

Estudo Avaliativo dos Biomarcadores Mais Influentes Para Doença de Alzheimer Com K-means e Floresta Aleatória

Fernanda Maria Lima Fernandes * Larissa Pereira Costa **
Suellen Mary Marinho dos Santos Andrade ***
José Maurício Ramos de Souza Neto **** José Hélio Bento da Silva †
Juan Moises Maurício Villanueva ‡

* Departamento de Engenharia Elétrica, Universidade Federal da Paraíba, PB, (e-mail: fernanda.fernandes@cear.ufpb.br).

** Departamento de Fisioterapia, Universidade Federal da Paraíba, PB, (e-mail: costaplarissa29@gmail.com)

*** Departamento de Fisioterapia, Universidade Federal da Paraíba, PB, (e-mail: suellenandrade@gmail.com)

**** Departamento de Engenharia Elétrica, Universidade Federal da Paraíba, PB, (e-mail: mauricio@cear.ufpb.br)

† Departamento de Engenharia Elétrica, Universidade Federal da Paraíba, PB, (e-mail: jose.dasilva@cear.ufpb.br)

‡ Departamento de Engenharia Elétrica, Universidade Federal da Paraíba, PB, (e-mail: jmauricio@cear.ufpb.br)

Abstract: Alzheimer's is a neurodegenerative disease that is related not only to memory, but to impairment of other important mental functions. From this perspective, the health care sector has shown one of the largest increases in the amount of digital data, which has led to the emergence of different techniques for exploring and treating them. These technologies are of fundamental importance for the health professional to ensure more accurate diagnoses, in addition to optimizing time, reducing costs and choosing the best interventions in the clinical decision-making process. In addition to contributing to a broader understanding of information and detection of hidden patterns. Therefore, a machine learning model was proposed to identify the best functional biomarkers of cognitive response to Transcranial Direct Current Stimulation (tDCS) capable of differentiating responders and non-responders to treatment.

Resumo: Alzheimer é uma doença neurodegenerativa que tem relação não apenas com a memória, mas com comprometimento de outras funções mentais importantes. Sob essa perspectiva, o setor de cuidado à saúde apresentou um dos maiores crescimentos na quantidade de dados digitais, o que propiciou o surgimento de diferentes técnicas para a exploração e tratamento deles. Essas tecnologias são de fundamental importância para que o profissional de saúde possa garantir diagnósticos mais precisos, além de otimizar tempo, reduzir custos e escolher as melhores intervenções no processo de tomada de decisão clínica. Além de contribuir para uma compreensão mais ampla das informações e detecção de padrões ocultos. Sendo assim, foi proposto um modelo de aprendizado de máquina para identificar os melhores biomarcadores funcionais de resposta cognitiva à Estimulação Transcraniana por Corrente Contínua (ETCC) capazes de diferenciar respondedores e não respondedores ao tratamento.

Keywords: Machine Learning; Artificial Intelligence; Alzheimer; Neuroscience; Prediction.

Palavras-chaves: Aprendizado de Máquina; Inteligência Artificial; Alzheimer; Neurociência; Predição.

1. INTRODUÇÃO

A Doença de Alzheimer (DA) é definida como uma desordem neurodegenerativa que se apresenta com declínio progressivo da memória associado com uma degradação gradativa das funções cognitivas, que afetam a autonomia e influenciam negativamente a realização de atividades no cotidiano dos pacientes Knopman et al. (2021); Marceglia et al. (2016).

Projeta-se que existam hoje aproximadamente 50 milhões de pessoas vivendo com essa demência e que em 2050 esse número possa chegar a 152 milhões, o que equivale a uma pessoa diagnosticada a cada 3 segundos. Diante disso, a DA é considerada a forma mais comum de demência no mundo, representando cerca de 60 a 70% de todos os casos registrados Huang et al. (2020).

