



Faculdade de Medicina de São José do Rio Preto
Programa de Pós-graduação em Enfermagem

DANIELA DENIS GUIDO

**PREDIÇÃO DE TRATAMENTOS
ORTODÔNTICOS COMO ESTIMATIVA DO
ALINHAMENTO DA DENTIÇÃO: UMA
ABORDAGEM DE INTELIGÊNCIA ARTIFICIAL
APLICADA**

**São José do Rio Preto
2024**

DANIELA DENIS GUIDO

**PREDIÇÃO DE TRATAMENTOS
ORTODÔNTICOS COMO ESTIMATIVA DO
ALINHAMENTO DA DENTIÇÃO: UMA
ABORDAGEM DE INTELIGÊNCIA ARTIFICIAL
APLICADA**

Dissertação apresentada ao Programa de Pós-Graduação Stricto Sensu em Enfermagem da Faculdade de Medicina de São José do Rio Preto, para obtenção do Título de Mestre. Área de Concentração: Cuidados de enfermagem e de saúde nos ciclos de vida. Linha de Pesquisa: Gestão e educação em enfermagem e saúde. Grupo de Pesquisa: Educação em Saúde (EDUS).

Orientador: Prof. Dr. Alexandre Lins Werneck
Coorientadora: Prof^a. Dr^a. Rita de Cássia Helu Mendonça Ribeiro

**São José do Rio Preto
2024**

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESSE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Ficha Catalográfica

Guido, Daniela Denis.

Predição de Tratamentos Ortodônticos como Estimativa do Alinhamento da Dentição: uma abordagem de inteligência artificial aplicada / Daniela Denis Guido. São José do Rio Preto 2024.
57 p.

Dissertação (Mestrado) – Programa de Pós-Graduação *Stricto Sensu* em Enfermagem da Faculdade de Medicina de São José do Rio Preto.
Área de Concentração: Cuidados de enfermagem e de saúde nos ciclos de vida.
Linha de Pesquisa: Gestão e educação em enfermagem e saúde.
Grupo de Pesquisa: Educação em Saúde (EDUS).

Orientador: Prof. Dr. Alexandre Lins Werneck

1. Predição; 2. Ortodontia Corretiva; 3. Inteligência Artificial; 4. Atenção à Saúde; 5. Tomada de Decisões; 6. Gestão em Saúde

BANCA EXAMINADORA

BANCA EXAMINADORA

Presidente: Profa. Dra. Rita de Cássia Helu Mendonça Ribeiro
FAMERP - SJRP

2º Examinador: Prof. Dr. João Paulo Lemos Escola
IFSP - Barretos

3º Examinador Profa. Dra. Luciene Cavalcanti Rodrigues FATEC
- SJRP

Suplente: Profa. Dra. Ana Carla Raphaelli Nahás Scocate UNORTE -
SJRP

Suplente: Prof. Dr. Marli de Carvalho Jericó
FAMERP - SJRP

São José do Rio Preto, __/__/__

A Deus, pois sem Ele eu nem mesmo existiria...

Ao meu orientador, Professor Alexandre Lins Werneck, que com enorme carinho dedicou-me o seu tempo e a sua atenção para viabilizar este trabalho.

à Professora Dra. Rita de Cássia Helu Mendonça Ribeiro, pelo expressivo apoio.

SUMÁRIO

Lista de Figuras.....	i
Lista de Tabelas e Quadros.....	ii
Resumo.....	iii
Abstract	iv
Resumen.....	v
1 INTRODUÇÃO.....	1
2 OBJETIVOS.....	8
3 METODOLOGIA.....	10
3.1 Modelamento do Preditor por Regressão Linear.....	14
3.2 Modelamento do Preditor por Regressão Logística.....	20
3.3 Modelamento do Preditor por Perceptrons.....	21
4 RESULTADOS.....	24
5 DISCUSSÃO.....	33
6 CONCLUSÃO.....	48
7 REFERÊNCIAS.....	51

Figura 1.	A sequência de passos para a abordagem proposta.....	11
Figura 2.	Coleta de imagens intrabucais para o tratamento clássico com Invisalign.....	12
Figura 3.	Interface on-line da base Invisalign.....	12
Figura 4.	a) Imagens radiográficas de um paciente da base Invisalign; b) Imagens sintéticas geradas a partir da coleta automática de fotos intrabucais utilizando o instrumento iTero, desenvolvido pela Align para viabilizar o tratamento com Invisalign.....	13
Figura 5.	Erros positivos e negativos, respectivamente, nas cores azul-petróleo e verde.....	15
Figura 6.	Exemplo de regressão logística. A curva se adequa aproximadamente aos pontos em verde, minimizando o erro quadrático médio.....	21
Figura 7.	A estrutura usual do perceptron de três camadas, similar ao que está sendo utilizado neste trabalho.....	22
Figura 8.	Modelo de rede neural perceptron adaptado para predição e usado neste trabalho. Notavelmente, a última camada possui apenas um conjunto de pesos.....	22
Figura 9.	Hiperplano tridimensional hipotético, nas coordenadas $\{x, y, z\}$, capaz de ser modelado por uma estrutura de perceptron multicamadas.....	23
Figura 10.	Resultados para o melhor caso de validação cruzada na análise de regressão linear.....	25
Figura 11.	Resultados para o melhor caso de validação cruzada na análise de regressão logística, com o eixo vertical normalizado em função do domínio da função logística, isto é, $0 < f(x) = \frac{1}{1 + e^{-x}} < 1$	26
Figura 12.	Resultados para o melhor caso de validação cruzada na análise de perceptron com uma camada oculta.....	27
Figura 13.	Resultados para o melhor caso de validação cruzada na análise de perceptron com duas camadas ocultas.....	27
Figura 14.	Resultados para o melhor caso de validação cruzada na análise de perceptron com três camadas ocultas.....	28
Figura 15.	Exemplo de um alinhador Invisalign na boca de um paciente.....	32

Quadro 1.	Algoritmo Proposto.....	14
Tabela 1.	Erros quadráticos médios para cada uma das estratégias de predição testadas.....	28

GUIDO, D. D. Predição de Tratamentos Ortodônticos como Estimativa do Alinhamento da Dentição: uma abordagem de inteligência artificial aplicada. 57 f. Dissertação (Mestrado) – Pós-Graduação Stricto Sensu em Enfermagem. Faculdade de Medicina de São José do Rio Preto, São José do Rio Preto, 2024.

Objetivo: O estudo visa aplicar e comparar modelos inteligentes otimizados para estimar a duração dos tratamentos ortodônticos com alinhadores transparentes removíveis Invisalign, para melhorar a precisão no planejamento de cuidados bucais na enfermagem, otimizar os recursos e tempo dos profissionais de saúde, além de proporcionar um melhor acompanhamento dos pacientes. O intuito é melhorar a precisão no planejamento de cuidados, otimizar recursos e tempo dos profissionais de saúde e proporcionar um melhor acompanhamento dos pacientes.

Metodologia: Foi realizado um estudo quantitativo para validar e comparar a eficácia de diferentes modelos preditivos na estimativa do tempo de duração dos tratamentos ortodônticos com alinhadores Invisalign. A análise dos dados foi conduzida utilizando regressão linear com validação cruzada para assegurar a robustez dos resultados e evitar overfitting. A validação cruzada foi implementada dividindo o conjunto de dados em múltiplas subamostras, utilizando cada subconjunto para o treinamento e teste do modelo em diferentes combinações. Além da regressão linear, outros modelos preditivos foram avaliados, incluindo regressão logística e perceptrons com diferentes camadas ocultas. A métrica de desempenho utilizada foi o erro quadrático médio (EQM), para quantificar a acurácia das previsões. Uma análise comparativa identificou o modelo mais adequado considerando a precisão preditiva e a aplicabilidade prática na área de enfermagem. **Resultados:** A regressão Linear apontou um EQM de 1,78, mostrando um desempenho razoável. A Regressão Logística, com um EQM de 5,33, teve o pior desempenho. O Perceptron com uma camada oculta, apresentou um EQM de 1,56, mostrando melhor desempenho. Com relação ao Perceptron com duas camadas ocultas, o EQM foi de 1,61, com um desempenho ligeiramente inferior ao de uma camada. O Perceptron com três camadas ocultas e um EQM de 1,63, apresentou desempenho similar ao de duas camadas. Os resultados mostram que o perceptron com uma camada oculta foi o modelo mais eficaz para a predição dos dados, com o menor erro quadrático médio. A regressão logística teve o pior desempenho, enquanto a regressão linear apresentou um desempenho intermediário. Os perceptrons com múltiplas camadas ocultas não apresentaram melhorias significativas em relação ao modelo com uma camada oculta, sugerindo overfitting ou complexidade desnecessária. **Conclusão:** A utilização de inteligência artificial (IA) na avaliação do cuidado ortodôntico, bem como na educação e pesquisa em saúde e enfermagem, permite uma análise mais robusta e detalhada dos dados. As predições geradas pela IA podem informar políticas e práticas, garantindo uma base científica sólida para a tomada de decisões. A predição de tratamentos ortodônticos com IA, alinhada aos referenciais teórico-filosóficos e metodológicos da gestão e avaliação em enfermagem e saúde, mostra como a inovação tecnológica pode ser eficazmente integrada em contextos de saúde, promovendo avanços significativos na qualidade dos cuidados prestados. Para este conjunto de dados, o modelo de perceptron simples com uma camada oculta foi o mais adequado para a predição, destacando a importância de escolher o modelo correto para otimizar os recursos e melhorar o acompanhamento dos pacientes em tratamentos ortodônticos.

Descritores: 1. Predição; 2. Ortodontia corretiva; 3. Inteligência artificial; 4. Atenção à saúde, 5. Tomada de decisões; 6. Gestão em saúde

GUIDO, D. D. Prediction of Orthodontic Treatments as an Estimate of Dental Alignment: an Applied Artificial Intelligence Approach. 57 p. Master Thesis – Stricto Sensu Graduate Nursing Program. Faculdade de Medicina de São José do Rio Preto, São José do Rio Preto, 2024.

Objective: The study aims to apply and compare optimized intelligent models to estimate the duration of orthodontic treatments with Invisalign removable transparent aligners. The goal is to improve the accuracy of oral care planning in nursing, optimize the resources and time of health professionals, and provide better patient follow-up. In summary, the aim is to improve precision in care planning, optimize resources and time for health professionals, and provide better patient follow-up. **Methodology:** We carried out a quantitative study to validate and compare the effectiveness of different predictive models in estimating the duration of orthodontic treatments with Invisalign aligners. We conducted Data analysis using cross-validation to ensure the robustness of the results. Cross-validation was implemented by dividing the data set into multiple subsamples, using each subset to train and test the model in different combinations. In addition to linear regression, we evaluated other predictive models, including logistic regression and Perceptrons with different hidden layers. The performance metric used was the root mean square error (RMSE) to quantify the accuracy of the predictions. A comparative analysis identified the most suitable model regarding predictive accuracy and practical applicability in nursing. **Results:** Linear Regression pointed to an MSE of 1.78, showing reasonable performance. Logistic Regression, with an MSE of 5.33, had the worst performance. The Perceptron with one hidden layer had an MSE of 1.56, showing better performance. About the Perceptron with two hidden layers, the MSE was 1.61, with a slightly lower performance than the one layer. The Perceptron performed similarly to the two-layer with three hidden layers and an EQM of 1.63. The results show that the Perceptron with one hidden layer was the most effective model for predicting the data, with the lowest mean square error. This conclusion is clear and unequivocal: the Perceptron with one hidden layer is the best model choice. Logistic Regression had the worst performance, while linear Regression showed intermediate performance. Perceptrons with multiple hidden layers showed no significant improvement over the model with one hidden layer. **Conclusion:** The use of artificial intelligence (AI) in the evaluation of orthodontic care, as well as in health and nursing education and research, is a significant step towards improving patient care. It allows for a more robust and detailed data analysis, leading to better-informed decisions. Predictions generated by AI can inform policy and practice, ensuring a sound scientific basis for decision-making. The prediction of orthodontic treatments with AI, aligned with the theoretical-philosophical and methodological frameworks of nursing and health management and evaluation, shows how technological innovation can be effectively integrated into health contexts, promoting significant advances in the quality of care. For this data set, the simple Perceptron model with a hidden layer was the most suitable for prediction, highlighting the importance of choosing the right model to optimize resources and improve patient follow-up in orthodontic treatments.

Descriptors: 1. Forecasting; 2. Orthodontics, Corrective; 3. Artificial Intelligence; 4. Delivery of Health Care; 5. Decision Making; 6. Health Management.

GUIDO, D. D. Predicción de Tratamientos Ortodónticos como Estimación del Alineamiento Dental: un Enfoque de Inteligencia Artificial Aplicada. 57 p. Dissertação (Mestrado) – Stricto Sensu Estudios de Posgrado en Enfermería. Faculdade de Medicina de São José do Rio Preto, São José do Rio Preto, 2024.

Objetivo: El estudio pretende aplicar y comparar modelos inteligentes optimizados para estimar la duración de los tratamientos de ortodoncia con alineadores transparentes removibles Invisalign. El objetivo es mejorar la precisión de la planificación de la atención bucodental en enfermería, optimizar los recursos y el tiempo de los profesionales de la salud y proporcionar un mejor seguimiento del paciente. En resumen, se pretende mejorar la precisión en la planificación de la atención, optimizar los recursos y el tiempo de los profesionales de la salud y proporcionar un mejor seguimiento del paciente. **Metodología:** Realizamos un estudio cuantitativo para validar y comparar la efectividad de diferentes modelos predictivos en la estimación de la duración de los tratamientos de ortodoncia con alineadores Invisalign. Realizamos un análisis de datos con validación cruzada para asegurar la robustez de los resultados. La validación cruzada se implementó dividiendo el conjunto de datos en múltiples submuestras, utilizando cada subconjunto para entrenar y probar el modelo en diferentes combinaciones. Además de la regresión lineal, evaluamos otros modelos predictivos, incluyendo la regresión logística y los Perceptrones con diferentes capas ocultas. La métrica de desempeño utilizada fue el error cuadrático medio (RMSE) para cuantificar la precisión de las predicciones. Un análisis comparativo identificó el modelo más adecuado en cuanto a precisión predictiva y aplicabilidad práctica en enfermería. **Resultados:** La regresión lineal apuntó a un MSE de 1,78, mostrando un rendimiento razonable. La regresión logística, con un MSE de 5,33, tuvo el peor rendimiento. El perceptrón con una capa oculta tuvo un MSE de 1,56, mostrando un mejor rendimiento. Sobre el perceptrón con dos capas ocultas, el MSE fue de 1,61, con un rendimiento ligeramente inferior al de una capa. El perceptrón se desempeñó de manera similar al de dos capas con tres capas ocultas y un EQM de 1,63. Los resultados muestran que el perceptrón con una capa oculta fue el modelo más eficaz para predecir los datos, con el error cuadrático medio más bajo. Esta conclusión es clara e inequívoca: el perceptrón con una capa oculta es la mejor opción de modelo. La regresión logística tuvo el peor desempeño, mientras que la regresión lineal mostró un desempeño intermedio. Los perceptrones con múltiples capas ocultas no mostraron una mejora significativa sobre el modelo con una capa oculta. **Conclusión:** El uso de inteligencia artificial (IA) en la evaluación de la atención de ortodoncia, así como en la educación e investigación en salud y enfermería, es un paso significativo hacia la mejora de la atención al paciente. Permite un análisis de datos más sólido y detallado, lo que conduce a decisiones mejor informadas. Las predicciones generadas por IA pueden informar la política y la práctica, asegurando una base científica sólida para la toma de decisiones. La predicción de tratamientos de ortodoncia con IA, alineada con los marcos teórico-filosóficos y metodológicos de la enfermería y la gestión y evaluación de la salud, muestra como la innovación tecnológica se puede integrar de manera efectiva en los contextos de salud, promoviendo avances significativos en la calidad de la atención. Para este conjunto de datos, el modelo de perceptrón simple con una capa oculta fue el más adecuado para la predicción, lo que resalta la importancia de elegir el modelo adecuado para optimizar los recursos y mejorar el seguimiento del paciente en los tratamientos de ortodoncia.

