X.; WANG, C.; ZHAO, Y. Llr: Learning learning rates by lstm for training neural networks. **Neurocomputing**, v. 394, 02 2020. Citado 2 vezes nas páginas 34 e 42.

ZHANG, Z. Introduction to machine learning: k-nearest neighbors. **Ann Transl Med**, China, v. 4, n. 11, p. 218, jun. 2016. Citado na página 16.

ZHAO, A.-X.; TANG, X.-J.; ZHANG, Z.-H.; LIU, J.-H. The parameters optimization selection of savitzky-golay filter and its application in smoothing pretreatment for ftir spectra. In: **2014 9th IEEE Conference on Industrial Electronics and Applications**. [S.l.: s.n.], 2014. p. 516–521. Citado 2 vezes nas páginas 24 e 41.

ZLOTOGORSKI-HURVITZ, A.; DEKEL, B. Z.; MALONEK, D.; YAHALOM, R.; VERED, M. FTIR-based spectrum of salivary exosomes coupled with computational-aided discriminating analysis in the diagnosis of oral cancer. **J Cancer Res Clin Oncol**, Germany, v. 145, n. 3, p. 685–694, jan 2019. Citado 3 vezes nas páginas 22, 41 e 42.