Nesse cenário, diante de uma progressão acelerada da doença, terapias não farmacológicas vêm sendo investigadas. Elas surgem como uma alternativa de suporte a fim de colaborar com o tratamento já utilizado nas abordagens tradicionais Andrade et al. (2022). As evidências sugerem que essas intervenções podem produzir plasticidade neural, possibilitando ganhos significativos na cognição, ou até mesmo na prevenção e/ou atraso na progressão causada pela DA. Dentre elas, encontramos a Estimulação Transcraniana por Corrente Contínua (ETCC) e a Intervenção Cognitiva (IC).

A ETCC é uma técnica não invasiva de neuromodulação, que consiste na aplicação de uma corrente elétrica contínua de baixa intensidade, capaz de modular a atividade neuronal através da regulação da frequência de potenciais de ação desencadeados Brunoni et al. (2012); Fregni et al. (2021). Enquanto a IC é considerada uma modalidade pertencente a reabilitação neuropsicológica, caracterizada por uma intervenção que busca estimular funções cognitivas Buschert et al. (2010); Sá et al. (2019).

Mesmo diante de inúmeros benefícios, os achados sobre a ETCC em pacientes com DA ainda são heterogêneos, e seus parâmetros de estimulação ideais no Alzheimer ainda permanecem inexplorados. Diante disso, estudos na área devem considerar os fatores físicos de variabilidade (idade, sexo, variações anatômicas) e de parâmetros técnicos (tamanho, orientação e montagem dos eletrodos, fluxo da corrente e número de sessões) para elaboração de novos protocolos de ETCC, uma vez que esses interagem para afetar a resposta à estimulação Lefaucheur et al. (2017); Guerra et al. (2020).

Desse modo, percebe-se a necessidade de uma maior compreensão de seus mecanismos, para que ela possa ser utilizada com mais especificidade De Berker et al. (2013); Nascimento et al. (2020). Desse modo, alguns recursos podem ser usados para guiar a terapia de maneira eficaz e individualizada, e nesse cenário, destacamos a Inteligência Artificial (IA), em especial, a Aprendizagem de Máquina.

A literatura já apresenta estudos que utilizam essa abordagem em conjunto com a ETCC para que essa seja capaz de orientar intervenções mais personalizadas, a fim de prever respostas a terapêutica Kambeitz et al. (2020) ou para investigar os parâmetros de dosagem adequados Albizu et al. (2020). Além disso, essa técnica também é capaz de identi-

ficar previamente quais pacientes serão mais responsivos ao tratamento, possibilitando assim, um subagrupamento da amostra de forma a reduzir a sua variância, e assim, elevar a eficácia da estimulação nos participantes do estudo Boyd et al. (2017). Com isso, percebe-se a escassez de trabalhos relacionados ao estudo da identificação dos biomarcadores mais influentes na diferenciação dos respondedores e não respondedores ao tratamento.

2. FUNDAMENTOS TEÓRICOS

Com o aumento de dados digitais no setor de saúde, técnicas de aprendizado de máquina têm sido muito utilizadas no campo da neurociência cognitiva. Além disso, árvores de decisão e algoritmos de floresta aleatória mostraram resultados proeminentes para classificação e prognóstico de padrões baseados em eletroencefalograma (EEG) em vários estudos que podem ser encontrados em Amorim et al. (2019) e Al Zoubi et al. (2018). Ainda nesse contexto, métodos não supervisionados têm sido usados para visualizar o comportamento de agrupamento do espaço de recursos multidimensional como em Hasanzadeh et al. (2019) e Khodayari-Rostamabad et al. (2013).

2.1 Árvore de Decisão (Decision Tree)

A Árvore de Decisão é um dos métodos de aprendizado supervisionado mais práticos e amplamente utilizados. Ele cria um mecanismo de previsão por meio de dados de treinamento conhecidos e, em seguida, aplica esse conhecimento para prever resultados para dados não rotulados. Esse processo identifica maneiras de dividir o conjunto de dados com base no valor de seus dados e variáveis. Em cada nó interno, os dados são separados em duas ramificações e continuam a se dividir até atingir o agrupamento final conhecido como nó folha Breiman et al. (2017); Cutler et al. (2012), conforme mostrado na figura 1.