Descriptor: 1. Predicción; 2. Ortodoncia Correctiva; 3. Inteligencia Artificial; 4. Atenciones a la Salud; 5. Toma de Decisiones; 6. Gestión em Salud

1 INTRODUÇÃO

1 INTRODUÇÃO

A área de enfermagem possui estreita relação com a área de odontologia, particularmente em unidades de terapia intensiva (UTIs).¹⁻³ O suporte que o profissional de enfermagem pode oferecer à saúde bucal dos pacientes internados na unidade de terapia intensiva (UTI), por exemplo, é um nítido apontamento do inter-relacionamento das áreas.⁴ Há um foco no papel da equipe de enfermagem na prestação de cuidados bucais aos pacientes, os desafios que enfrentam e a necessidade de treinamento e conhecimento adequados nesta área.² Existe uma necessidade de maior colaboração, recursos e experiência em enfermagem intensiva para melhorar as práticas de saúde bucal e os resultados dos pacientes.³

Dessa forma, assim como ocorre na área de enfermagem e nas demais áreas das ciências da saúde, a área de odontologia, e mais especialmente a ortodontia, tem se beneficiado com os expressivos reflexos que o avanço tecnológico vem proporcionando ao longo dos anos.⁵ Além dos inúmeros eventos internacionais dedicados à revelar as novidades tecnológicas na área,⁶⁻⁸ organizados em âmbitos científico e comercial, a marca registrada dessa modernidade pode ser encontrada atualmente em qualquer consultório odontológico: os clássicos e antigos aparelhos de “Raios-X”, por exemplo, têm dado lugar aos sistemas computadorizados de aquisição e análise digital de imagens bucais.

Paralelamente ao avanço observável na ortodontia, faz-se também cristalino o desabrochar dos sistemas de Inteligência Artificial (IA) em todos os mecanismos que, previamente, dependiam de intervenção humana direta.⁹ Nítidos exemplos contemplam desde os aplicativos existentes nos *smartphones*, tais como aqueles destinados ao reconhecimento facial, passando pelos sistemas bancários, de reconhecimento de fala e de identificação de emoções, e chegando aos que controlam veículos e mísseis intercontinentais.⁹

As tecnologias de IA podem ajudar a monitorizar o estado de saúde oral dos doentes, a fornecer análises de dados em tempo real e a melhorar os processos de tomada de decisão dos prestadores de cuidados de saúde.³ Um estudo realçou o uso de IA para prever a experiência do paciente durante o tratamento *Invisalign*, com o objetivo de melhorar a adesão do paciente em aplicações clínicas. A pesquisa construiu redes neurais artificiais (ANNs) para prever os níveis de conforto do paciente, particularmente em termos de dor, ansiedade e qualidade de vida (QV) durante o tratamento com *Invisalign*.¹⁰

Assim, a possibilidade de utilizar métodos de IA para planejar os tratamentos ortodônticos é, claramente, de extrema valia para melhorar a condução e aumentar a satisfação dos clientes submetidos aos tratamentos. Além disso, observa-se que as estimativas fornecidas pelos fabricantes dos alinhadores transparentes têm se mostrado distantes da realidade clínica, conforme constatado por Lombardo et al.¹¹ e Morales-Burruezo et al.¹² Sistemas de IA desenvolvidos para prever o conforto do paciente tem potencial para aplicação clínica para melhorar a adesão do paciente.¹⁰

Visando viabilizar esta discussão, registre-se que foi realizada a busca por artigos científicos, publicados a partir de 2019 em periódicos internacionais da área de ortodontia com fator de impacto entre 0.341 e 7.600 no *Clarivate Analytics*, a qual retornou oito registros para os quais os temas específicos envolvem: previsão do movimento dos dentes, análise de traçado cefalométrico, predição de crescimento cervical, segmentação e detecção de padrões em imagens, objetivos educacionais em ortodontia e prognóstico de tratamento. No caso particular do prognóstico, foi encontrado um único artigo de revisão fundamentado em outros 49 artigos anteriores à 2019. Entretanto, nenhum dos artigos encontrados trata da previsão do tempo de tratamento com alinhadores transparentes. Apesar disso, visando substanciar esta pesquisa, apresenta-se, a seguir, um resumo de tais tratativas científicas.

Na revisão de literatura apresentada por Faber et al.¹³ os autores destacam o papel relevante da IA no reconhecimento de padrões e na predição de resultados em ortodontia, enfatizando que muito ainda resta a ser feito. Particularmente, considerando as abordagens não supervisionadas e as supervisionadas dos sistemas de IA, os autores comentam que as principais subáreas nas quais os sistemas inteligentes podem ser empregados são a de diagnóstico e de planejamento do movimento dos dentes. Nesse sentido, Rao et al.¹⁴ apresentam um sistema de IA baseado em aprendizagem profunda que permite analisar e identificar pontos-chave para o traçado cefalométrico a partir da imagem da face do paciente sob exame. A margem de erro do sistema foi considerada aceitável e condizente com as expectativas: dentre 220 pontos identificados, erros na faixa de 0 a 1 mm estavam presentes em 48 deles; erros na faixa de 2 a 3 mm puderam ser observados em 75 deles; erros na faixa de 4 a 5 mm fizeram-se presentes em 92 pontos e, finalmente, apenas 6 pontos distavam mais de 6 mm dos locais reais.

Kök et al.¹⁵ descreveram o desenvolvimento de uma estratégia de IA para predição do crescimento cervical em 300 indivíduos entre 8 e 17 anos. Sete algoritmos, a saber, k-vizinhos mais próximos, Naive Bayes, árvore de decisão, Perceptron multicamadas, máquina de vetor de suporte, floresta de caminhos aleatórios e regressão logística, foram implementados e comparados, concluindo-se que o *Perceptron* multicamadas permitiu contabilizar uma acurácia de até 97.1%, sendo portanto o método adotado, ao contrário dos k-vizinhos mais próximos que, pelo seu teor modesto, permitiu alcançar não mais do que 60.9% de precisão.

Indicações clínicas foram contrastadas àquelas tomadas por um sistema do tipo k-vizinhos mais próximos, assim como àquelas oriundas de um do tipo árvore de decisão,¹⁶ com o objetivo de avaliar a confiabilidade dos sistemas automatizados de análise cefalométrica quando comparados às decisões clínicas.

Os resultados, obtidos com 144 pacientes portadores de má-oclusão classe III, permitiram concluir que os algoritmos testados atuam de maneira próxima àquela esperada pelos odontólogos, possibilitando, desse modo, o uso na rotina clínica visando verificar as indicações clínico-cirúrgicas antes das tomadas de decisões.

De acordo com Bichu et al.¹⁷ as aplicações da IA na ortodontia têm ocorrido modestamente, apesar do avanço daquela área de pesquisa, mesmo considerando as três estratégias essenciais de aprendizado de máquina, isto é, procedimentos supervisionados, não supervisionados e por reforço, as quais contemplam árvores de decisão, abordagens Bayseanas, redes neurais artificiais, máquinas de vetores de suporte, algoritmos genéticos e lógica Fuzzy. Nesse sentido, o pré-diagnóstico, a classificação de padrões esqueléticos e a previsibilidade de tratamentos ortodônticos considerando a avaliação de riscos têm sido as três linhas nas quais a maioria dos esforços tem sido observadas. No que tange ao problema da previsibilidade, notam-se recentes progressos, mas ainda considerados modestos,¹⁸ por meio do qual os autores reportam os seus esforços no sentido de realizar análises cefalométricas automatizadas com base em um tipo particular de rede neural artificial conhecido como Rede Neural Convolucional (RNC), que é um tipo de estratégia de aprendizado profundo. Esforços similares foram também realizados por Chen et al.¹⁹ visando segmentar imagens tri-dimensionais para avaliar atresia dos maxilares.

Yamashiro e Ko,²⁰ editores convidados para a edição especial que trata de IA com aplicações na ortodontia registraram as suas opiniões sobre as sub-áreas nas quais os sistemas computacionais inteligentes têm possibilitado os mais marcantes avanços na ortodontia. Entre elas, destacam-se a análise automática de imagens cefalométricas, de maturação esquelética, de formato palatal, de tomada de decisões relacionadas com possíveis extrações dentárias e, além disso, a previsibilidade nos tratamentos. Tais comentários são corroborados pela revisão sistemática e de metanálise constante de Monill-González,²¹ no qual os autores selecionaram

17 publicações em meio ao montante de 311 registros do *Web of Science* e bases correlatas no período de 2010 a 2021. Particularmente, os autores destacaram que as RNCs foram as mais aplicadas na detecção automática de pontos anatômicos em imagens radiográficas, ao passo que maturações cervicais foram estimadas com sucesso com base em redes neurais artificiais rasas, que são as mais clássicas. Congruentemente, a revisão sistemática de Bichu et al. que seguiu os mesmos padrões da mencionada anteriormente considerando a base PubMed, foi formulada com base em 62 artigos. Neles, notou-se que as RNCs, as redes rasas e as Máquinas de Vetores de Suporte têm sido as técnicas mais comuns para solucionar os problemas mencionados.

Com um objetivo mais desafiador, Retrouvey²² destaca que o próximo passo da aplicação dos sistemas inteligentes na ortodontia é a criação de modelos de IA com potencialidade para aprender a realizar diagnósticos em ortodontia.

Para isso, o autor revisa as potencialidades dos modelos de aprendizado profundo em IA destacando aplicações como as citadas nos parágrafos anteriores. Uma observação relevante realizada pelo autor é que as redes neurais profundas requerem quantidades significantes de dados para que sejam treinadas convenientemente, o que corrobora com todo o conhecimento disponível na área de IA.⁹ Do mesmo modo, Leonardi et al.²³ destacam a potencialidade das estratégias recentes de IA quando associadas a conjuntos volumosos de dados, mostrando, inclusive, aplicações daquelas técnicas para segmentação de imagens ortodônticas

A segmentação de imagens de tomografia computadorizada por meio de técnicas de IA é o tema de Wang et al.²⁴ no qual os autores relatam resultados promissores com o uso de RNCs para isolar mandíbula, dentes e plano de fundo, reduzindo o trabalho manual do ortodontista. Adicionalmente, a revisão de literatura apresentada por Mohammad-Rahimi et al.²⁵ realizada por meio de uma seleção de 49 artigos das bases PubMed, Google, Scopus e Embase no período de 2010 a 2020, permite também concluir que a IA pode auxiliar os ortodontistas a economizarem tempo e realizar prognósticos com expressiva precisão. Kök et

al.²⁶ confirma esse fato pois, os autores usam com sucesso técnicas de IA para determinar alterações nas estruturas ósseas durante a fase ortodôntica. Complementarmente, Narayan et al.²⁷ mencionam, inclusive, o papel fundamental da IA na aprendizagem de ortodontia nos cursos de graduação, contemplando os sistemas tutores inteligentes, o ensino pré-clínico, e os mecanismos de gerenciamento de informação ortodôntica, em associação com estratégias de realidade virtual e realidade aumentada.

Diante da revisão apresentada, nota-se que nenhum dos trabalhos mencionados, e nem mesmo qualquer outro que se tenha notícia por meio de registro formal na literatura acadêmica, foi dedicado à solução do caso em tela. Desse modo, a aplicação e comparação dos modelos de IA experimentados nesta dissertação, objetivando a predição do tempo de tratamento com alinhadores *Invisalign* é, de fato, inovadora.

Conforme as observações registradas, a pesquisa proposta abre caminho para, caracterizando a solução de um problema relevante para a subárea na qual se encaixa. Adicionalmente, o uso de uma base de dados pública, conforme descrito adiante, permite assegurar a realização da pesquisa e evita a perda de sequência nos segmentos do trabalho em função da falta de controle sob a dedicação e colaboração dos pacientes envolvidos.

Apesar do destacado progresso disjuncto em ambas as áreas mencionadas, isto é, ortodontia e IA, é possível notar que as aplicações desta no aprimoramento daquela ainda são relativamente modestos. A busca pelos termos “artificial intelligence” ou “machine learning” associados ao termo “orthodontics” na base do *Web of Science* retorna somente algumas dezenas de artigos publicados em periódicos de impacto. Desse modo, é nítida a existência de uma lacuna a ser explorada e, portanto, um potencial caminho a ser percorrido. Para especificá-lo e balizá-lo, as publicações científicas dos três anos mais recentes foram selecionadas e estudadas.

2 OBJETIVOS

2 OBJETIVOS

O estudo visa aplicar e comparar modelos inteligentes otimizados para estimar a duração dos tratamentos ortodônticos com alinhadores transparentes removíveis *Invisalign*, para melhorar a precisão no planejamento de cuidados bucais na enfermagem, otimizar os recursos e tempo dos profissionais de saúde, além de proporcionar um melhor acompanhamento dos pacientes. O intuito é melhorar a precisão no planejamento de cuidados, otimizar recursos e tempo dos profissionais de saúde e proporcionar um melhor acompanhamento dos pacientes.

3 METODOLOGIA

3 METODOLOGIA

Foi realizado um estudo quantitativo para validar e comparar a eficácia de diferentes modelos preditivos na estimativa do tempo de duração dos tratamentos ortodônticos com alinhadores Invisalign. A análise dos dados foi conduzida com validação cruzada para assegurar a robustez dos resultados. A validação cruzada foi implementada dividindo o conjunto de dados em múltiplas subamostras, utilizando cada subconjunto para o treinamento e teste do modelo em diferentes combinações. Além da regressão linear, outros modelos preditivos foram avaliados, incluindo regressão logística e Perceptrons com diferentes camadas ocultas. A métrica de desempenho utilizada foi o erro quadrático médio (EQM) para quantificar a acurácia das previsões. Uma análise comparativa identificou o modelo mais adequado considerando a precisão preditiva e a aplicabilidade prática na área de enfermagem. A abordagem proposta para viabilizar as previsões almejadas para o tempo de tratamento com *Invisalign* ocorre de acordo com o diagrama exibido na Figura 1, detalhada no *Algoritmo Invisa*.

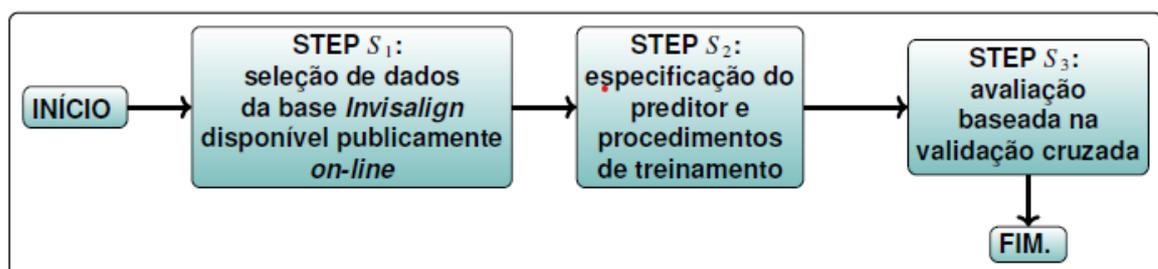


Figura 1. A seqüência de passos para a abordagem proposta.

Fonte: <https://site.invisa.org.br/>

Solicitamos ao Comitê de Ética e Pesquisa em Seres Humanos da Instituição, a dispensa de apresentação do Termo de Consentimento Escrito e Informado, porque, para a realização dos experimentos descritos, os dados são de caráter público e irrestrito da base de dados e das técnicas utilizadas. A interface *on-line* está ilustrada na Figuras 2, 3 e 4 respectivamente.



