

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

Paulo D. S. Souza<sup>1</sup>, Paola G. Souza<sup>2</sup>, Paulo V. Soares<sup>2</sup>,  
Robinson Sabino-Silva<sup>3</sup>, Murillo G. Carneiro<sup>1</sup>

<sup>1</sup>Faculdade de Computação  
Universidade Federal de Uberlândia, Uberlândia, MG, Brasil

<sup>2</sup>Faculdade de Odontologia  
Universidade Federal de Uberlândia, Uberlândia, MG, Brasil

<sup>3</sup>Departamento de Fisiologia, Instituto de Ciências Biomédicas  
Universidade Federal de Uberlândia, Uberlândia, MG, Brasil

{pdodonto, paulovsoares, mgcarneiro}@ufu.br

{paolagomessouza, robinsonsabino}@gmail.com

**Abstract.** *In this study we developed an end-to-end pipeline to the diagnosis of dentin hypersensitivity. The pipeline has two major components: data collection and preparation, and machine learning. In the former, patient samples of gingival crevicular fluid underwent Fourier-transform infrared spectroscopy (FTIR) analysis, with preprocessing steps encompassing normalization via Amide I and application of Savitzky-Golay with derivatives. In the latter, several traditional and state-of-the-art machine learning models were developed and evaluated. Our results evidenced the feasibility of dentin hypersensitivity detection for XGBoost and BiLSTM, as well as the importance of proper data preparation and parameter selection in order to deal with the inherent challenges of FTIR data.*

**Resumo.** *Neste estudo desenvolvemos um pipeline de ponta a ponta para o diagnóstico da hipersensibilidade dentinária. O pipeline possui dois componentes principais: coleta e preparação de dados, e aprendizado de máquina. No primeiro, amostras do fluido crevicular gengival de pacientes foram submetidas à análise por espectroscopia de infravermelho por transformada de Fourier (FTIR), com etapas de pré-processamento incluindo normalização via Amida I e aplicação de Savitzky-Golay com derivadas. No segundo, diversos modelos de aprendizado de máquina foram desenvolvidos e avaliados. Nossos resultados evidenciaram a viabilidade dos modelos de XGBoost e BiLSTM e a importância da adequada seleção de parâmetros e preparação dos dados de FTIR.*

## 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. Essa condição 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. 2021].

O fluido crevicular gengival (FCG) é uma mistura complexa de biomarcadores celulares e moleculares, refletindo a interação entre biofilme bacteriano e tecido periodontal. Sua coleta é não invasiva e pode indicar doença periodontal, cicatrização e, recentemente, hipersensibilidade dentinária, com alterações químicas na composição, detectadas pela Espectroscopia Infravermelha com Transformada de Fourier (FTIR) [Souza 2018].

A espectroscopia FTIR é uma tecnologia verde empregada de forma eficaz na detecção de diversas doenças devido à sua capacidade de análise molecular precisa com preparação mínima de amostra e sustentável, destacando-se pela ausência do uso de reagentes químicos, o que contribui significativamente para a redução de resíduos e impactos ambientais na prática clínica [Caixeta et al. 2023]. É 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 doença periodontal [Pereira de Souza et al. 2023].

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].

O algoritmo XGBoost 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 [Chen and Guestrin 2016, Li et al. 2024].

Outro algoritmo notável é o da arquitetura BiLSTM (Bidirectional Long Short-Term Memory), que representa uma extensão da LSTM (Long Short-Term Memory), uma arquitetura especial de rede neural recorrente. 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 maior. São especialmente úteis nas ocasiões em que o contexto da entrada é necessário, muito útil em trabalhos de classificação de sentimentos [Sharfuddin et al. 2018].

A grande parte dos modelos criados 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 subjacentes. 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 and Longo 2021].

O objetivo deste trabalho é desenvolver modelos de aprendizado supervisionado do estado-da-arte para o diagnóstico da hipersensibilidade dentinária a partir de dados de FTIR. A hipótese investigada é que técnicas capazes de aprender representações mais robustas dos dados, tais como XGBoost e BiLSTM, podem superar o desempenho preditivo de outros algoritmos de classificação. Ademais, também é objetivo dessa pesquisa fornecer interpretações sobre os melhores modelos obtidos de modo que possam auxiliar os especialistas na compreensão das decisões geradas por eles.