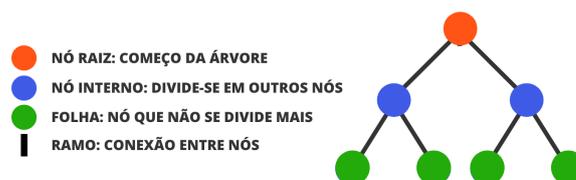


Figura 1. Representação gráfica de uma árvore de decisão

Modelos de árvore são usados para problemas de classificação e regressão Myles et al. (2004), e sua aplicabilidade é realizada em diversos setores para prever as causas mais prováveis de distúrbios florestais, realizar análise de detecção de fraude e, em nosso caso, detecção precoce de comprometimento cognitivo.

O algoritmo para construção de árvores de decisão geralmente aumenta sua complexidade junto com o número de divisões. Com efeito, o desempenho do modelo depende de selecionar as melhores variáveis e saber quando parar de dividir. Para isso, são levadas em consideração todas as características e a melhor é escolhida com base em uma função de custo Breiman et al. (2017).

$$\sum p_k \times (1 - p_k), \quad (1)$$

Como se trata de um modelo de classificação, o ganho de informação de cada divisão é calculado por uma pontuação de Gini, onde o termo "pk" é a proporção das mesmas entradas de classe k em um determinado nó, como na Equação (1). Ela mede a quantidade de importância que a variável fornece sobre uma classe em cada iteração no nó interno que continuará a dividir até que o ganho de informação seja zero, o que significa que o nó atingiu a pureza que estamos procurando e contém todas as entradas da mesma classe Cutler et al. (2012).

2.2 Floresta Aleatória (Random Forest)

Floresta Aleatória é um algoritmo de aprendizado de máquina que combina preditores de árvore de decisão de forma que cada uma contenha valores independentes de vetores aleatórios igualmente distribuídos para todas as árvores da floresta Breiman (2001). Fornecendo, como consequência, um modelo mais preciso com previsão estável cuja principal característica é a tendência de evitar o overfitting. Tal fato ocorre quando um modelo avalia um conjunto de dados muito bem no treinamento, mas sua performance é ineficaz para prever novos resultados.

Essa técnica atua construindo árvores menores que diferem umas das outras pelo conjunto de dados aleatoriamente escolhidos, isto é, cada árvore é alocada em uma amostra de *bootstrap* independente da original usada para criar o banco de dados no qual as árvores são construídas. A cada iteração, as variáveis dessas subamostra são avaliadas para identificar aquela que providencia maior informação de ganho, garantindo que a divisão da árvore ocorra com o menor percentual de erro. Dessa forma, para problemas de classificação, o resultado final é obtido por meio da votação das predições individuais de cada árvore Cutler et al. (2012), como é visto na Figura 2.

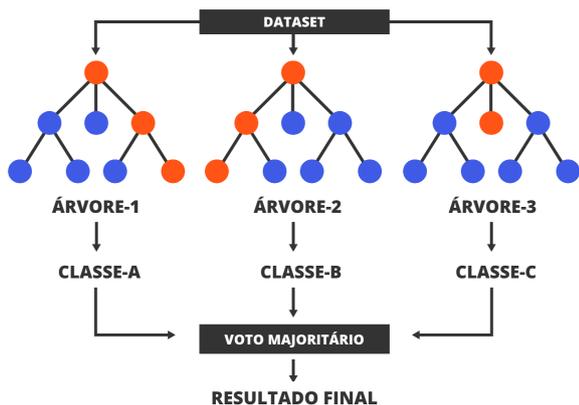


Figura 2. Representação gráfica de uma floresta aleatória

Ainda no que tange as predições, as florestas aleatórias apresentam um ótimo funcionamento, uma vez que são capazes de modelar iterações e lidam bem com amostras de dados grandes. No entanto, em determinadas situações, alguns hiperparâmetros podem ser otimizados para melhorar a performance do algoritmo, como o número de variáveis predictoras em cada nó, a quantidade de árvores

do modelo e o tamanho mínimo que um nó deve ter para se dividir Probst et al. (2019); Cutler et al. (2012).