Figura 2. Coleta de imagens intrabucais para o tratamento clássico com *Invisalign*

Fonte: <https://samantanigro.com.br/itero-element-5d-plus-4-beneficios-para-o-paciente/>

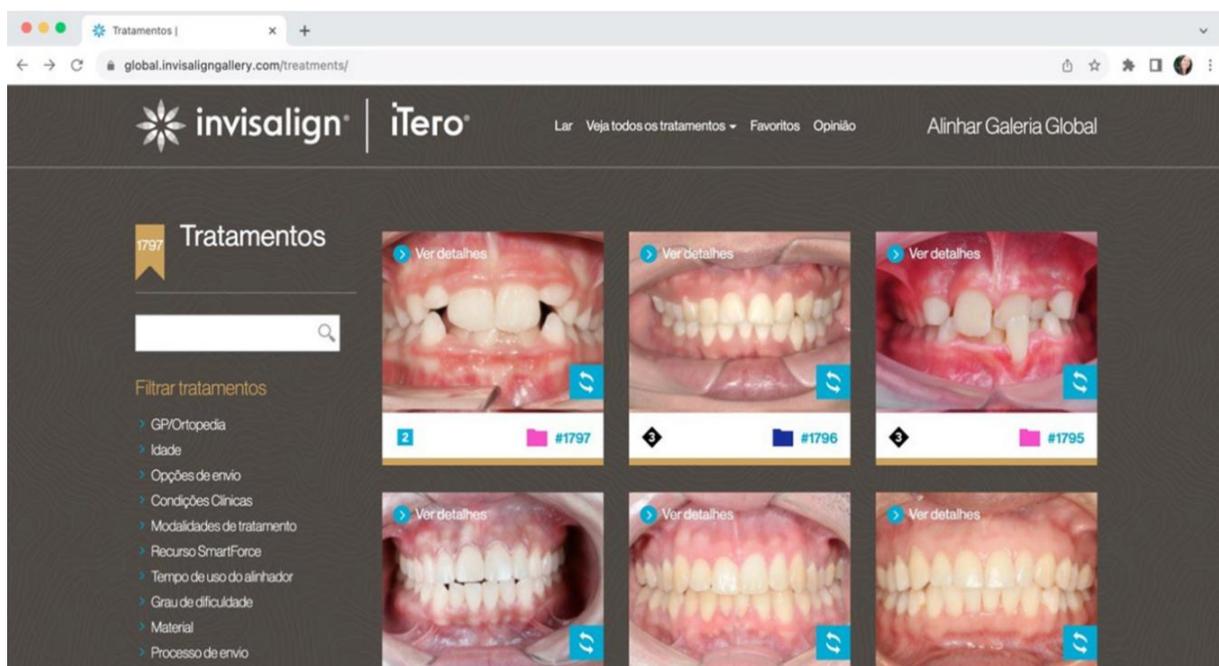


Figura 3. Interface *on-line* da base *Invisalign*

Fonte: Invisalign. Invis is straightening teeth [homepage na Internet]. San Jose: Align Technology, Inc.; 2024 [acesso em 2024 Nov 2]. Disponível em: <https://www.invisalign.com>

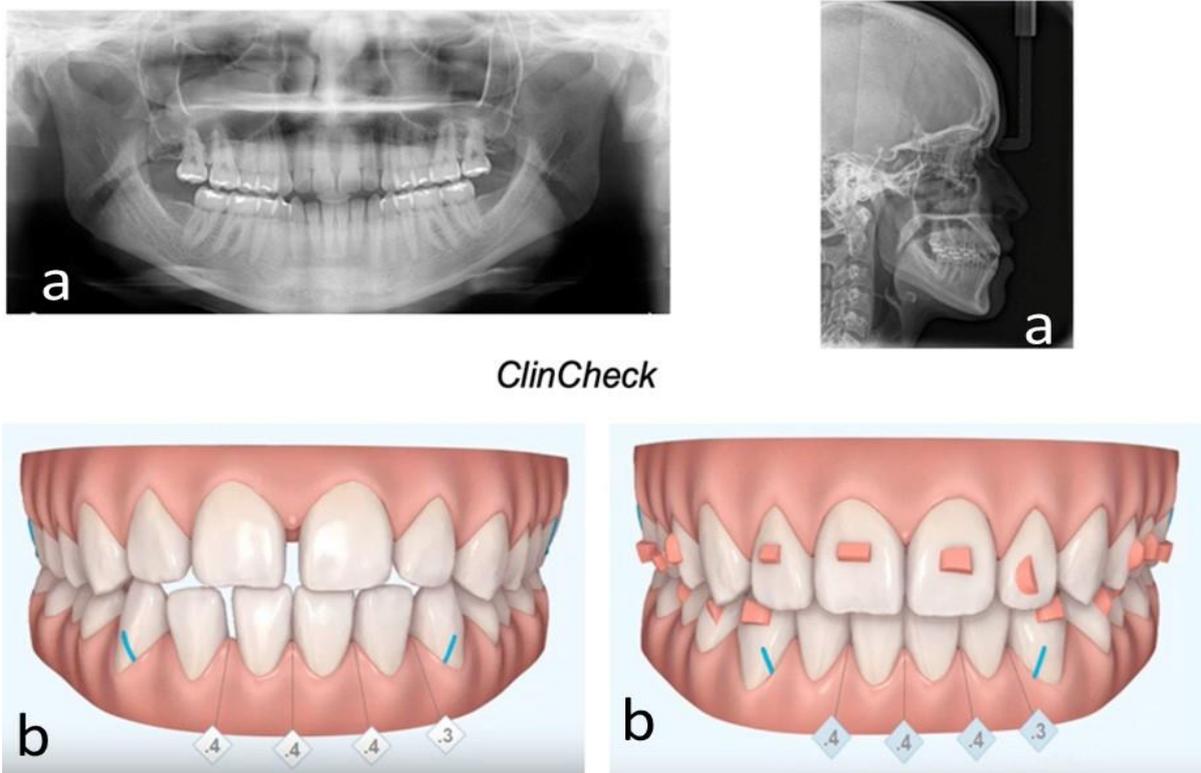


Figura 4. Imagens radiográficas de um paciente da base Invisalign; **b)** Imagens sintéticas geradas a partir da coleta automática de fotos intrabucais utilizando o instrumento *iTero*, desenvolvido pela *Align* para viabilizar o tratamento com *Invisalign*

Fonte: <https://samantanigro.com.br/itero-element-5d-plus-4-beneficios-para-o-paciente/>

Os procedimentos envolvendo a base de dados, conforme mencionado no Passo S_1 , são modestos: para cada um dos 1491 indivíduos, são coletados os seguintes dados: idade (e_1), gênero (e_2) tipo específico de alinhador utilizado (e_3), número de alinhadores usados na mandíbula (e_4), número de alinhadores usados na maxila (e_5), intervalo de tempo em dias entre cada troca de alinhadores (e_6) e o tempo total que foi necessário para o tratamento (r).

Por fim, associado às imagens, também é coletado o tempo total de tratamento para cada paciente. Em suma, a descrição abaixo contém o passo-a-passo de todo o algoritmo proposto (Quadro 1) desenvolvido neste trabalho, desde a coleta da base *Invisalign* de imagens, até a avaliação dos modelos propostos no conjunto de validação.

Quadro 1. Algoritmo Proposto

----- ALGORITMO PROPOSTO -----
<ul style="list-style-type: none">• INÍCIO • PASSO S_1: A base de imagens <i>Invisalign</i>, disponível <i>on-line</i>,²⁸ é obtida e preparada para uso. • PASSO S_2: O preditor, que pode ser baseado em regressão linear, regressão logística ou no Perceptron multicamadas, é definido e treinado com base em uma parcela da base de dados. • PASSO S_3: O preditor é avaliado valendo-se da parcela da base de dados que não foi utilizada para o seu treinamento com base na divisão de treino e teste de maneira aleatoriamente estratificada. • FIM.

Assim, sete parâmetros foram isolados para cada indivíduo, ou seja:

- $S_1\{e_1, e_2, e_3, e_4, e_5, e_6, r\}$ são os sete parâmetros para o indivíduo S_1 ;

- $S_2\{e_1, e_2, e_3, e_4, e_5, e_6, r\}$ são os sete parâmetros para o indivíduo S_2 ;

- e assim por diante.

Uma vez que os dados tenham sido armazenados, uma parcela deles, foi usado para treinar o preditor escolhido, isto é, o preditor por regressão linear, por regressão logística ou baseado no perceptron multicamadas. Desse modo, antes de prosseguir com as especificações dos três preditores adotados, as três próximas Seções contêm as descrições generalizadas do funcionamento de cada um deles.

3.1 Modelamento do Preditor por Regressão Linear

Regressão Linear é uma técnica matemática usada para encontrar uma equação do tipo $y_i = \alpha x_i + \beta$, sendo $(\alpha, \beta \in \mathfrak{R})$ e $(i = 0, 1, 2, \dots, M - 1)$, que melhor casa com um conjunto de M pontos, minimizando o erro quadrático médio, isto é, a soma das distâncias entre cada ponto

ideal e sua respectiva predição. Tradicionalmente, esse procedimento estatístico é conhecido como Método dos Mínimos Quadrados (*Least Mean Squares (LMS)*).²⁹ Sempre que $M = 2$, os pontos (x_0, y_0) e (x_1, y_1) casam-se perfeitamente ao modelo de linha reta, sem erro (E). Para $M > 2$ e pontos não alinhados, entretanto, E não se anula pois os pontos não correspondem a uma reta exata.

Apesar disso, o LMS pode ser utilizado para determinar α e β que minimizam E . Para determiná-los, simplesmente anula-se a derivada da função de erro, conforme a Figura 5, na qual, para cada ponto no conjunto $\{x_i, y_i\}$, existe um erro em relação à linha reta. A intenção é a de minimizar a soma dos quadrados dos erros, sendo que os quadrados são considerados para evitar possíveis sinais negativos. Linhas das cores azul-petróleo e verde foram utilizadas para representar erros positivos e negativos, respectivamente.

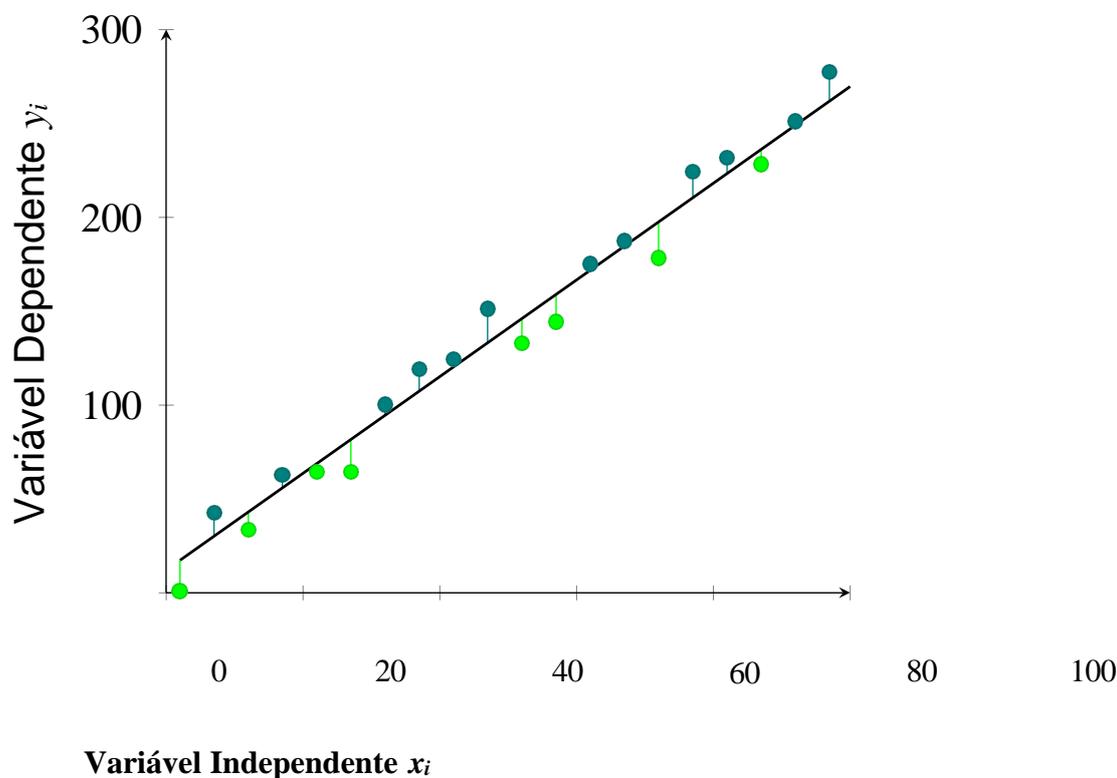


Figura 5. Erros positivos e negativos, respectivamente, nas cores azul-petróleo e verde.

Fonte: A própria pesquisadora

Para implementar o LMS, inicia-se escrevendo o sistema de equações lineares originado do conjunto de pontos a serem casados, ou seja, $\{x_i, y_i\}$, como segue:

$$\begin{cases} y_0 = \alpha x_0 + \beta \\ y_1 = \alpha x_1 + \beta \\ \dots \\ y_{M-1} = \alpha x_{M-1} + \beta \end{cases} .$$

Entendendo que α e β são as variáveis a serem determinadas pela solução do sistema de equações lineares, somente existe solução exata para $M = 2$. Para $M > 2$, entretanto, não existem α e β que resolvam exatamente o sistema. Por outro lado, é possível encontrar uma solução aproximada para o sistema. Para isso, o sistema é inicialmente expresso na forma matricial:

$$\underbrace{\begin{pmatrix} x_0 & 1 \\ x_1 & 1 \\ \dots & \dots \\ x_{M-1} & 1 \end{pmatrix}}_{\text{matriz } A} \cdot \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\text{variáveis}} = \underbrace{\begin{pmatrix} y_0 \\ y_1 \\ \dots \\ y_{M-1} \end{pmatrix}}_{\text{matriz } B} .$$

Calcular a derivada parcial de cada equação em relação à cada variável e anulá-las [13] é equivalente a multiplicar ambos os lados do sistema anterior pela transposta da matriz A , isto é, A_T , conforme segue:

$$\underbrace{\begin{pmatrix} x_0 & x_1 & \dots & x_{M-1} \\ 1 & 1 & \dots & 1 \end{pmatrix}}_{\text{matriz } A^T} \cdot \underbrace{\begin{pmatrix} x_0 & 1 \\ x_1 & 1 \\ \dots & \dots \\ x_{M-1} & 1 \end{pmatrix}}_{\text{matriz } A} \cdot \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\text{variáveis}} = \underbrace{\begin{pmatrix} x_0 & x_1 & \dots & x_{M-1} \\ 1 & 1 & \dots & 1 \end{pmatrix}}_{\text{matriz } A^T} \cdot \underbrace{\begin{pmatrix} y_0 \\ y_1 \\ \dots \\ y_{M-1} \end{pmatrix}}_{\text{matriz } B} .$$

Notavelmente, sendo A uma matriz de ordem $X \times 2$ então A^T é de ordem $2 \times X$, implicando que a multiplicação $A^T \cdot A$ é sempre possível e resulta em uma matriz quadrada de ordem 2. Do mesmo modo, no lado direito do sistema, a multiplicação de A^T , de ordem $S \times M$, por B , de ordem $M \times 1$, é também sempre possível, resultando em uma matriz de ordem 2×1 , como segue:

$$\underbrace{\begin{pmatrix} (x_0)^2 + (x_1)^2 + \dots + (x_{M-1})^2 & x_0 + x_1 + \dots + x_{M-1} \\ x_0 + x_1 + \dots + x_{M-1} & M \end{pmatrix}}_{A^T A \text{ é uma matriz } 2 \times 2} \cdot \underbrace{\begin{pmatrix} \alpha \\ \beta \end{pmatrix}}_{\text{variáveis}} = \underbrace{\begin{pmatrix} x_0 y_0 + x_1 y_1 + \dots + x_{M-1} y_{M-1} \\ y_0 + y_1 + \dots + y_{M-1} \end{pmatrix}}_{A^T B \text{ é uma matriz } 2 \times 1} .$$

Equivalentemente, tem-se:

$$\begin{cases} \alpha \sum_{i=0}^{M-1} (x_i)^2 + \beta \sum_{i=0}^{M-1} x_i = \sum_{i=0}^{M-1} x_i y_i \\ \alpha \sum_{i=0}^{M-1} x_i + M\beta = \sum_{i=0}^{M-1} y_i \end{cases} .$$

Resolvendo o sistema anterior, obtém-se as variáveis procuradas:

$$\alpha = \frac{\left(M \sum_{i=0}^{M-1} x_i y_i\right) - \left(\sum_{i=0}^{M-1} x_i \sum_{i=0}^{M-1} y_i\right)}{M \sum_{i=0}^{M-1} (x_i)^2 - \left(\sum_{i=0}^{M-1} x_i\right)^2}$$

$$\beta = \frac{\left(\sum_{i=0}^{M-1} (x_i)^2 \sum_{i=0}^{M-1} y_i\right) - \left(\sum_{i=0}^{M-1} x_i \sum_{i=0}^{M-1} x_i y_i\right)}{M \sum_{i=0}^{M-1} (x_i)^2 - \left(\sum_{i=0}^{M-1} x_i\right)^2}$$

Consequentemente, esses são α e β que minimizam o erro quadrático médio.