O restante do artigo está dividido em cinco seções principais: A Seção 2 apresenta uma revisão dos trabalhos relacionados, abordando pesquisas relevantes para o contexto do estudo. A Seção 3 detalha os materiais e métodos utilizados, incluindo a base de dados, o pré-processamento e o desenvolvimento dos modelos de classificação. Na Seção 4 são apresentados os resultados obtidos, uma comparação detalhada dos modelos considerados e uma análise de interpretabilidade do modelo de melhor desempenho; seguidos por uma discussão de tais resultados na Seção 5. Por fim, a Seção 6 apresenta as conclusões do estudo, consolidando as principais descobertas, contribuições, limitações e implicações, bem como propostas para trabalhos futuros.

## **2. Trabalhos Relacionados**

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 no 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. 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 and 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 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].

[Li et al. 2024] desenvolveram um modelo diagnóstico para distinguir entre doença de Crohn e tuberculose intestinal. Amostras de tecido foram analisadas usando ATR-FTIR e XGBoost, resultando em um modelo com sensibilidade, especificidade e precisão acima de 90%. Os autores preveem que essa abordagem, combinando análise

espectral e aprendizado de máquina, possa se tornar um novo método para o diagnóstico diferencial dessas doenças.

Por outro lado, tais estudos evidenciam a necessidade de se investigar técnicas mais sofisticadas de aprendizado de máquina ante aquelas amplamente consideradas pela literatura (e.g., kNN, MLP, etc). Mais do que isso, os estudos também sofrem pela falta de análises de interpretação (ou explicação) dos modelos, os quais dificultam a sua validação e adoção prática. Neste trabalho, tais lacunas são cobertas tanto pela investigação de arquiteturas de aprendizado mais sofisticadas como LSTM e XGBoost, quanto pela análise de interpretabilidade dos melhores modelos obtidos considerando o método SHAP.

### 3. Materiais e Métodos

A seguir, são apresentados os materiais e métodos utilizados, incluindo a base de dados, o pré-processamento e o desenvolvimento dos modelos de classificação. Os dois principais componentes do *pipeline* desenvolvido para o diagnóstico de hipersensibilidade dentinária, a coleta e preparação de dados e o aprendizado de máquina, estão ilustrados na Figura 1. No primeiro, a Etapa A descreve a coleta de amostras do fluido crevicular gengival de pacientes utilizando tiras de papel absorvente, devidamente preparadas para análise por espectroscopia no infravermelho por transformada de Fourier (Etapa B). Na Etapa C, os dados obtidos passam por um pré-processamento que inclui métodos de normalização pela Amida I e suavização utilizando o filtro *Savitzky-Golay*. No segundo componente, a Etapa D ilustra os algoritmos selecionados para o trabalho, incluindo alguns clássicos e outros mais robustos, como XGBoost e BiLSTM. A Etapa E ilustra a seleção e otimização dos hiperparâmetros para cada algoritmo por meio da biblioteca GridSearchCV. Considerando as métricas de avaliação na Etapa F, os modelos são analisados e comparados com base no desempenho (Etapa G). Por fim, a etapa H ilustra a utilização da Inteligência Artificial explicável no modelo de maior desempenho.