2.3 K-means

K-means é um método de clusterização não supervisionado do aprendizado de máquina, isto é, ele explora maneiras de agrupar a distribuição de dados não rotulados com base em características similares Ahmed et al. (2020). Sendo o *K-means* uma técnica fundamentada em distância, na qual quanto mais próximos os dados, mais prováveis de pertencer ao mesmo grupo (*cluster*).

Com isso, a separação dos dados se dá inicialmente por K números centrais do cluster (*centroids*) que são escolhidos aleatoriamente. Calcula-se, então, a distância de cada ponto do banco de dados com relação aos (*centroids*) inicialmente selecionados. Tendo isso em vista, uma vez que todos os objetos foram assimilados, as posições dos (*centroids*) são recalculadas com base na média de posição daquele determinado (*cluster*) e seus pontos. Sendo o processo repetido até que não haja mais nenhuma mudança, como é possível observar na Figura 3 Jain et al. (1999).

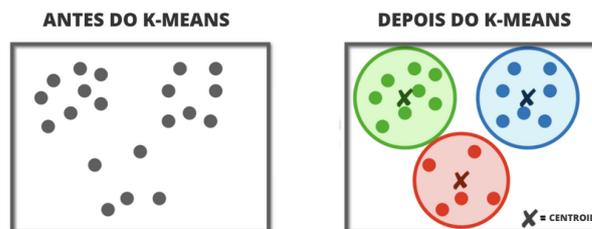


Figura 3. Representação gráfica do K-means

3. BANCO DE DADOS

O conjunto de dados de Alzheimer utilizado neste estudo foi coletado como parte de um ensaio clínico randomizado, duplo cego, controlado por placebo Andrade et al. (2018). Os pacientes foram recrutados de centros de referência clínica para o tratamento da DA. O estudo foi aprovado pelo Comitê de Ética da Instituição e os pacientes e seus responsáveis forneceram o consentimento para participação no estudo. O ensaio clínico foi composto por 60 pacientes, com diagnóstico de DA, sendo diagnosticados 30 respondedores e 30 não respondedores.

A gravação das medidas do EEG se deu em um ambiente controlado na condição de silêncio, com os pacientes sentados, utilizando o amplificador modelo actiCHamp da BrainVision Products (2018) de 32 canais, com taxa de amostragem de 200 Hz, bandpass de 1–70 Hz e notch filter de 60 Hz. Os eletrodos foram posicionados no escalpo, sendo os canais Fc1, Fc2, Fc5, Fc6, Fp1, Fp2, F3, F4, F7, F8, FT9, FT10, C3, C4, CP1, CP2, CP5, CP6, T7, T8, P3, P4, P7, P8, O1 e O2 utilizados para o monitoramento das regiões pré-frontal, frontal, parietal, temporal e occipital.

Os dados foram registrados durante aproximadamente 4 minutos, alternando períodos de repouso com os olhos fechados com períodos de alerta quando os olhos estavam abertos, cada período com duração de 1 minuto. Ainda nesse contexto, artefatos quanto a sintomas de sonolência

ou sono foram removidos pelo BrainVision Analyser. Após a segmentação dos dados, foi aplicada a Transformada Rápida de Fourier para o cálculo da potência das bandas de frequência delta, teta, alfa, beta e gama.

4. METODOLOGIA

Durante a etapa de pré-processamento, que consiste na limpeza, modelagem e estruturação dos dados de forma a garantir o melhor desempenho do modelo, as informações contidas no banco de dados não apresentaram irregularidades. Além disso, utilizando boxplots para visualizar a distribuição e possíveis discrepâncias na análise exploratória dos dados (AED), constatou-se a sua uniformidade, ausência de outliers e uma visível separação entre os valores das bandas de frequência do grupo dos não respondedores (1) e dos respondedores (2), como exemplificado na Figura 4 pelo eletrodo na posição F7 e seguindo o padrão nas demais regiões.