Particularmente, para a aplicação proposta neste trabalho, uma dimensão maior do que 2 é requerida pois o tempo de tratamento y_i não é somente função de um parâmetro independente x_i , mas de diversos. Assim, a formulação anterior pode ser estendida como segue

$$\left\{ \begin{array}{l} y_0 = \alpha x_0 + \beta z_0 + \gamma w_0 + \dots \\ y_1 = \alpha x_1 + \beta z_1 + \gamma w_1 + \dots \\ \dots \\ y_{M-1} = \alpha x_{M-1} + \beta z_{M-1} + \gamma w_{M-1} + \dots \end{array} \right.$$

onde $\{x_i, z_i, w_i, \dots\}$ são as variáveis procuradas. Para determiná-las, escrevem-se as equações para cada situação exemplo na base de dados usando matrizes. Em seguida, ambos os lados do sistema linear são multiplicados pela transposta da matriz principal e o sistema resultante, que é possível e determinado, é resolvido exatamente, provendo uma aproximação para o sistema original que minimiza o erro quadrático médio.

Particularmente, para a abordagem baseada na regressão linear, $T > 6$ vetores de tamanho sete foram usados de modo que os seis primeiros parâmetros ($\{e_1, e_2, e_3, e_4, e_5, e_6, r\}$)

serviram de entrada para o preditor, sendo que o sétimo (e_7) é o resultado desejado na saída do preditor. Consequentemente, a formulação resultou no seguinte sistema de equações:

$$\begin{cases} S_a(r) = \delta_1 S_a(e_1) + \delta_2 S_a(e_2) + \delta_3 S_a(e_3) + \delta_4 S_a(e_4) + \delta_5 S_a(e_5) + \delta_6 S_a(e_6) \\ S_b(r) = \delta_1 S_b(e_1) + \delta_2 S_b(e_2) + \delta_3 S_b(e_3) + \delta_4 S_b(e_4) + \delta_5 S_b(e_5) + \delta_6 S_b(e_6) \\ S_c(r) = \delta_1 S_c(e_1) + \delta_2 S_c(e_2) + \delta_3 S_c(e_3) + \delta_4 S_c(e_4) + \delta_5 S_c(e_5) + \delta_6 S_c(e_6) \\ \dots \\ \dots \\ \dots \end{cases}, \quad (3.2)$$

onde S_a , S_b , S_c , são os indivíduos da parcela da base de dados reservada para treinamento do preditor e, além disso, δ_1 , δ_2 , ..., δ_6 são os pesos a serem determinados, ou seja, as variáveis do sistema. Uma vez que se definiu $T > 6$, o sistema anterior possui mais equações do que incógnitas. Portanto, inexistente solução exata para ele, sendo apenas possível obter soluções aproximadas, possivelmente baseadas em Mínimos Quadrados (LMS).

Para encontrar uma solução aproximada, adotam-se os procedimentos explicados anteriormente, implicando que o sistema incompatível é escrito inicialmente na forma matricial, como na Equação 3.3:

$$\underbrace{\begin{pmatrix} S_a\{e_1\} & S_a\{e_2\} & S_a\{e_3\} & S_a\{e_4\} & S_a\{e_5\} & S_a\{e_6\} \\ S_b\{e_1\} & S_b\{e_2\} & S_b\{e_3\} & S_b\{e_4\} & S_b\{e_5\} & S_b\{e_6\} \\ S_c\{e_1\} & S_c\{e_2\} & S_c\{e_3\} & S_c\{e_4\} & S_c\{e_5\} & S_c\{e_6\} \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}}_{\text{matriz } A} \cdot \underbrace{\begin{pmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \end{pmatrix}}_{\text{variáveis}} = \underbrace{\begin{pmatrix} S_a\{r\} \\ S_b\{r\} \\ S_c\{r\} \\ \dots \end{pmatrix}}_{\text{matriz } B}. \quad (3.3)$$

Por fim, multiplicam-se ambos os lados do sistema por A^T , ou seja, o qual é resolvido para $\{\delta_1, \dots, \delta_6\}$ utilizando qualquer técnica regular para solução de sistemas lineares [13]. Uma vez que o sistema 6×6 tenha sido resolvido, o preditor para o tempo de tratamento (r) do indivíduo S_x é:

$$S x \{r\} = \delta 1 S x \{e1\} + \delta 2 S x \{e2\} + \delta 3 S x \{e3\} + \delta 4 S x \{e4\} + \delta 5 S x \{e5\} + \delta 6 S x \{e6\}. \quad (3.4)$$

3.2 Modelamento do Preditor por Regressão Logística

A segunda possibilidade planejada para experimentação é a regressão logística.

Ela funciona da mesma forma que a regressão linear, entretanto, a intenção é a de minimizar o erro quadrático médio em relação à uma função logística, ou seja, $y_i = \frac{1}{1+e^{-(\alpha x_i + \beta)}}$, em vez de uma função linear.²⁹ Equivalentemente, tem-se:

$$\frac{1}{y_i} = 1 + e^{-(\alpha x_i + \beta)} \rightarrow \frac{1}{y_i} - 1 = e^{-(\alpha x_i + \beta)} \rightarrow -\ln\left(\frac{1}{y_i} - 1\right) = \alpha x_i + \beta$$

Consequentemente e considerando uma dimensão mais alta, o sistema na Equação

3.1 se torna:

$$\left\{ \begin{array}{l} -\ln\left(\frac{1}{y_0} - 1\right) = \alpha x_0 + \beta z_0 + \gamma w_0 + \dots \\ -\ln\left(\frac{1}{y_1} - 1\right) = \alpha x_1 + \beta z_1 + \gamma w_1 + \dots \\ \dots \\ -\ln\left(\frac{1}{y_{M-1}} - 1\right) = \alpha x_{M-1} + \beta z_{M-1} + \gamma w_{M-1} + \dots \end{array} \right. , \quad (3.5)$$

que permite um modelamento baseado em uma função não linear tal qual a que consta na

Figura 6.

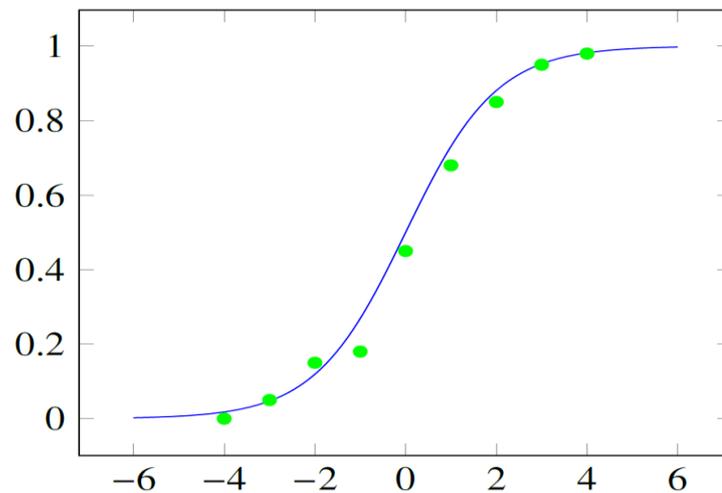


Figura 6. Exemplo de regressão logística. A curva se adequa aproximadamente aos pontos em verde, minimizando o erro quadrático médio.

Fonte: A própria pesquisadora

3.3 Modelamento do Preditor por Perceptrons

Os modelos de redes neurais perceptron multicamadas,⁹ tal qual o exibido na Figura 7 funcionam como aproximadores universais de funções e, portanto, são conhecidos pelos seus potenciais de solucionar tarefas complexas de inteligência artificial. Para treiná-los, ou seja, para ajustar os seus parâmetros de acordo com o problema específico a ser solucionado, a técnica de gradiente descendente⁹ é uma das estratégias utilizadas, assim como ocorre neste trabalho. Naquela figura, o conjunto de pesos determinado durante a fase de treinamento é $\{w_i j[\cdot]\}$ para $(i < X; j < Y)$ contemplando o número de neurônios em cada camada, ou seja, X e Y, exceto a última. Neste exemplo hipotético, a estrutura foi dedicada ao reconhecimento de padrões para distinção entre cinco, isto é, classes A, B, C, D e E.

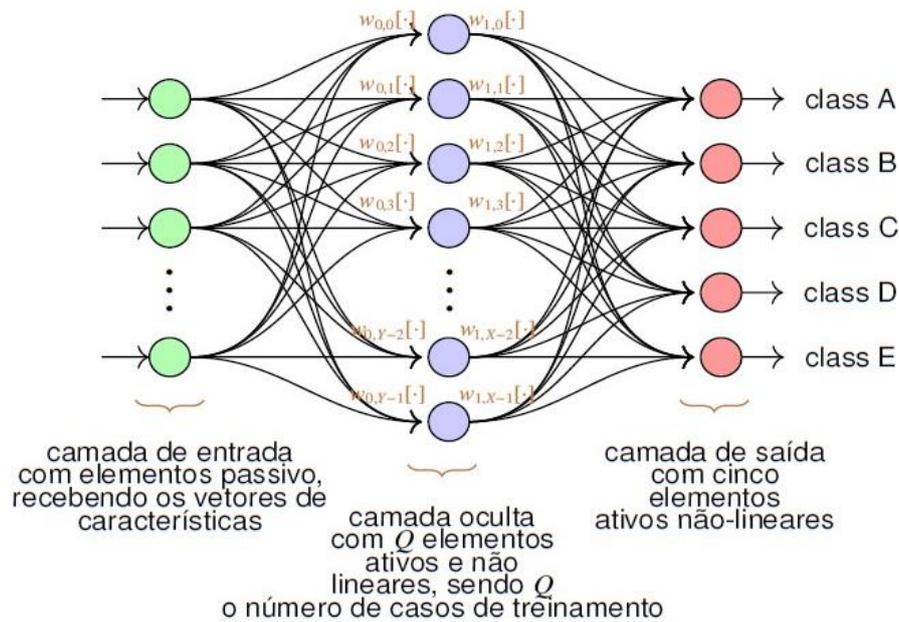


Figura 7. A estrutura usual do perceptron de três camadas, similar ao que está sendo utilizado neste trabalho.

Fonte: A própria pesquisadora

Nesta pesquisa, este modelo foi adaptado para predição, conforme a Figura 8.

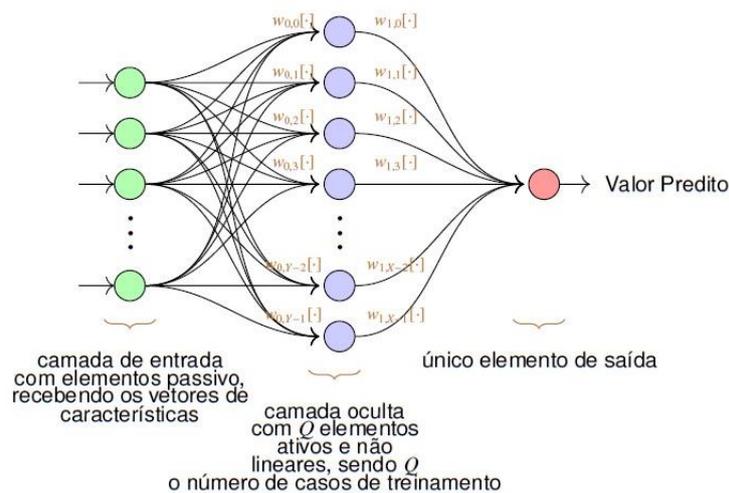


Figura 8. Modelo de rede neural perceptron adaptado para predicao e usado neste trabalho. Notavelmente, a ultima camada possui apenas um conjunto de pesos.

Fonte: A própria pesquisadora

Notavelmente, os Perceptrons multicamadas são frequentemente utilizados para reconhecimento de padrões, entretanto, neste trabalho eles são, diferentemente, usados para realizar previsões. Na Figura 9 consta um exemplo hipotético de um hiperplano modelado por um Perceptron multicamadas, sendo possível notar que a complexidade da envoltória que o define é muito superior àquela existente nos casos da regressão linear e logística.

Com base na estrutura de três camadas exibida na Figura 8, onde a primeira camada é passiva e as outras duas são ativas e não lineares, concebe-se uma função de transferência inteligente capaz de prever o tempo de tratamento de um indivíduo em função das seis características descritas anteriormente. Particularmente, a camada intermediária possui 100 neurônios e a camada de saída possui apenas um.

Todos os neurônios das camadas ativas são baseados em funções

sigmoide normalizadas 9

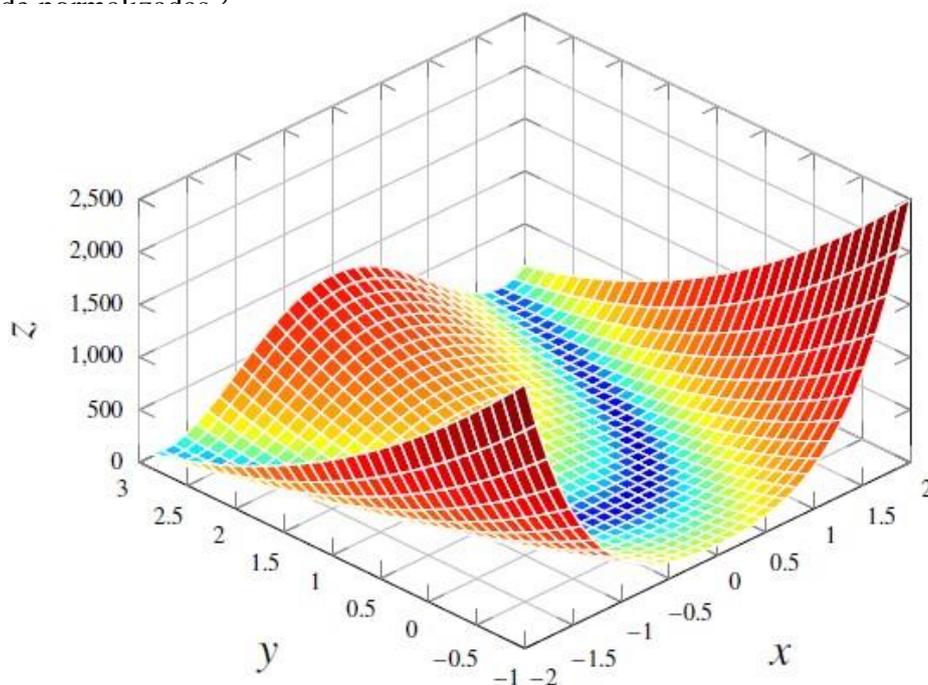


Figura 9. Hiperplano tridimensional hipotético, nas coordenadas $\{x, y, z\}$, capaz de ser modelado por uma estrutura de perceptron multicamadas.