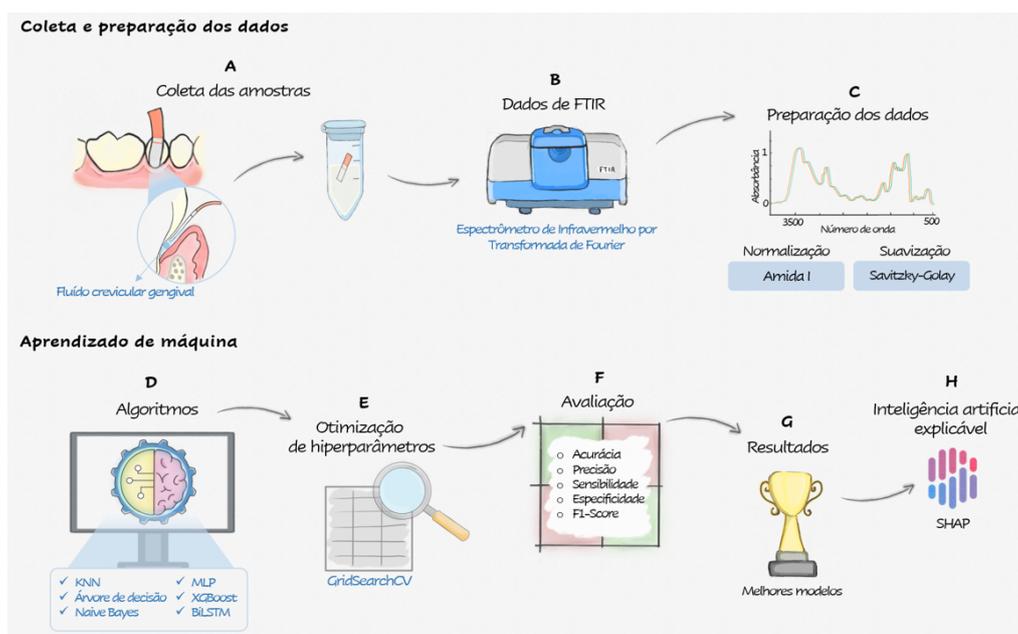
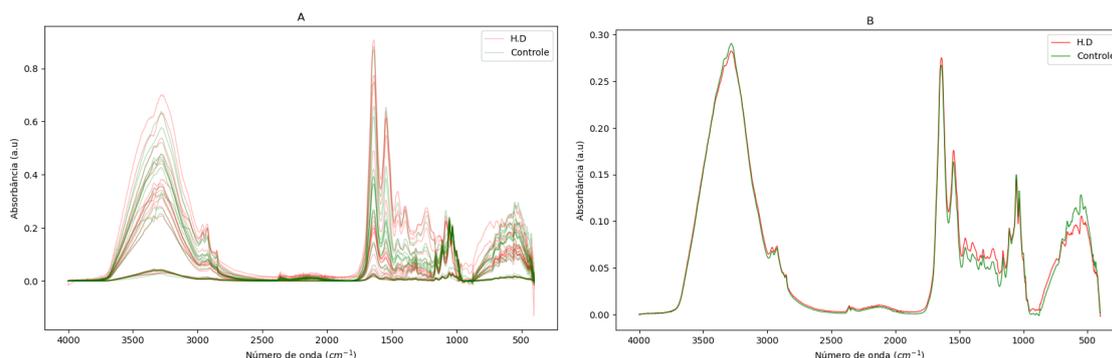


Figura 1. Componentes principais da solução fim-a-fim desenvolvida para o diagnóstico de hipersensibilidade dentinária.

### 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 (Figura 2). 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  $\mu m$  [Leal et al. 2018].



**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.**

### 3.2. Pré-processamento

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  $cm^{-1}$ ). Considerando 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 minimizar ruídos e outliers.

### 3.3. Desenvolvimento e avaliação 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 (k-vizinhos mais

próximos), 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.

O desempenho de um modelo está fortemente vinculado à configuração dos hiperparâmetros. A biblioteca GridSearchCV foi utilizada para automatizar o ajuste dos hiperparâmetros dos algoritmos KNN, Árvore de Decisão, Naive Bayes, MLP (Multi-Layer Perceptron) e XGBoost (Tabela 1). 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. A técnica ‘Early Stopping’ foi utilizada para monitorar o processo de treinamento e interrompê-lo quando determinados critérios fossem atendidos.

**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)
BiLSTM	activation units dropout_rate	relu, tanh {20, 21, ..., 200} {0, 0.1, ..., 1}

Foi utilizada a técnica de k-fold cross-validation, com valor de k igual a 10. O conjunto de dados foi dividido em 10 partições mutuamente 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 e minimizar limitações relacionadas à base de dados, foram realizadas 10 repetições, considerando o resultado médio e o desvio padrão de todas as execuções.

A análise e comparação de desempenho dos modelos foi realizada através das medidas de sensibilidade, especificidade, acurácia, precisão, análise da matriz de confusão e principalmente pela média harmônica entre sensibilidade e especificidade (F1-score).

#### 4. Resultados

A Tabela 2 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. Por outro lado, algoritmos como KNN, Árvore de Decisão (AD), Naive Bayes e MLP 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, provavelmente influenciado pela alta dimensionalidade dos dados. O algoritmo BiLSTM apresentou resultados intermediários, com especificidade próxima àquela obtida pelo XGBoost.

**Tabela 2. Resultados dos modelos de 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%
<b>XGBoost</b>	<b>74,7</b> $\pm$ 6,4%	<b>74,3</b> $\pm$ 7,9%	<b>77,0</b> $\pm$ 5,1%	<b>72,5</b> $\pm$ 10,8%	<b>74,2</b> $\pm$ 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%

A Tabela 3 apresenta os resultados obtidos pelas técnicas de classificação considerando o pré-processamento dos dados usando o método *Savitzky-Golay*. O BiLSTM foi o grande destaque aqui, com 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. Em relação aos resultados da Tabela 2 com a normalização pela Amida I, observamos alguns pontos interessantes. Houve um aumento geral no desempenho do KNN, MLP e BiLSTM após a aplicação do método *Savitzky-Golay*. Naive Bayes e XGBoost tiveram seu desempenho preditivo diminuído. Apesar da queda, o XGBoost obteve desempenho intermediário, consistente e com alta sensibilidade.