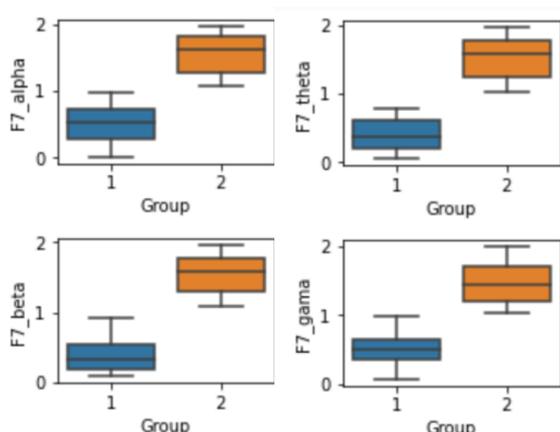


Figura 4. Distribuição do eletrodo F7 por boxplot

A fim de validar os achados do boxplot, realizou-se uma análise com o método não supervisionado *K-means*. Com isso, o modelo comprovou a distinção entre o grupo de respondedores e não respondedores para o banco de dados como um todo. Em adição, pelo procedimento da Análise de Componentes Principais (PCA) foi possível a visualização do espaçamento entre os dados, estando descrito na Figura 5. Esse artifício permite a realização da técnica de *feature extraction* de forma que os dados são organizados pela variância entre eles, sendo criada uma nova configuração da dimensão em que eles se apresentam Jamal et al. (2018).

Dessa forma, tais fatores corroboraram para a escolha do método da floresta aleatória para o desenvolvimento do algoritmo de predição dos melhores pontos de neuroestimulação pelo estudo das bandas de frequência do EEG.

5. RESULTADOS

Os hiperparâmetros envolvidos na construção do algoritmo da floresta aleatória influenciam na importância das variá-

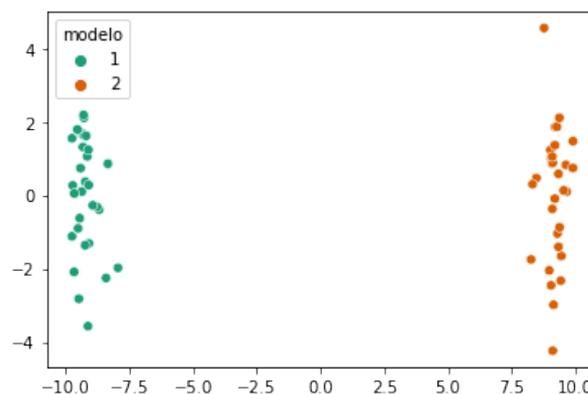


Figura 5. Distribuição do dataset por PCA

veis e em sua performance de predição Probst et al. (2019). Para tal, dentre os métodos, optou-se pela *Random Search* cuja aplicação se baseia em uma busca aleatória pelos pontos de amostragem selecionados avaliando o desempenho deles e comparando o comportamento obtido com base no histórico de tentativas antes de seguir para a próxima iteração Andradóttir (2006).

Sendo utilizado em conjunto a técnica de validação cruzada *K-fold*, nesse caso o K foi especificado como 3. Assim, o dataset foi dividido em K subamostras igualmente participadas e, posteriormente, sendo iterado o modelo k vezes utilizando K-1 subamostras como o banco de treinamento e a subamostra restante como dados de validação para o teste do modelo. Tal procedimento se repetiu até que cada uma das subamostras tivesse sido utilizada como dado de validação para as demais. Diante desse contexto, obteve-se os valores para os hiperparâmetros otimizados na Tabela 1.

Tabela 1. Configurações dos hiperparâmetros.

Hiperparâmetros	Resultado obtido
Quantidade de árvores	223
Critério de parada	Gini
Número máximo de árvores	50
Nível máximo da árvore	sqrt (square root)
Quantidade mínima de amostras para dividir o nó	2
Quantidade mínima de amostras em um nó folha	4
Presença de bootstrap	False

Com isso, a identificação dos melhores biomarcadores foi realizada por meio da ferramenta de *feature importance* da floresta aleatória, sendo essa métrica caracterizada como a importância relativa de determinada variável para a predição Probst et al. (2019). Sendo encontrados os pontos *OZ*, *CP1_gamma*, *F8_gamma*, *Fc1_gamma* e *P7_theta*, com a ordem do valor de divisão entre os grupos alvos expressos nas Figuras 6, 7 8, 9 e 10.