Fonte: A própria pesquisadora

4 RESULTADOS

4 RESULTADOS

Os testes para as três estratégias propostas, ou seja, predição por regressão linear, por regressão logística e por perceptron multicamadas, foram realizados e os resultados constam nas Figuras 10 a 14.

Particularmente, em todos os testes exibidos nas referidas Figuras, considerou-se a estratégia de validação cruzada na qual 80% da base de dados, ou seja, $Q = 1491 \cdot 0.8 = 1193$, foi utilizada para treinamento ou modelamento e os 20% restantes, isto é, $1491 - 1193 = 298$, foram utilizados para os testes, valendo-se de 1000 repetições aleatórias. Nos eixos horizontais estão exibidos o número de ordem do teste, ou caso, e nos eixos verticais estão exibidos o tempo de tratamento em dias, com os casos ordenados em ordem crescente de tempo de tratamento. Além disso, os pontos marcados em azul referem-se ao tempo real de tratamento constante na base de dados. Para contrapô-los, os pontos vermelhos correspondem ao tempo predito pelo algoritmo em questão.

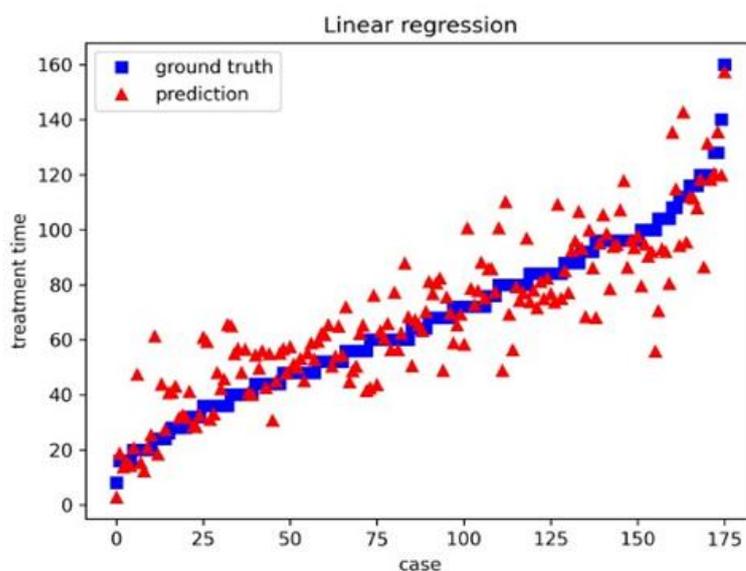


Figura 10. Resultados para o melhor caso de validação cruzada na análise de regressão linear.

Fonte: A própria pesquisadora

A Figura 11 representa uma análise de regressão logística, na qual o eixo vertical foi normalizado de acordo com o domínio da função logística $0 < f(x) = \frac{1}{1+e^{-x}} < 1$. Essa função é comumente utilizada em problemas de classificação binária, mas pode também ser aplicada nos casos em que se deseja modelar uma probabilidade ou uma transformação não linear dos dados. No gráfico apresentado, o eixo vertical, que representa o “tempo de tratamento”, foi normalizado para estar dentro deste intervalo, o que reflete a saída da função logística. Os quadrados azuis representam os valores reais normalizados do tempo de tratamento. Os triângulos vermelhos representam as previsões feitas pelo modelo de regressão logística. O modelo de regressão logística não conseguiu capturar adequadamente a relação entre os casos e os tempos de tratamento no contexto deste conjunto de dados. a análise sugere que o modelo de regressão logística, conforme apresentado, não foi eficaz para prever o tempo de tratamento no contexto dos dados fornecidos. Seria recomendável explorar modelos alternativos ou reavaliar a abordagem de normalização e seleção de características para melhorar o desempenho das previsões.

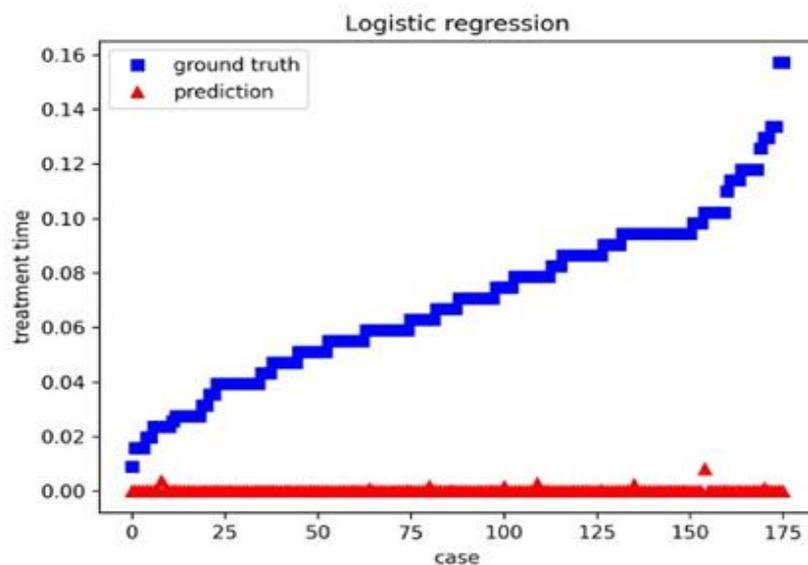


Figura 11. Resultados para o melhor caso de validação cruzada na análise de regressão logística, com o eixo vertical normalizado em função do domínio da função logística, isto é,

Fonte: A própria pesquisadora

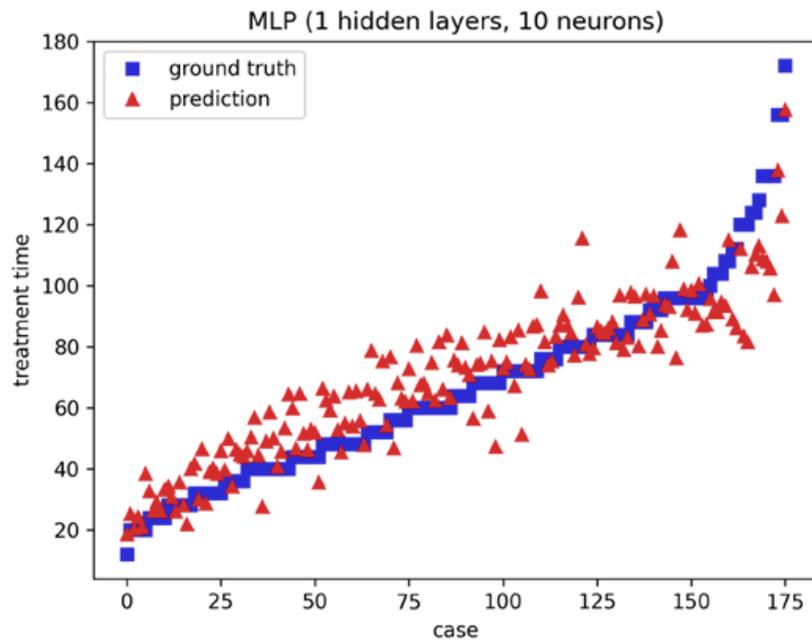


Figura 12. Resultados para o melhor caso de validação cruzada na análise de perceptron com uma camada oculta.

Fonte: A própria pesquisadora

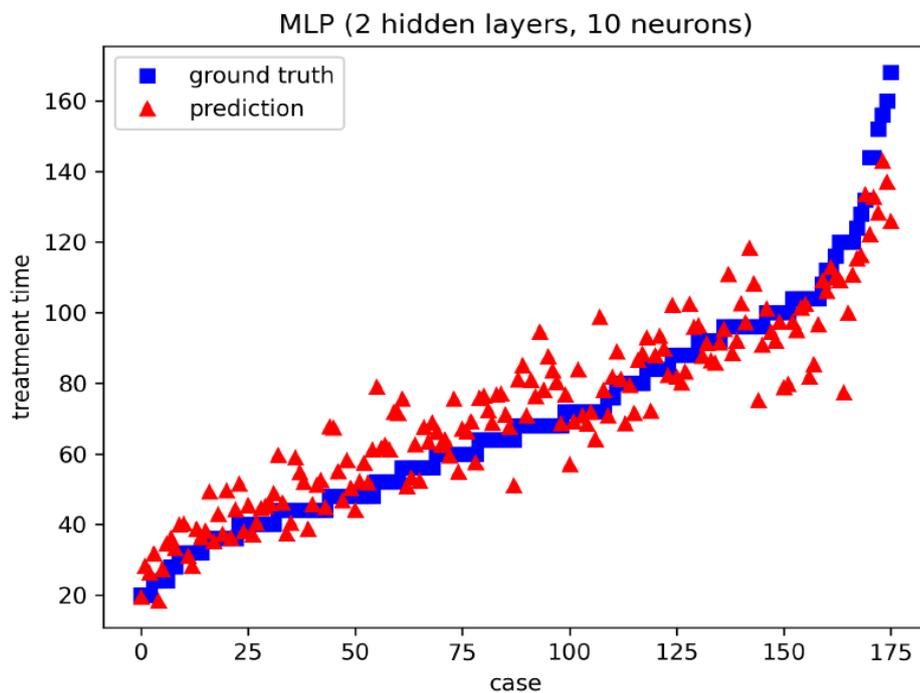


Figura 13. Resultados para o melhor caso de validação cruzada na análise de Perceptron com duas camadas ocultas.

Fonte: A própria pesquisadora

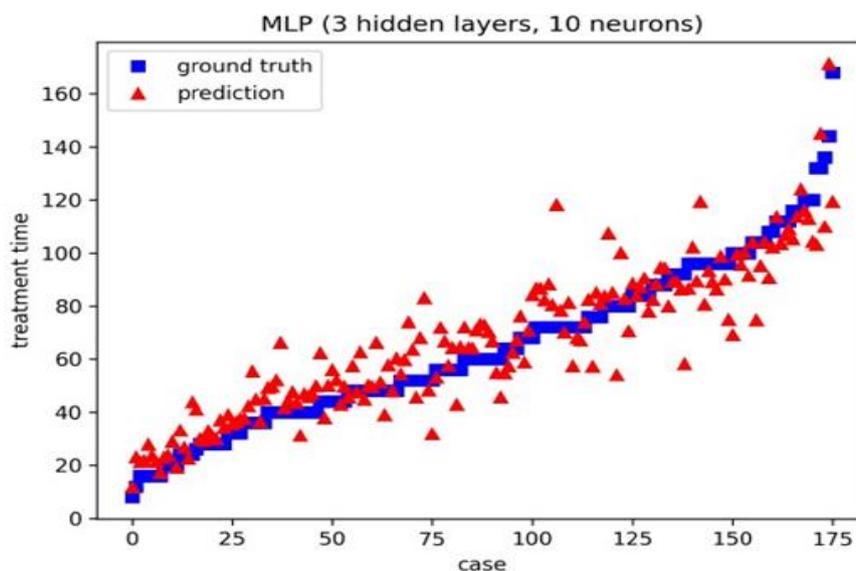


Figura 14. Resultados para o melhor caso de validação cruzada na análise de Perceptron com três camadas ocultas.

Fonte: A própria pesquisadora

Por fim, na Tabela 1, estão exibidos os resultados numéricos mensurados pelo critério do erro quadrático médio para cada um dos casos expostos nas Figuras. Nitidamente, tanto pelo exame das Figuras quanto pelo exame das Tabelas, é fácil perceber que o perceptron com três camadas, isto é, com uma camada oculta, foi o que proporcionou os melhores resultados, isto é, os menores erros locais e, conseqüentemente, o menor erro global. Diferentemente, a regressão logística foi a estratégia que proporcionou os piores resultados. Assim, a principal conclusão é que é possível realizar as predições com baixo e aceitável erro baseadas em um modelo de Perceptron com uma camada oculta tal qual o da Figura 8.

Tabela 1. Erros quadráticos médios para cada uma das estratégias de predição testadas.

Método	Erro Quadrático Médio
Regressão Linear	1,78
Regressão Logística	5,33
Perceptron com uma camada oculta	1,56
Perceptron com duas camadas ocultas	1,61
Perceptron com três camadas ocultas	1,63

A Tabela 1 apresenta os valores dos Erros Quadráticos Médios (MSE) para diferentes métodos de predição testados. O MSE avalia a média dos quadrados das diferenças entre os valores previstos e os observados, sendo que menores valores indicam melhor desempenho preditivo. A regressão linear obteve um MSE de 1.78, demonstrando desempenho satisfatório, mas não superior às demais estratégias. O modelo de regressão logística apresentou o maior MSE (5.33), sugerindo inadequação para esse conjunto de dados, uma vez que é mais apropriado para problemas de classificação. O Perceptron com uma camada oculta registrou o menor MSE (1.56), mostrando-se mais eficaz na previsão, equilibrando complexidade e capacidade de generalização. O modelo com duas camadas ocultas obteve MSE de 161, próximo ao de uma camada, sem ganhos significativos de desempenho. Já o modelo com três camadas ocultas apresentou MSE de 1.63, reforçando que o aumento da complexidade não trouxe melhorias claras e pode ter introduzido ruído. Assim, o Perceptron com uma camada oculta foi a melhor opção, enquanto a regressão logística se mostrou menos recomendada para a tarefa em questão.

A validação cruzada é uma técnica fundamental para avaliar a capacidade de generalização de um modelo. O gráfico da Figura 10 representa os resultados de um caso específico da validação cruzada, no qual o modelo foi treinado em um conjunto de dados e testado em outro. Na figura, os triângulos vermelhos representam as predições feitas pelo modelo com base nos dados de teste (que não foram usados no treinamento), e os quadrados azuis representam os valores reais desses dados.

Quando as predições (vermelho) estão próximas dos valores reais (azul), o modelo está desempenhando bem. No entanto, quanto maior a distância entre esses pontos, maior o erro de predição.

A Figura 12 mostra os resultados de um modelo de Perceptron Multicamadas (MLP) com uma camada oculta de 100 neurônios, comparando o tempo de tratamento real com as

previsões do modelo. A análise envolve a comparação da linha azul (que representa o valor real) com a linha vermelha (que representa as previsões do modelo) ao longo dos diferentes casos analisados. Dizer que este é o melhor caso indica que, entre várias iterações de treinamento e teste com diferentes divisões dos dados, essa configuração específica (uma camada oculta com 100 neurônios) apresentou o melhor desempenho em termos de erro de acurácia. O gráfico mostra que o modelo MLP com uma camada oculta e 100 neurônios tem um desempenho razoável em prever o tempo de tratamento, embora com algumas variações notáveis. A frase melhor caso de validação cruzada indica que essa foi a melhor configuração encontrada durante o processo de otimização do modelo.

A Figura 13 apresentada mostra os resultados de um modelo de Perceptron Multicamadas (MLP) com duas camadas ocultas com 100 neurônios, comparando o tempo de tratamento real com as previsões do modelo. A análise realizada envolve a observação da correspondência entre os valores reais e as predições do modelo, refletida no gráfico. Assim como no caso anterior com uma camada oculta, o modelo com duas camadas ocultas também consegue seguir a tendência geral dos dados reais.

Apesar da similaridade na tendência, observa-se uma variabilidade nas previsões em relação aos valores reais, com alguns desvios notáveis. Isso é visível nas diferenças entre os triângulos vermelhos (predições) e os quadrados azuis (valores reais). Esses desvios indicam a presença de erros, onde o modelo nem sempre acerta o tempo de tratamento com acurácia condizente. A validação cruzada indica que este foi o melhor resultado obtido com essa configuração de duas camadas ocultas ao testar diferentes divisões dos dados. Isso significa que, entre as diferentes tentativas, este modelo teve a melhor acurácia em prever os dados de teste. O modelo MLP com duas camadas ocultas e 100 neurônios em cada camada segue bem a tendência dos dados reais, embora ainda existam erros de predição notáveis. A frase “melhor caso de validação cruzada” reforça que este foi a melhor acurácia observada para esta

configuração específica, embora a adição de uma segunda camada oculta não tenha eliminado completamente as discrepâncias entre as previsões e os valores reais.