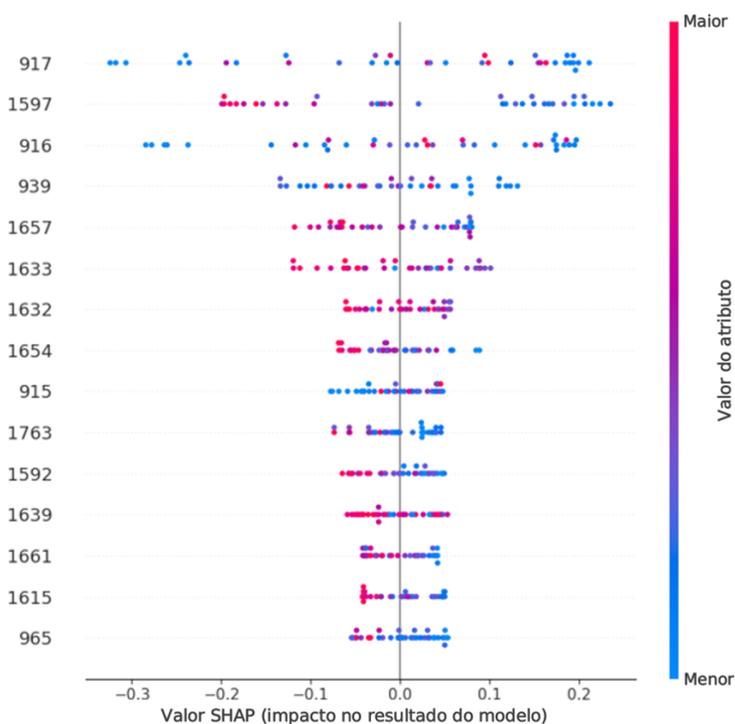
**Tabela 3. Resultados dos modelos de classificação em porcentagem (média e desvio padrão), utilizando o pré-processamento *Savitzky-Golay* e destacando os resultados mais promissores de derivação.**

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

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 banda infravermelho 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].

A Figura 3 oferece uma visualização das contribuições das características no modelo de XGBoost, onde as *features* (números de onda entre  $1800\text{ cm}^{-1}$  e  $900\text{ cm}^{-1}$ ) 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.



**Figura 3. Explicabilidade do modelo de XGBoost utilizando a técnica SHAP.**

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 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. A hipótese de pesquisa foi também confirmada, uma vez que modelos mais robustos de aprendizado, tais como XGBoost e BiLSTM, alcançaram os melhores desempenhos preditivos.

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 and 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 [Salehinejad 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 et al. 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 podem ser mais sensíveis aos ruídos e à alta dimensionalidade presente 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 técnicas de pré-processamento com hiperparâmetros otimizados pode contribuir em melhor desempenho.

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 e essa técnica 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 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. Esses modelos se revelam aliados promissores nos esforços para melhorar a precisão e a acessibilidade dos diagnósticos, proporcionando, quando incorporados à rotina clínica odontológica, uma abordagem mais simples, segura e de baixo custo 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 FTIR, uma plataforma sustentável que realiza análise baseada na luz infravermelha sem uso de reagentes químicos, tem liberação rápida de resultados e alta sensibilidade. 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 utilizando as 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 sugere um potencial significativo para a aplicação dessas abordagens na detecção 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.

## Agradecimentos

Os autores agradecem o apoio financeiro do Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq (processo n. 420212/2023-0), da Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG (processo APQ-00410-21), do INCT em Teranóstica e Nanobiotecnologia (processo n. CNPq-465669/2014-0), e do INCT em Saúde Oral e Odontologia (processo n. CNPq-406840/2022-9).