6. CONCLUSÃO

Atualmente a literatura já demonstra os benefícios terapêuticos que as técnicas não invasivas como ETCC e IC podem apresentar quando utilizadas em pacientes com DA Andrade et al. (2022); Khedr et al. (2014). A maioria dos ensaios envolvendo ETCC na DA normalmente

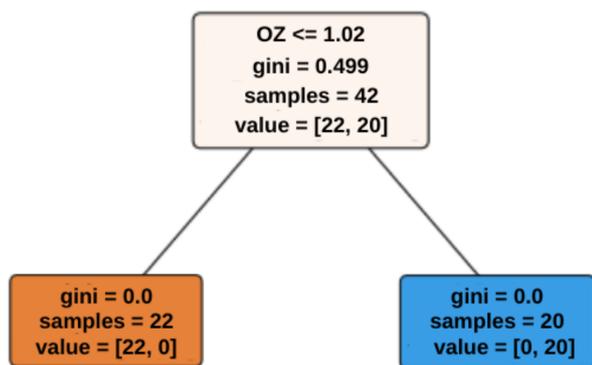


Figura 6. Distribuição do ponto *OZ*

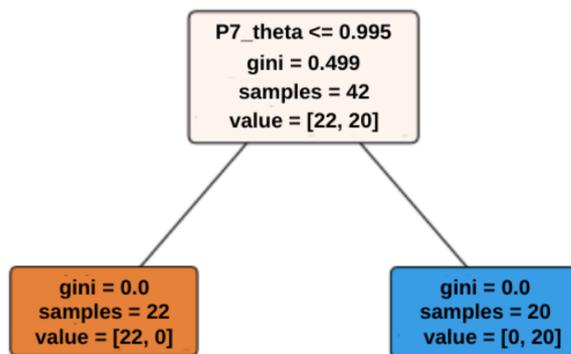


Figura 10. Distribuição do ponto *P7_theta*

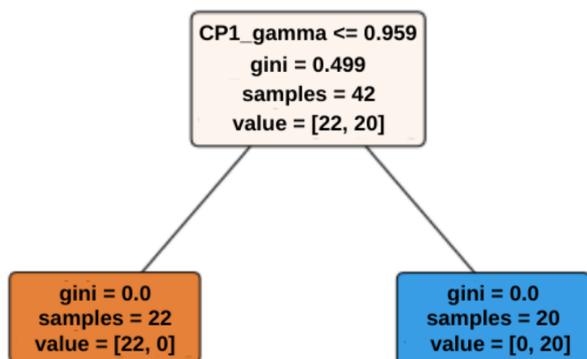


Figura 7. Distribuição do ponto *CP1_gamma*

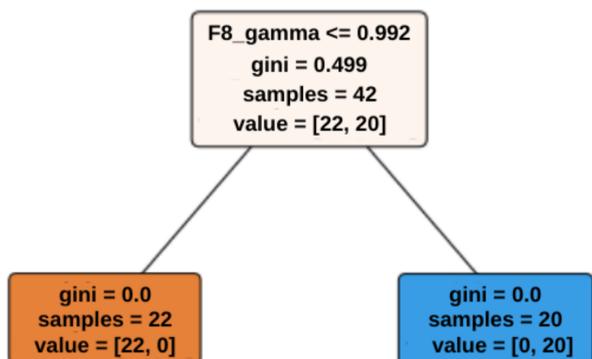


Figura 8. Distribuição do ponto *F8_gamma*

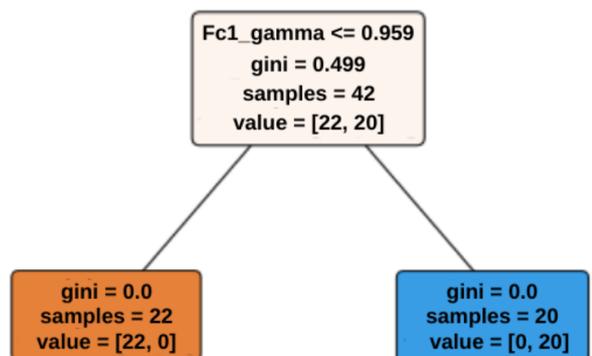


Figura 9. Distribuição do ponto *Fc1_gamma*

avalia os efeitos da neuroestimulação em apenas uma área cortical. Esse fato é alarmante, pois a demência difere de