A Figura 14 mostra os resultados de um modelo de Perceptron Multicamadas (MLP) com três camadas ocultas com 100 neurônios comparando o tempo de tratamento real com as previsões do modelo. Assim como nas configurações anteriores (com uma e duas camadas ocultas), o modelo com três camadas ocultas segue a tendência geral dos dados reais. Os triângulos vermelhos (previsões) nem sempre estão alinhados com os quadrados azuis (valores reais), indicando a presença de erros de previsão. A dispersão parece mais evidente em pontos onde o tempo de tratamento é maior, o que pode sugerir dificuldades do modelo em prever corretamente casos mais extremos ou complexos. Em comparação com as configurações de uma e duas camadas ocultas, o modelo de três camadas ocultas apresenta uma variação semelhante ou até maior nas previsões em relação aos valores reais. Isso pode indicar que o aumento na complexidade do modelo (mais camadas) não necessariamente melhora a acurácia das previsões e pode, em alguns casos, aumentar a variabilidade.

O termo “melhor caso de validação cruzada” sugere que, dentre várias execuções e divisões dos dados para treinamento e teste, este foi o resultado mais favorável obtido com essa configuração específica de três camadas ocultas. O modelo MLP com três camadas ocultas e 100 neurônios em cada camada continua a seguir a tendência dos dados reais, mas com uma dispersão significativa em torno dos valores reais. O aumento da complexidade do modelo, com a adição de uma terceira camada oculta, não resultou em uma melhoria clara na acurácia, e, em alguns casos, pode ter aumentado a variabilidade dos erros. Portanto, embora o modelo capture a tendência geral dos dados, ele ainda apresenta desafios na acurácia das previsões para casos específicos, particularmente em cenários mais complexos onde o tempo de tratamento é maior.

Essa análise sugere que, ao aumentar a complexidade de um modelo, é essencial verificar se isso realmente resulta em melhor acurácia. Nesse caso, o modelo com três camadas

ocultas não demonstrou uma vantagem clara sobre as configurações mais simples. Complementarmente, para fins de ilustração e conforme exibido na Figura 15, é possível notar uma dentição plenamente alinhada com o uso de *Invisalign*.



Figura 15. Exemplo de um alinhador *Invisalign* na boca de um paciente.

Fonte: Casarim S [homepage na Internet]. Juiz de Fora: Ortodontia e Estética Sergio Casarim; 2017 [acesso em 2024 Out 30]. 9 Vantagens do Aparelho Invisalign®; [aproximadamente 5 telas]. Disponível em: <https://www.ortodontiasergiocasarim.com/9-vantagens-do-aparelho-invisalign/>

5 DISCUSSÃO

Para confrontar os nossos resultados com a literatura, precisamos analisar o desempenho do modelo preditivo no contexto de outras pesquisas sobre a estimativa da duração do tratamento ortodôntico com alinhadores transparentes. Os principais pontos a serem considerados incluem a acurácia do modelo, a metodologia utilizada para validação, e a relevância das variáveis incluídas no processo de modelagem.

Nossos dados mostraram que a utilização de uma divisão de 80% dos dados para treinamento e 20% para teste, com 1.000 repetições aleatórias, é uma estratégia robusta que se alinha com práticas comuns na literatura. A repetição de testes e a aleatoriedade ajudam a garantir que o modelo não esteja superajustado a um subconjunto específico dos dados, aumentando a confiabilidade dos resultados. Um estudo destacou a importância da validação cruzada para reduzir a variância nos resultados e melhorar a generalização dos modelos de predição.³⁰

A literatura frequentemente utiliza o erro quadrático médio (MSE) como métrica de avaliação, como demonstrado por Volovic et al.³¹ A comparação entre os tempos reais e preditos, conforme descrito nos pontos azuis (tempos reais) e vermelhos (tempos preditos), permite avaliar visualmente a eficácia do modelo. Em estudos como o de Prasad et al.³⁰ a precisão preditiva é crucial para avaliar a viabilidade de modelos de IA na ortodontia. O desempenho do modelo pode ser confrontado com outros estudos que também utilizam IA para prever o tempo de tratamento ortodôntico. Por exemplo, um estudo demonstrou que modelos baseados em redes neurais profundas podem prever o tempo de tratamento com uma precisão bastante alta, mas ainda há desafios em casos com variações significativas nos dados clínicos.³²

Os resultados da nossa pesquisa estão bem alinhados com a literatura existente. Os métodos utilizados para validação e os resultados comparativos reforçam a relevância da nossa pesquisa no campo da predição de tempo de tratamento ortodôntico com alinhadores.

Nossos dados mostraram que foi fácil perceber que o perceptron com três camadas, isto é, com uma camada oculta, foi o que proporcionou os melhores resultados, isto é, os menores erros locais e, conseqüentemente, o menor erro global. Diferentemente, a regressão logística foi a estratégia que proporcionou os piores resultados. Assim, a principal conclusão é que é possível realizar as predições com baixo e aceitável erro baseadas em um modelo de Perceptron com uma camada oculta.

Os modelos de Perceptron multicamadas (MLP) são amplamente reconhecidos na literatura por sua capacidade de capturar relações não lineares nos dados, o que muitas vezes resulta em uma melhor acurácia preditiva em comparação com modelos lineares.

Um estudo mostrou que modelos com uma camada oculta podem alcançar um equilíbrio entre complexidade e capacidade de generalização, proporcionando resultados superiores em comparação com modelos mais simples ou mais complexos. Esse equilíbrio é crucial, pois evita o problema de *overfitting*, que pode ocorrer em modelos com muitas camadas.³³

A regressão logística é uma técnica poderosa para tarefas de classificação binária, mas tem limitações quando aplicada a problemas de predição complexos com múltiplas variáveis de entrada que interagem de maneiras não lineares. A literatura aponta que, embora a regressão logística seja intuitiva e fácil de interpretar, ela pode falhar em capturar padrões complexos nos dados, resultando em maiores erros preditivos.

Um outro estudo destacou que a regressão logística pode ser insuficiente em contextos nos quais os dados apresentam alta dimensionalidade ou não linearidade, confirmando os

achados de que o perceptron multicamadas pode superar significativamente a regressão logística em termos de acurácia preditiva.³⁴

A observação de que o Perceptron multicamadas proporcionou menores erros locais e globais é consistente com a literatura, que frequentemente relata que redes neurais, mesmo com estruturas relativamente simples, podem superar modelos lineares clássicos em tarefas de previsão complexas. Em um estudo realizado em 2023, modelos de MLP com uma camada oculta foram avaliados em várias tarefas de previsão e consistentemente demonstraram menores erros em comparação com a regressão logística, especialmente em contextos nos quais

as variáveis possuem interações complexas.³⁵

Nossos dados são consistentes com o que é relatado em estudos sobre redes neurais e regressão logística. O uso de um perceptron com uma única camada oculta demonstra ser uma escolha eficaz para reduzir erros preditivos em problemas complexos, confirmando a vantagem desse modelo sobre a regressão logística em tais cenários.

Nossa pesquisa não mostrou que a validação cruzada é uma técnica fundamental para avaliar a capacidade de generalização de um modelo. O gráfico representa os resultados de um caso específico da validação cruzada, no qual o modelo foi treinado em um conjunto de dados e testado em outro. Os triângulos vermelhos representam as previsões feitas pelo modelo com base nos dados de teste (que não foram usados no treinamento), e os quadrados azuis representam os valores reais desses dados. Quanto à distribuição e dispersão e observando a sobreposição entre os pontos azuis e vermelhos, podemos inferir a acurácia do modelo.

A predição final mostra que, apesar de o modelo linear ter capturado a tendência básica, ele apresenta erros que aumentam em magnitude para certos casos. Isso pode indicar que uma abordagem mais complexa ou uma maior quantidade de dados pode ser necessária para melhorar a precisão das previsões.

Para confrontar os resultados apresentados com a literatura recente, é importante considerar as limitações inerentes à regressão logística, especialmente em contextos nos quais se tenta modelar variáveis contínuas, como o tempo de tratamento, e discutir alternativas metodológicas que possam oferecer melhor desempenho preditivo.

A regressão logística é amplamente reconhecida como uma técnica eficaz para problemas de classificação binária. No entanto, seu uso para prever variáveis contínuas, mesmo após normalização, pode não ser adequado. Um estudo destacou que, apesar da flexibilidade da função logística, sua aplicação em previsões de variáveis contínuas muitas vezes resulta em previsões enviesadas, especialmente quando a distribuição dos dados não se alinha bem com o modelo logístico. Isso se alinha com os resultados apresentados, onde a concentração das previsões em valores próximos a zero indica uma falha na captura da complexidade dos dados.

A normalização dos dados para ajustá-los ao intervalo de $[0, 1]$, como no caso da função logística, é uma prática comum. No entanto, dependendo da distribuição dos dados, essa normalização pode introduzir distorções que afetam a acurácia do modelo.³⁶ Um estudo recente, mostrou que técnicas alternativas de normalização ou a utilização de funções de ativação diferentes podem melhorar a precisão preditiva em cenários onde a regressão logística falha.³⁷

A divergência significativa entre os valores reais normalizados e as previsões do modelo, conforme evidenciado pelos resultados, sugere uma inadequação do modelo de regressão logística para este conjunto de dados. Um estudo discutiu como essa discrepância pode ocorrer devido à incapacidade do modelo em capturar interações complexas entre variáveis, recomendando o uso de técnicas mais avançadas, como redes neurais ou modelos de regressão não lineares, para problemas similares.³⁸

A literatura recente oferece diversas alternativas que podem superar as limitações da regressão logística em cenários como o descrito. Por exemplo, um estudo propôs o uso de modelos baseados em redes neurais profundas para capturar melhor a complexidade dos dados,³⁹ enquanto outro estudo explorou o uso de métodos de ensemble learning para melhorar a precisão preditiva em comparação com métodos lineares simples.⁴⁰

Nossa pesquisa demonstrou que os resultados indicam que a regressão logística, especialmente quando aplicada a variáveis contínuas como o tempo de tratamento, pode não ser a melhor escolha. A exploração de modelos alternativos ou a reavaliação das técnicas de normalização e seleção de características são recomendadas para melhorar o desempenho preditivo.

Nossos resultados mostram a análise de um modelo de Perceptron Multicamadas (MLP) com uma camada oculta e 100 neurônios, comparando o tempo de tratamento real com as previsões do modelo. A análise que realizamos envolve a comparação da linha azul (que representa o valor real) com a linha vermelha (que representa as previsões do modelo) ao longo dos diferentes casos analisados. O modelo seguiu a tendência geral dos dados reais. Há uma variação significativa em torno da linha do valor real. Isso sugere que, embora o modelo capture a tendência geral, ele apresenta um erro considerável em alguns pontos específicos, o que é visível pelos desvios dos triângulos vermelhos (predições) em relação aos quadrados azuis (valores reais). Dizer que este é o “melhor caso” indica que, entre várias iterações de treinamento e teste com diferentes divisões dos dados, essa configuração específica (uma camada oculta com 100 neurônios) apresentou a melhor acurácia em termos de erro de predição. O gráfico mostra que o modelo MLP com uma camada oculta e 100 neurônios tem uma acurácia razoável em prever o tempo de tratamento, embora com algumas variações notáveis. A frase “melhor caso de validação cruzada” indica que essa foi a melhor configuração encontrada durante o processo de otimização do modelo.

Os modelos de MLP, especialmente aqueles com uma camada oculta, são amplamente utilizados em tarefas de previsão em função de sua capacidade de modelar relações não lineares entre as variáveis. Um estudo recente, mostrou que uma configuração com uma camada oculta pode capturar adequadamente a tendência geral dos dados, como observado nos resultados apresentados. No entanto, a presença de variação significativa ao redor da linha de valores reais, como evidenciado pelos desvios das previsões em relação aos valores reais, indica que o modelo pode ter dificuldades em capturar nuances específicas dos dados ou lidar com outliers.⁴¹

A variação significativa observada entre as previsões e os valores reais é um desafio comum em modelos de MLP, especialmente quando o número de neurônios na camada oculta é limitado. De acordo com outro estudo, modelos MLP com uma única camada oculta e um número limitado de neurônios podem ter um desempenho razoável, mas podem falhar em capturar complexidades mais profundas dos dados, resultando em erros maiores para certos casos. Isso sugere que, para melhorar a precisão, pode ser necessário ajustar a arquitetura do modelo, como aumentar o número de neurônios ou adicionar mais camadas ocultas.⁴²

A validação cruzada é uma técnica fundamental para avaliar a capacidade de generalização de um modelo, especialmente para evitar o superajuste (*overfitting*). No contexto dos MLPs, a validação cruzada ajuda a garantir que o modelo performe bem em diferentes

subconjuntos dos dados.⁴³ A frase "melhor caso de validação cruzada" indica que, entre várias configurações testadas, essa foi a que minimizou o erro de previsão. No entanto, a presença de variação significativa em torno dos valores reais sugere que, mesmo no melhor caso, o modelo pode ainda estar limitado em termos de capacidade preditiva, possivelmente devido à simplicidade da arquitetura escolhida.

Para abordar as variações observadas, a literatura sugere várias abordagens. Um estudo de recomenda a experimentação com diferentes arquiteturas de rede, como a introdução de mais camadas ocultas ou a utilização de técnicas de regularização, como *dropout*, para melhorar a generalização do modelo. Além disso, o ajuste fino dos hiperparâmetros, como o número de neurônios ou a taxa de aprendizado, pode contribuir para reduzir o erro de predição em casos específicos.⁴⁴

Nossa pesquisa mostrou que o MLP com uma camada oculta e 100 neurônios tem um desempenho razoável em capturar a tendência geral dos dados, mas ainda enfrenta desafios em prever com acurácia certos casos. A validação cruzada valida a escolha da arquitetura, mas os erros observados indicam que ajustes na complexidade do modelo ou em técnicas de regularização podem ser necessários para melhorar o desempenho preditivo.

Nossos resultados mostraram um modelo de Perceptron Multicamadas (MLP) com duas camadas ocultas e 100 neurônios em cada camada, comparando o tempo de tratamento real com as previsões do modelo. A análise estatística realizada envolveu a observação da correspondência entre os valores reais e as previsões do modelo, refletida no gráfico. Assim como no caso anterior com uma camada oculta, o modelo com duas camadas ocultas também conseguiu seguir a tendência geral dos dados reais. Isso é visível nas diferenças entre os triângulos vermelhos (previsões) e os quadrados azuis (valores reais). Esses desvios indicaram a presença de erros, nos quais o modelo nem sempre acerta o tempo de tratamento corretamente. A validação cruzada indicou que este foi o melhor resultado obtido com essa configuração de duas camadas ocultas ao testar diferentes divisões dos dados. Isso significa que, entre as diferentes tentativas, este modelo teve a melhor acurácia em prever os dados de teste. O modelo MLP com duas camadas ocultas e 100 neurônios em cada camada seguiu bem a tendência dos dados reais, embora ainda existam erros de predição notáveis. A frase “melhor caso de validação cruzada” reforça que este foi o melhor desempenho observado para esta configuração específica,

embora a adição de uma segunda camada oculta não tenha eliminado completamente as discrepâncias entre as predições e os valores reais.

