

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

Paulo Diego S. Souza<sup>1</sup>, Robinson S. Silva<sup>2</sup>, Murillo G. Carneiro<sup>1</sup>

<sup>1</sup>Faculdade de Computação – Universidade Federal de Uberlândia (UFU)  
Uberlândia – MG – Brasil

<sup>2</sup>Instituto de Ciências Biomédicas – Universidade Federal de Uberlândia (UFU)  
Uberlândia - MG- Brasil

{pdodonto,mgcarneiro}@ufu.br, robinsonsabino@gmail.com

**Abstract.** *Dentin hypersensitivity can significantly impact the life quality of patients. The lack of non-invasive diagnostic techniques makes the problem even worse. To address this, we developed an end-to-end pipeline composed of two major components. In the first, patient samples of gingival crevicular fluid underwent Fourier-transform infrared spectroscopy (FTIR) analysis, followed by preprocessing steps including normalization and smoothing methods. In the second, several machine learning models were developed and evaluated. The results evidenced the feasibility of non-invasive dentin hypersensitivity detection for XGBoost and BiLSTM, as well as the importance of data preparation and parameter selection to deal with the inherent challenges of FTIR data.*

**Resumo.** *A hipersensibilidade dentinária pode afetar a qualidade de vida dos pacientes. A falta de métodos de diagnóstico não-invasivos torna o problema ainda pior. Como alternativa, nós desenvolvemos uma solução fim-a-fim composta de dois componentes principais. No primeiro, amostras do fluido crevicular gengival de pacientes passaram por espectroscopia de infravermelho por transformada de Fourier (FTIR), seguidas por etapas de pré-processamento incluindo métodos de normalização e suavização. No segundo, vários 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 cemento 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].

O aprendizado de máquina objetiva a construção de programas com melhoria 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. Muitos modelos de aprendizado de máquina e aprendizado profundo são considerados “caixa-preta” devido à sua complexidade e dificuldade de interpretação. Diante desse cenário, surgiu a demanda por Inteligência Artificial Explicável para tornar esses modelos mais transparentes e confiáveis [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.

## **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 FTIR 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]. Outro estudo colaborou para a detecção do câncer oral através do desenvolvimento de modelos preditivos de bom desempenho, combinando métodos de correção de linha de base e técnicas de classificação [Filho et al. 2023]. [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.

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.

#### **3.1. Base de dados**

A base de dados, proveniente do trabalho de [Souza 2018], consiste em 40 amostras de fluido crevicular gengival de 20 pacientes, com idades entre 20 e 50 anos, sem periodontite e boa higiene oral. O diagnóstico de hipersensibilidade dentinária foi feito por meio do teste de jato de ar e escala visual analógica, resultando em 20 amostras de controle e 20 de diagnóstico positivo. Cada amostra inclui valores de infravermelho vibracional obtidos por espectrômetro de infravermelho por transformada de Fourier de refletância total.

#### **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**

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. 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.

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 no conjunto de dados, foram realizadas 10 repetições utilizando a função ‘repeatedkfold’, considerando o resultado médio e o desvio padrão de todas as execuções.

A análise e comparação de desempenho dos modelos foi realizada com base nas medidas de sensibilidade, especificidade, acurácia, precisão e pela média harmônica entre sensibilidade e especificidade (F1-score).

#### 4. Resultados

A Tabela 1 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 1. 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 ±5,6%	48,8 ±6,5%	42,5 ±5,1%	55,0 ±7,7%	47,7 ±5,5%
AD	57,0 ±7,7%	56,5 ±7,4%	61,5 ±9,5%	52,5 ±10,1%	56,0 ±8,0%
Naive Bayes	53,0 ±3,1%	53,6 ±3,6%	46,0 ±4,9%	60,0 ±5,5%	51,8 ±3,3%
<b>XGBoost</b>	<b>74,7</b> ±6,4%	<b>74,3</b> ±7,9%	<b>77,0</b> ±5,1%	<b>72,5</b> ±10,8%	<b>74,2</b> ±7,9%
MLP	58,5 ±3,6%	59,7 ±4,7%	52,5 ±5,6%	64,5 ±4,7%	57,6 ±3,6%
BiLSTM	70,0 ±6,3%	70,1 ±5,4%	69,5 ±10,5%	70,5 ±7,6%	68,8 ±6,9%

A Tabela 2 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 1 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.

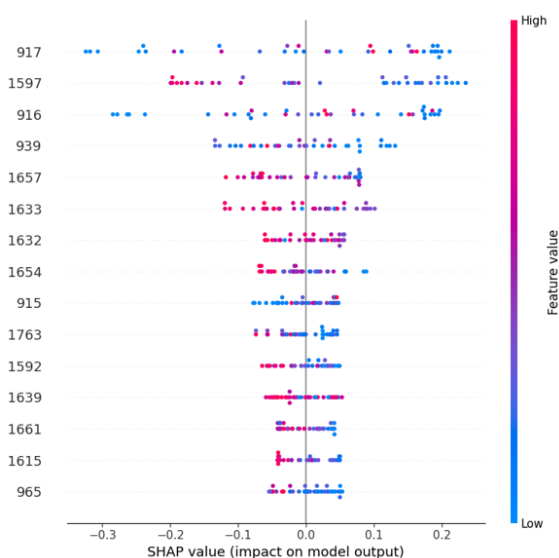
**Tabela 2. Resultados dos modelos de classificação em porcentagem (média e desvio padrão), utilizando o pré-processamento *Savitzky-Golay* e destacando os resultados mais promissores de derivação.**

	Acurácia	Precisão	Sensibilidade	Especificidade	F1-Score
KNN	54,7 ±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%	<b>74,5</b> ±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%
<b>BiLSTM</b>	<b>71,5</b> ±5,0%	<b>73,5</b> ±4,9%	67,5 ±11,9%	<b>75,5</b> ±6,1%	<b>70,4</b> ±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 [Vilone and Longo 2021].

A Figura 1 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 1. 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.

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.

## 5. Considerações Finais

Neste estudo, foram desenvolvidos modelos de aprendizado supervisionado para o diagnóstico da hipersensibilidade dentinária utilizando dados de FTIR. 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 não invasiva e de baixo custo da 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, considerar métodos de correção de *baseline*, como *rubberband* e *polynomial*, bem como explorar outras técnicas de aprendizado profundo tais como redes convolucionais. Um banco de dados público é fundamental para fornecer uma base ampla, sólida e acessível, facilitando a validação e comparação de diferentes técnicas e, consequentemente, a melhoria no desempenho dos classificadores.

## Referências

- 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.
- 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.
- 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.
- 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.