## Referências

- Ali, S., Abuhmed, T., El-Sappagh, S., Muhammad, K., Alonso-Moral, J. M., Confalonieri, R., Guidotti, R., Del Ser, J., Díaz-Rodríguez, N., and Herrera, F. (2023). Explainable artificial intelligence (XAI): What we know and what is left to attain trustworthy artificial intelligence. *Information Fusion*, 99:101805.
- Caixeta, D. C., Carneiro, M. G., Rodrigues, R., Alves, D. C. T., Goulart, L. R., Cunha, T. M., Espindola, F. S., Vitorino, R., and Sabino-Silva, R. (2023). Salivary ATR-FTIR spectroscopy coupled with support vector machine classification for screening of type 2 Diabetes Mellitus. *Diagnostics*, 13(8).
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. pages 785–794.
- Filho, A. C. M., Fernandes, J. M., Sabino-Silva, R., and Carneiro, M. G. (2023). OCANSpectra: an oral cancer detection system from salivary ATR-FTIR spectroscopy. In *Anais do XX Encontro Nacional de Inteligência Artificial e Computacional*, pages 984–996, Porto Alegre, RS, Brasil. SBC.
- Filho, R. B. L. and Carneiro, M. G. (2023). Diagnóstico do câncer oral através da classificação de alto nível. In *Anais Estendidos do XXIII Simpósio Brasileiro de Computação Aplicada à Saúde*, pages 54–59, Porto Alegre, RS, Brasil. SBC.
- Kiser, A. C. et al. (2023). Developing an LSTM model to identify surgical site infections using electronic healthcare records. *AMIA Joint Summits on Translational Science proceedings*, 2023:330–339.
- Leal, L., Nogueira, M., Canevari, R., and Carvalho, L. (2018). Vibration spectroscopy and body biofluids: Literature review for clinical applications. *Photodiagnosis and Photodynamic Therapy*, 24:237–244.

- Li, Y.-P., Lu, T.-Y., Huang, F.-R., Zhang, W.-M., Chen, Z.-Q., Guang, P.-W., Deng, L.-Y., and Yang, X.-H. (2024). Differential diagnosis of crohn's disease and intestinal tuberculosis based on ATR-FTIR spectroscopy combined with machine learning. *World J. Gastroenterol.*, 30(10):1377–1392.
- Lima, J. B., Nascimento, A. R. F., Nascimento, V. R. P., and Oliveira, I. L. M. (2021). Hipersensibilidade dentinária: etiologia, diagnóstico e tratamento. *Odontol. Clín.-Cient*, 20(2):46–51.
- Lima, T. C., Vieira-Barbosa, N. M., Grasielle de Sá Azevedo, C., de Matos, F. R., Douglas de Oliveira, D. W., de Oliveira, E. S., Ramos-Jorge, M. L., Gonçalves, P. F., and Flecha, O. D. (2016). Oral Health-Related quality of life before and after treatment of dentin hypersensitivity with cyanoacrylate and laser. *J Periodontol*, 88(2):166–172.
- Pereira de Souza, N. M., Machado, B. H., Padoin, L. V., Prá, D., Fay, A. P., Corbellini, V. A., and Rieger, A. (2023). Rapid and low-cost liquid biopsy with ATR-FTIR spectroscopy to discriminate the molecular subtypes of breast cancer. *Talanta*, 254(123858):123858.
- Salehinejad, H., Sankar, S., Barfett, J., Colak, E., and Valaee, S. (2017). Recent advances in recurrent neural networks.
- Sharfuddin, A., Tiham, N., and Islam, M. S. (2018). A deep recurrent neural network with BiLSTM model for sentiment classification. pages 1–4.
- Sichman, J. S. (2021). Inteligência artificial e sociedade: avanços e riscos. *Estudos Avançados*, 35(101):37–50.
- Souza, P. G. (2018). Análise da composição química do fluido crevicular gengival em dentes com hipersensibilidade dentinária através das espectroscopias ATR-FTIR e Raman. Dissertação de mestrado, Universidade Federal de Uberlândia, Uberlândia.
- Vilone, G. and Longo, L. (2021). Notions of explainability and evaluation approaches for explainable artificial intelligence. *Information Fusion*, 76.
- Wang, H., Rehmetulla, A., Guo, S., Kong, X., Lü, Z., Guan, Y., Xu, C., Sulaiman, K., Wei, G., and Liu, H. (2022). Machine learning based on structural and FTIR spectroscopic datasets for seed autotclassification. *RSC Adv*, 12(18):11413–11419.
- Wang, T., Tian, Y., and Qiu, R. G. (2019). Long Short-Term memory recurrent neural networks for multiple diseases risk prediction by leveraging longitudinal medical records. *IEEE J Biomed Health Inform*, 24(8):2337–2346.
- Yu, C., Qi, X., Ma, H., He, X., Wang, C., and Zhao, Y. (2020). LLR: Learning learning rates by LSTM for training neural networks. *Neurocomputing*, 394.
- Zhao, A.-X., Tang, X.-J., Zhang, Z.-H., and Liu, J.-H. (2014). 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*, pages 516–521.
- Zlotogorski-Hurvitz, A., Dekel, B. Z., Malonek, D., Yahalom, R., and Vered, M. (2019). FTIR-based spectrum of salivary exosomes coupled with computational-aided discriminating analysis in the diagnosis of oral cancer. *J Cancer Res Clin Oncol*, 145(3):685–694.