A inclusão de uma segunda camada oculta em um MLP pode, teoricamente, aumentar a capacidade do modelo de capturar padrões mais complexos nos dados. No entanto, como apontado por um estudo, enquanto adicionar mais camadas pode aumentar a flexibilidade do modelo, isso também pode levar a uma maior variabilidade nas predições, especialmente se o número de neurônios por camada for limitado, como no caso apresentado. Os desvios observados entre as predições e os valores reais sugerem que, embora o modelo siga bem a tendência geral, ele não consegue capturar todas as nuances dos dados.⁴⁵ A variabilidade nas predições em relação aos valores reais é um aspecto comum em modelos de MLP, especialmente quando há apenas 100 neurônios por camada oculta. Essa variabilidade pode ser atribuída à incapacidade do modelo de generalizar adequadamente em todos os casos, resultando em erros de predição. Isso é particularmente relevante quando o número de neurônios é insuficiente para capturar toda a complexidade dos dados, o que pode explicar os desvios notáveis observados em alguns casos.⁴³

A validação cruzada é uma prática padrão para avaliar a robustez de um modelo, especialmente em modelos de MLP com múltiplas camadas ocultas. O fato de este ser o “melhor caso de validação cruzada” indica que, entre várias divisões dos dados, essa configuração específica apresentou o melhor desempenho em termos de minimização do erro de predição. No entanto, a validação cruzada, embora útil, não garante que o modelo esteja completamente livre de *overfitting*, especialmente em modelos mais complexos com várias camadas ocultas.⁴³

Embora a adição de uma segunda camada oculta tenha melhorado o desempenho do modelo em relação a uma única camada, os erros persistentes sugerem que ainda há espaço

para melhorias. Um estudo propõe que, em casos como esse, a utilização de técnicas adicionais, como regularização ou ajuste fino dos hiperparâmetros, pode ajudar a reduzir os erros de predição. Além disso, explorar arquiteturas alternativas, como redes profundas com mais neurônios por camada ou mesmo outras técnicas de *machine learning*, pode oferecer melhores resultados.⁴⁶

Os nossos resultados indicaram que, embora o MLP com duas camadas ocultas e 100 neurônios em cada camada capturou bem a tendência geral dos dados, ainda existem erros notáveis nas predições. A validação cruzada confirmou que esta foi a melhor configuração encontrada, mas as discrepâncias entre predições e valores reais sugerem a necessidade de explorar técnicas adicionais ou arquiteturas mais complexas para melhorar o desempenho preditivo.

Os nossos resultados foram realizados com base em um modelo de Perceptron Multicamadas (MLP) com três camadas ocultas, com 100 neurônios. Assim como nas configurações anteriores (com uma e duas camadas ocultas), o modelo com três camadas ocultas seguiu a tendência geral dos dados reais. O modelo MLP com três camadas ocultas e 100 neurônios, continuou a seguir a tendência dos dados reais, mas com uma dispersão significativa em torno dos valores reais. O aumento da complexidade do modelo, com a adição de uma terceira camada oculta, não resultou em uma melhoria clara na acurácia das predições, e, em alguns casos, pode ter aumentado a variabilidade dos erros. Portanto, embora o modelo capturasse a tendência geral dos dados, ele ainda apresenta desafios na precisão das predições para casos específicos, particularmente em cenários mais complexos nos quais o tempo de tratamento é maior. Essa análise indica que, ao aumentar a complexidade de um modelo, é essencial verificar se isso realmente resulta em melhor desempenho. Nesse caso, o modelo com três camadas ocultas não demonstrou uma vantagem clara sobre as configurações mais simples.

A introdução de múltiplas camadas ocultas em um MLP é geralmente feita com a expectativa de que o modelo será capaz de capturar padrões mais complexos nos dados. No entanto, o aumento da complexidade do modelo nem sempre se traduz em uma melhoria na precisão das predições. Em muitos casos, adicionar mais camadas pode levar a um aumento na variabilidade dos erros, especialmente se o número de neurônios por camada for limitado. Isso está em linha com os resultados apresentados, onde o modelo de três camadas ocultas mostrou uma maior dispersão em torno dos valores reais, particularmente em casos extremos.⁴⁷

A maior dispersão observada nas predições em relação aos valores reais, especialmente em casos de tempo de tratamento mais longo, pode indicar que o modelo está tendo dificuldades em generalizar adequadamente para todos os casos. Adicionar camadas ocultas pode, em alguns casos, exacerbar os problemas de *overfitting*, onde o modelo se ajusta excessivamente aos dados de treinamento, mas falha em generalizar para novos dados. Isso pode explicar a variabilidade nas predições observada no modelo de três camadas ocultas.⁴⁸

O fato de que a adição de uma terceira camada oculta não resultou em uma melhoria clara na precisão das predições está alinhado com a literatura, que sugere que há um ponto de retorno decrescente na complexidade do modelo. Aumentar a complexidade do modelo, sem um aumento correspondente na capacidade de dados ou técnicas de regularização adequadas, pode levar a uma piora no desempenho. Isso é especialmente crítico em cenários onde os dados são limitados ou onde a variabilidade intrínseca é alta.⁴⁹

A validação cruzada é uma ferramenta crucial para garantir que o modelo não esteja superajustado aos dados de treinamento. O fato de este ser o “melhor caso de validação cruzada” sugere que, apesar da maior complexidade do modelo, as variações observadas ainda estão dentro do melhor desempenho possível para essa configuração. No entanto, mesmo o melhor caso de validação cruzada pode não garantir que o modelo seja ideal, especialmente se

a complexidade adicional não estiver resultando em melhorias significativas na acurácia preditiva.⁴³

Os resultados indicaram que simplesmente aumentar a complexidade do modelo, como adicionar mais camadas ocultas, não garante necessariamente uma melhor performance. Uma abordagem mais eficaz pode envolver o uso de técnicas de regularização, como *dropout* ou *batch normalization*, para controlar a complexidade do modelo e melhorar sua capacidade de generalização.⁵⁰

Assim, nossa pesquisa mostra que, embora o MLP com três camadas ocultas siga bem a tendência geral dos dados, a adição de complexidade não resultou em uma melhoria clara na acurácia das predições. Na verdade, pode ter aumentado a variabilidade dos erros, especialmente em casos complexos. Isso indica a necessidade de uma abordagem mais equilibrada na escolha da arquitetura do modelo, garantindo que o aumento da complexidade seja acompanhado por técnicas adequadas de regularização e otimização.

A tabela mostra os resultados dos Erros Quadráticos Médios (MSE) para diferentes métodos de predição testados. O MSE é uma métrica que mede a média dos quadrados das diferenças entre os valores previstos pelo modelo e os valores reais. Quanto menor o valor do MSE, melhor é o desempenho do modelo, pois indica que as previsões estão mais próximas dos valores reais. Ao fazer a análise, a regressão linear apresentou um MSE de 1,78, o que é um resultado relativamente bom, mas não o melhor entre as estratégias testadas. Isso indica que, embora o modelo linear capture a tendência geral dos dados, ele não é o mais preciso entre as opções disponíveis. A análise da regressão logística apresentou o maior MSE entre todos os métodos testados, com um valor de 5,33. Isso indica que este modelo tem a maior discrepância entre as previsões e os valores reais, sugerindo que não é adequado para este conjunto específico de dados. Vale ressaltar que a regressão logística é geralmente mais apropriada para problemas

de classificação, não de regressão, o que pode explicar o desempenho inferior. A análise do O modelo de Perceptron com uma camada oculta apresentou o menor MSE, com um valor de 1,56. Isso indica que esta configuração foi a mais eficaz em prever os valores reais entre todas as estratégias testadas, sugerindo um bom equilíbrio entre a complexidade do modelo e a capacidade de generalização. A análise do modelo com duas camadas ocultas teve um MSE de 1,61, ligeiramente superior ao do modelo com uma camada oculta. Embora ainda seja um bom resultado, indica que a adição de uma segunda camada oculta não proporcionou uma melhoria significativa na acurácia e, na verdade, pode ter introduzido uma ligeira complexidade adicional que não se traduziu em melhor acurácia. A análise do modelo com três camadas ocultas apresentou um MSE de 1,63, o que é muito próximo do resultado obtido com duas camadas ocultas. Isso reforça a observação de que aumentar a complexidade do modelo não trouxe benefícios claros em termos de acurácia e pode, na verdade, ter adicionado erro às previsões. Com base nos resultados apresentados, o modelo de Perceptron com uma camada oculta foi o mais eficaz, com o menor MSE (1,56), indicando que esta configuração conseguiu prever os dados com maior acurácia. A regressão linear também teve uma acurácia aceitável, mas inferior ao do Perceptron com uma camada oculta. Em contrapartida, a regressão logística foi a menos eficaz, apresentando o maior MSE, o que a torna a menos recomendada para este problema específico de predição. A adição de camadas ocultas extras nos modelos de Perceptron não resultou em uma melhoria significativa, sugerindo que uma camada oculta já era suficiente para capturar a complexidade dos dados.

Para confrontar os resultados apresentados com a literatura recente, é essencial analisar como diferentes modelos de predição se comportam em termos de Erro Quadrático Médio (MSE) e discutir as implicações de aumentar a complexidade do modelo, especialmente ao adicionar camadas ocultas em redes neurais, como o Perceptron Multicamadas (MLP).

A regressão linear apresentou um MSE de 1,78, o que indica que, embora capture bem a tendência geral dos dados, ela pode ser limitada em termos de precisão para conjuntos de dados mais complexos. Estudo recente, mostrou que a regressão linear é uma escolha sólida para problemas simples e lineares, mas seu desempenho diminui à medida que a complexidade dos dados aumenta. Este resultado sugere que, embora a regressão linear tenha um desempenho razoável, modelos mais complexos, como o Perceptron Multicamadas, podem oferecer maior precisão.⁵¹

A regressão logística apresentou o maior MSE (5,33), reforçando a ideia de que este modelo não é adequado para problemas de predição de valores contínuos, mas sim para classificação binária ou multiclasse. A literatura destaca que a regressão logística é ineficaz para problemas de regressão, pois não foi projetada para prever variáveis contínuas. Este resultado é consistente com a expectativa teórica, indicando que a regressão logística deve ser evitada em cenários como este.⁵²

O modelo de Perceptron com uma camada oculta apresentou o menor MSE (1,56), o que sugere que essa configuração foi a mais eficaz para prever os valores reais. Um estudo confirmou que MLPs com uma única camada oculta podem capturar de maneira eficiente as complexidades dos dados sem introduzir ruído adicional. Esse equilíbrio entre simplicidade e capacidade preditiva é um ponto forte deste modelo, indicando que uma camada oculta é suficiente para capturar as características relevantes dos dados neste caso.

A adição de camadas ocultas extras nos modelos de Perceptron (com duas e três camadas ocultas) resultou em um aumento ligeiro no MSE (1,61 e 1,63, respectivamente). Este resultado sugere que aumentar a complexidade do modelo não trouxe melhorias significativas e, na verdade, pode ter introduzido ruído ou *overfitting*. A literatura demonstra que, em muitos casos, adicionar mais camadas ocultas sem um aumento correspondente na quantidade de dados

ou na regularização pode resultar em uma piora no desempenho do modelo, devido à maior complexidade que não é compensada por um ganho significativo em capacidade preditiva.⁵³

Com base nos resultados apresentados, o modelo de Perceptron com uma camada oculta foi o mais eficaz para este conjunto de dados, conforme medido pelo MSE. Um estudo apoia a ideia de que redes neurais simples, mas bem configuradas, podem superar tanto modelos lineares quanto redes mais complexas quando os dados não justificam a complexidade adicional. A escolha de um modelo mais simples também tem a vantagem de ser mais interpretável e menos propenso a overfitting, especialmente quando o volume de dados é limitado.⁵⁰

Os nossos resultados indicaram que o modelo de Perceptron com uma camada oculta foi o mais eficaz em termos de precisão preditiva, enquanto a regressão logística foi a menos adequada para este problema específico. A adição de camadas ocultas extras não proporcionou melhorias significativas e pode ter introduzido maior variabilidade, reforçando a importância de equilibrar a complexidade do modelo com a capacidade dos dados de suportar essa complexidade.

.

.

6 CONCLUSÃO

Nesta dissertação, após uma etapa de revisão da literatura, apresentou-se uma metodologia para viabilizar o uso de técnicas de IA no planejamento de tratamentos ortodônticos baseados em alinhadores transparentes *Invisaling*.

As implementações permitiram comparar três modelos inteligentes para solucionar o problema em questão, analisando, ainda, o potencial de generalização da base de dados pública disponibilizada pela *Invisaling*. Por conseguinte, os resultados permitiram concluir que é viável realizar as predições baseadas em IA, sem levar em conta imagens da arcada dentária e da cavidade oral dos pacientes, com erro médio quadrático de 1.56, considerado aceitável, com base no perceptron de uma camada oculta.

No âmbito prático, referindo-se ao dia a dia de trabalho no consultório odontológico, entende-se que é perfeitamente possível aplicar a técnica desenvolvida para estimar tempo dos tratamentos ortodônticos necessários para o devido alinhamento da dentição.

Finalmente, registre-se que as maiores dificuldades para o desenvolvimento foram as que se estão relacionadas com as implementações computacionais dos modelos de IA, as quais contaram com o auxílio de profissionais da área de Computação Aplicada.

Destaca-se, ainda, no âmbito dos trabalhos futuros, a intenção de estudar e desenvolver outros modelos computacionais inteligentes, possivelmente baseados em estratégias de aprendizado profundo, assim que a base de dados utilizada, ou outra que venha a ser concebida, possua maior volume de dados, compatibilizando-a com as necessidades de tal estratégia computacional.

A aplicação de IA no planejamento de tratamentos ortodônticos com Invisalign e na enfermagem compartilha muitas semelhanças em termos de coleta de dados, análise, personalização, e monitoramento contínuo. Em ambos os casos, a IA não substitui os profissionais humanos, mas aumenta sua capacidade de fornecer cuidados precisos e personalizados, melhorando os resultados para os pacientes. Este paralelo destaca a importância crescente da IA na personalização e eficiência dos cuidados de saúde, seja em tratamentos dentários ou no gerenciamento geral da saúde dos pacientes.

7 REFERÊNCIAS

1. Alberton E, Maestrelli B, Oliveira JC, Ribeiro DM, Scarparo A. Conhecimento, percepções e práticas de uma equipe de enfermagem sobre saúde bucal em ambiente hospitalar: uma abordagem qualitativa. *Rev Flum Odontol* [periódico na Internet] 2022 Maio/Ago [acesso em 2024 Out 24];2(58):146-58. Disponível em: <https://docs.bvsalud.org/biblioref/2022/08/1391065/58-13.pdf>
2. Siqueira JS, Piardi R, Butze JP. Importancia da assistencia odontologica dentro do ambiente hospitalar. *Rev Odontol Araçatuba* [periódico na Internet]. 2024 Maio/Ago [acesso em 2024 Out 23];45(2):58-64. Disponível em: <https://revaracatuba.odo.br/revista/2024/04/trabalho08.pdf>
3. Silva Junior AC, Xavier IP, Silveira LM, Stabile AM, Cárnio EC, Gusmão JL, et al. Higiene oral: atuação da equipa de enfermagem em ambiente hospitalar. *Rev Enf Ref* [periódico na Internet]. 2020 Jan [acesso em 2024 Out 24];serV(1):e19099-e19099. Disponível em: <https://scielo.pt/pdf/ref/vserVn1/vserVn1a10.pdf>
4. Martins HDD, Brito Junior AF, Araújo GLBD, Silva EM, Ferreira MAF, Medeiros DS, et al. Práticas e percepção da saúde bucal das equipes de enfermagem em internados na unidade de terapia intensiva. *RSBO* [periódico na Internet]. 2021 Dez [acesso em 2024 Out 23];18(2):209-14. Disponível em: <https://periodicos.univille.br/RSBO/article/view/1599/1361>
5. Proffit WR, Fields H, Larson B, Sarver DM. *Contemporary Orthodontics*. 6th ed. Philadelphia (PA): Mosby; 2018.
6. Liu J, Zhang C, Shan Z. Application of Artificial Intelligence in Orthodontics: Current State and Future Perspectives. *Healthcare (Basel)* [periódico na Internet]. 2023 Out [acesso em 2024 Nov 2];11(20):2760. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10606213/pdf/healthcare-11-02760.pdf>
7. Sipiyyaruk K, Kaewsirirat P, Santiwong P. Technology-enhanced simulation-based learning in orthodontic education: A scoping review. *Dental Press J Orthod* [periódico na Internet]. 2023 Jul [acesso em 2024 Out 31];28(3):e2321354. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10353848/pdf/2177-6709-dpjo-28-03-e2321354.pdf>
8. Daud A, Matoug-Elwerfelli M, Daas H, Zahra D, Ali K. Enhancing learning experiences in pre-clinical restorative dentistry: the impact of virtual reality haptic simulators. *BMC Med Educ* [periódico na Internet]. 2023 Dez [acesso em 2024 Nov 2];23(1):948. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC10717008/pdf/12909_2023_Article_4904.pdf
9. Russel S, Norvig P. *Artificial Intelligence: A Modern Approach*. 4th ed. Upper Saddle River (NJ): Pearson; 2020.

10. Xu L, Mei L, Lu R, Li Y, Li H, Li Y. Predicting patient experience of Invisalign treatment: An analysis using artificial neural network. *Korean J Orthod [periódico na Internet]*. 2022 Jul [acesso em 2024 Out 23];52(4):268-77. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9314214/pdf/kjod-52-4-268.pdf>
11. Lombardo L, Arreghini A, Ramina F, Huanca Ghislanzoni LT, Siciliani G. Predictability of orthodontic movement with orthodontic aligners: a retrospective study. *Prog Orthod [periódico na Internet]*. 2017 Nov [acesso em 2024 Out 24];18(1):35. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC5682257/pdf/40510_2017_Article_190.pdf
12. Morales-Burruezo I, Gandía-Franco JL, Cobo J, Vela-Hernández A, Bellot-Arcís C. Arch expansion with the Invisalign system: Efficacy and predictability. *PLoS One [periódico na Internet]*. 2020 Dez [acesso em 2024 Out 23];15(12):e0242979. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC7728268/pdf/pone.0242979.pdf>
13. Faber J, Faber C, Faber P. Artificial intelligence in orthodontics. *APOS Trends Orthod*. 2019;9(4):201-5.
14. Rao GKL, Srinivasa AC, Iskandar YHP, Mokhtar N. Identification and analysis of photometric points on 2D facial images: a machine learning approach in orthodontics. *Health Technol [periódico na Internet]*. 2019 Mar [acesso em 2024 Out 23];9:715-24. Disponível em: <https://link.springer.com/content/pdf/10.1007/s12553-019-00313-8.pdf>
15. Kök H, Acilar AM, İzgi MS. Usage and comparison of artificial intelligence algorithms for determination of growth and development by cervical vertebrae stages in orthodontics. *Prog Orthod [periódico na Internet]*. 2019 Nov [acesso em 2024 Out 24];20(1):41. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC6856254/pdf/40510_2019_Article_295.pdf
16. Auconi P, McNamara JA Jr, Franchi L. Computer-aided heuristics in orthodontics. *Am J Orthod Dentofacial Orthop [periódico na Internet]*. 2020 Dez [acesso em 2024 Out 23];158(6):856-67. Disponível em: <https://doi.org/10.1016/j.ajodo.2019.10.018>
17. Bichu YM, Hansa I, Bichu AY, Premjani P, Flores-Mir C, Vaid NR. Applications of artificial intelligence and machine learning in orthodontics: a scoping review. *Prog Orthod [periódico na Internet]*. 2021 Jul [acesso em 2024 Out 24];22(1):18. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC8255249/pdf/40510_2021_Article_361.pdf
18. Kunz F, Stellzig-Eisenhauer A, Zeman F, Boldt J. Artificial intelligence in orthodontics: Evaluation of a fully automated cephalometric analysis using a customized convolutional neural network. *J Orofac Orthop [periódico na Internet]*. 2020 Jan [acesso em 2024 Out 23];81(1):52-68. Disponível em: <https://link.springer.com/content/pdf/10.1007/s00056-019-00203-8.pdf>
19. Chen S, Wang L, Li G, Wu TH, Diachina S, Tejera B, et al. Machine learning in orthodontics: Introducing a 3D auto-segmentation and auto-landmark finder of CBCT images to assess maxillary constriction in unilateral impacted canine patients. *Angle Orthod [periódico na Internet]*. 2020 Jan [acesso em 2024 Out 23];90(1):77-84.

- Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC8087054/pdf/i0003-3219-90-1-77.pdf>
20. Yamashiro T, Ko CC. Artificial intelligence and machine learning in orthodontics. *Orthod Craniofac Res* [periódico na Internet]. 2021 Dez [acesso em 2024 Out 24];24 Suppl 2):3-5. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/ocr.12543>
 21. Monill-González A, Rovira-Calatayud L, d'Oliveira NG, Ustrell-Torrent JM. Artificial intelligence in orthodontics: Where are we now? A scoping review. *Orthod Craniofac Res* [periódico na Internet]. 2021 Dez [acesso em 2024 Out 24];24 Suppl 2:6-15. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/ocr.12517>
 22. Retrouvey JM. The role of AI and machine learning in contemporary orthodontics. *APOS Trends Orthod* [periódico na Internet] 2021 Jan/Mar [acesso em 2024 Out 23];11(1):74-80. Disponível em: <https://apospublications.com/view-pdf/?article=42256d691b8a5c5a1e4ff5fb5a55fd88UgjUVmdjeCE=>
 23. Leonardi R, Lo Giudice A, Isola G, Spampinato C. Deep learning and computer vision: two promising pillar, powering the future in orthodontics. *Semin Orthod* [periódico na Internet]. 2021 Jun [acesso em 2024 Out 23];27(2):62-9. Disponível em: <https://doi.org/10.1053/j.sodo.2021.05.002>
 24. Wang H, Minnema J, Batenburg KJ, Forouzanfar T, Hu FJ, Wu G. Multiclass CBCT Image Segmentation for Orthodontics with Deep Learning. *J Dent Res* [periódico na Internet]. 2021 Ago [acesso em 2024 Out 24];100(9):943-9. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC8293763/pdf/10.1177_00220345211005338.pdf
 25. Mohammad-Rahimi H, Nadimi M, Rohban MH, Shamsoddin E, Lee VY, Motamedian SR. Machine learning and orthodontics, current trends and the future opportunities: A scoping review. *Am J Orthod Dentofacial Orthop* [periódico na Internet]. 2021 Ago [acesso em 2024 Out 23];160(2):170-92. Disponível em: <https://doi.org/10.1016/j.ajodo.2021.02.013>
 26. Kök H, Izgi MS, Acilar AM. Determination of growth and development periods in orthodontics with artificial neural network. *Orthod Craniofac Res* [periódico na Internet]. 2021 Dez [acesso em 2024 Out 23];24 Suppl 2:76-83. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1111/ocr.12443>
 27. Narayan H, Gandedkar, Matthew T Wong, and M Ali Darendeliler. Role of virtual reality (vr), augmented reality (ar), and artificial intelligence (ai) in tertiary education and research of orthodontics: an insight. *Semin Orthod* [periódico na Internet]. 2021 Jun [acesso em 2024 Out 24];27(2):69-77. Disponível em: <https://www.semortho.com/action/showPdf?pii=S1073-8746%2821%2900026-8>
 28. Invisalign. Invis is straightening teeth [homepage na Internet]. San Jose: Align Technology, Inc.; 2024 [acesso em 2024 Nov 2]. Disponível em: <https://www.invisalign.com>

29. Helene O. Metodo dos Minimos Quadrados. São Paulo: Livraria da Fisica; 2013.
30. Prasad J, Mallikarjunaiah DR, Shetty A, Gandedkar N, Chikkamuniswamy AB, Shivashankar PC. Machine Learning Predictive Model as Clinical Decision Support System in Orthodontic Treatment Planning. *Dent J (Basel)* [periódico na Internet]. 2022 Dez [acesso em 2024 Out 26];11(1):1. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC9858447/pdf/dentistry-11-00001.pdf>
31. Volovic J, Badirli S, Ahmad S, Leavitt L, Mason T, Bhamidipalli SS, et al. A Novel Machine Learning Model for Predicting Orthodontic Treatment Duration. *Diagnostics (Basel)* [periódico na Internet]. 2023 Ago [acesso em 2024 Out 29];13(17):2740. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC10486486/pdf/diagnostics-13-02740.pdf>
32. Surendran A, Daigavane P, Shrivastav S, Kamble R, Sanchla AD, Bharti L, Shinde M. The Future of Orthodontics: Deep Learning Technologies. *Cureus* [periódico na Internet]. 2024 Jun [acesso em 2024 Out 26];16(6):e62045. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11234326/pdf/cureus-0016-00000062045.pdf>
33. Ramsey SM, Bergtold JS. Examining Inferences from Neural Network Estimators of Binary Choice Processes: Marginal Effects, and Willingness-to-Pay. *Comput Econ* [periódico na Internet]. 2021 Jun [acesso em 2024 Out 26];58:1137-65. Disponível em: <https://link.springer.com/content/pdf/10.1007/s10614-020-09998-w.pdf>
34. Lewis RM, Battey HS. On inference in high-dimensional logistic regression models with separated data. *Biometrika* [periódico na Internet]. 2024 Set [acesso em 2024 Out 25];(111):989–1011. Disponível em: <https://academic.oup.com/biomet/article/111/3/989/7338235>
35. Manno, A., Rossi, F., Smriglio, S. et al. Comparing deep and shallow neural networks in forecasting call center arrivals. *Soft Comput* [periódico na Internet]. 2023 Abr [acesso em 2023 Abr 26];27(18):12943-957. Disponível em: <https://link.springer.com/content/pdf/10.1007/s00500-022-07055-2.pdf>
36. Peláez-Rodríguez C, Torres-López R, Pérez-Aracil J, López-Laguna N, Sánchez-Rodríguez S, Salcedo-Sanz S. An explainable machine learning approach for hospital emergency department visits forecasting using continuous training and multi-model regression. *Comput Methods Programs Biomed* [periódico na Internet]. 2024 Mar [acesso em 2024 Out 26];245:108033. Disponível em: <https://doi.org/10.1016/j.cmpb.2024.108033>
37. Xiao TZ, Bamler R, Schölkopf B, Liu W. Verbalized Machine Learning: Revisiting Machine Learning with Language Models. In: *Proceedings of the 1st Workshop on In-Context Learning at the 41st International Conference on Machine Learning, Vienna, Austria. 2024.*
38. Georgios Charizanos, Haydar Demirhan, and Duygu İçen. A monte carlo fuzzy logistic regression framework against imbalance and separation. *Inf Sci* [periódico na Internet]. 2024 Jan [acesso em 2024 Out 29];655(C):119893. Disponível em: <https://doi.org/10.1016/j.ins.2023.119893>

39. Zhao X, Wang L, Zhang Y, Han X, Deveci M, Parmar M. A review of convolutional neural networks in computer vision. *Artif Intell Ver* [periódico na Internet]. 2024 Mar [acesso em 2024 Out 26];57:99. Disponível em: <https://link.springer.com/content/pdf/10.1007/s10462-024-10721-6.pdf>
40. Mahesh TR, Vinoth Kumar V, Vivek V, Karthick Raghunath KM, Sindhu Madhuri G. Early predictive model for breast cancer classification using blended ensemble learning. *Int J Syst Assur Eng Manag* [periódico na Internet]. 2024 Jun [acesso em 2024 Out 29];15(1):188-97. Disponível em: <https://link.springer.com/content/pdf/10.1007/s13198-022-01696-0.pdf>
41. Ali Ali AS, Jazaei F, Babakhani P, Ashiq MM, Bakhshae A, Waldron B. An overview of deep learning applications in groundwater level modeling: Bridging the gap between academic research and industry applications. *Appl Comput Intell Soft Comput* [periódico na Internet]. 2024 Jan [acesso em 2024 Out 29];(1):9480522. Disponível em: <https://onlinelibrary.wiley.com/doi/epdf/10.1155/2024/9480522>
42. Ziming Li, Bin Chen, Shengbiao Wu, Mo Su, Jing M. Chen, and Bing Xu. Deep learning for urban land use category classification: A review and experimental assessment. *Remote Sens Environ* [periódico na Internet]. 2024 Set [acesso em 2024 Out 26];311:114290. Disponível em: <https://doi.org/10.1016/j.rse.2024.114290>
43. Hayashi-Sakai S, Nishiyama H, Hayashi T, Sakai J, Shimomura-Kuroki J. Deep learning with convolution neural network detecting mesiodens on panoramic radiographs: comparing four models. *Odontology* [periódico na Internet]. 2024 Jul [acesso em 2024 Out 29]. Online ahead of print. Disponível em: <https://link.springer.com/content/pdf/10.1007/s10266-024-00980-8.pdf>
44. Pascal I. A novel Swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images. *Int J Mach Learn Cyber* [periódico na Internet]. 2024 Mar [acesso em 2024 Out 30];(15):3579-97. Disponível em: <https://link.springer.com/content/pdf/10.1007/s13042-024-02110-w.pdf>
45. Mfetoum IM, Ngoh SK, Molu RJJ, N de Kenfack BF, Onguene R, Naoussi SRD, et al. A multilayer perceptron neural network approach for optimizing solar irradiance forecasting in Central Africa with meteorological insights. *Sci Rep* [periódico na Internet]. 2024 Fev [acesso em 2024 Out 29];14(1):3572. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC10861485/pdf/41598_2024_Article_54181.pdf
46. Gemechu WF, Sitek W, Batalha GF. Improving Hardenability Modeling: A Bayesian Optimization Approach to Tuning Hyperparameters for Neural Network Regression *Appl Sci* [periódico na Internet]. 2024 Mar [acesso em 2024 Out 28];14(6):[aproximadamente 31 telas]. Disponível em: <https://www.mdpi.com/2076-3417/14/6/2554>
47. Hasnain MA, Malik H, Asad MM, Sherwani F. Deep learning architectures in dental diagnostics: a systematic comparison of techniques for accurate prediction of dental disease through x-ray imaging. *Int J Intell Comput Cybern.*2024;17:161-80.
48. Hon KW. Artificial neural networks. In: *Technology and Security for Lawyers and*

-
- Other Professionals. Dewey Court Northampton (MA): Edward Elgar Publishing; 2024. p. 490-511.
49. Jingyang Li and Guoqiang Li. The triangular trade-off between robustness, accuracy and fairness in deep neural networks: A survey. *ACM Comput Surv* [periódico na Internet]. 2024 Fev [acesso em 2024 Out 30]. In Press. Disponível em: <https://dl.acm.org/doi/pdf/10.1145/3645088>
 50. Miikkulainen R, Liang J, Meyerson E, Rawal A, Fink D, Francon O, et al. Evolving deep neural networks. In: Kozma R, Alippi C, Choe Y, Morabito FC, editores. *Artificial Intelligence in the Age of Neural Networks and Brain Computing*. Amsterdam: Elsevier; 2019. p. 293-312.
 51. Zhou W, Yan Z, Zhang L. A comparative study of 11 non-linear regression models highlighting autoencoder, DBN, and SVR, enhanced by SHAP importance analysis in soybean branching prediction. *Sci Rep* [periódico na Internet]. 2024 Mar [acesso em 2024 Out 30];14(1):5905. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC10928191/pdf/41598_2024_Article_55243.pdf
 52. Muse AH, Mwalili S, Ngesa O, Chesneau C, Al-Bossly A, El-Morshedy M. Bayesian and Frequentist Approaches for a Tractable Parametric General Class of Hazard-Based Regression Models: An Application to Oncology Data. *Mathematics* [periódico na Internet]. 2022 Out [acesso em 2024 Out 29];10(20):3813.. Disponível em: <https://www.mdpi.com/2227-7390/10/20/3813>
 53. Beltran-Royo C, Llopis-Ibor L, Pantrigo JJ, Ramírez I. Dc neural networks avoid overfitting in one-dimensional nonlinear regression. *Knowledge-Based Systems* [periódico na Internet]. 2024 Mar [acesso em 2024 Out 30];283:111154. Disponível em: <https://doi.org/10.1016/j.knosys.2023.111154>