outras doenças neurológicas, nas quais apenas uma região do cérebro deve ser modulada, como acidente vascular cerebral ou depressão Iannone et al. (2016); Ilić et al. (2016).

Diante disso, o presente estudo possibilitou a ampliação do leque de conhecimento sobre tal área de investigação, pois propôs a estimulação de diferentes regiões cerebrais especificamente relacionadas à neuropatologia da DA.

A corrente ativa foi aplicada em seis áreas corticais que representam centros primários envolvidos na manifestação dos sintomas clínicos da doença. Elas incluíram as porções esquerda e direita do córtex pré-frontal dorsolateral (F3 e F4), relacionados à memória de curto e longo prazo, capacidade de julgamento e funções executivas. A área de Broca (F5) e a área de Wernicke (CP5), que estão localizadas no lobo temporal, e são responsáveis pela linguagem, além do córtex de associação somatossensorial esquerdo e direito (P3 e P4), no lobo parietal, responsável pela orientação topográfica, espacial e práxis. A escolha dessas áreas seguiu protocolos de neuroestimulação já estabelecidos e previamente testados em pacientes com DA Bentwich et al. (2011); Nguyen et al. (2017); Rabey et al. (2013) e que concordam com estudos já publicados na literatura.

Observou-se então, uma notável afinidade entre os resultados encontrados pelo modelo de árvore de decisão, e o protocolo utilizado em nosso ensaio clínico. Estimulamos a região de compreensão da linguagem (CP5), e consequentemente, devido a dispersão de corrente gerada pela aplicação da ETCC, atingimos a região de CP1, alvo sugerido pelos achados do modelo. De maneira semelhante, a região frontal também foi um dos focos de aplicação da ETCC, em especial o córtex pré frontal dorsolateral, reforçando assim os dados encontrados pelo modelo de decisão. Enquanto que, a região de *OZ* proposta, pode ter sido ativada por meio de vias subcorticais, uma vez que o eletrodo ocupa uma grande extensão, e tendo sido colocado em CP5, pode ter gerado uma comunicação na região parietooccipital.

Além de concordar com os achados descritos pela avaliação da IA, o presente estudo concorda também com outros artigos previamente descritos na literatura Andrade et al. (2016); de Almeida Rodrigues et al. (2020); Zhao et al. (2017) que demonstram os benefícios alcançados pela aplicação da neuromodulação em áreas específicas, proporcionando assim a melhoria das funções cognitivas dos pacientes.

Portanto, diante desse contexto, o uso da ETCC associada com IA pode fornecer uma boa ferramenta que indique previamente os pacientes que serão mais responsivos ao tratamento, possibilitando assim uma redução na heterogeneidade encontrada em diversos estudos, para que assim possamos aumentar o tamanho de efeito da estimulação no estudo Boyd et al. (2017), a fim de que intervenções personalizadas sejam usadas para prever melhores respostas à terapia

AGRADECIMENTOS

Os autores agradecem à Universidade Federal da Paraíba (UFPB) pelo apoio fornecido para o desenvolvimento desta pesquisa.

REFERÊNCIAS

- Ahmed, M., Seraj, R., and Islam, S.M.S. (2020). The k-means algorithm: A comprehensive survey and performance evaluation. *Electronics*, 9(8), 1295.
- Al Zoubi, O., Ki Wong, C., Kuplicki, R.T., Yeh, H.w., Mayeli, A., Refai, H., Paulus, M., and Bodurka, J. (2018). Predicting age from brain eeg signals—a machine learning approach. *Frontiers in aging neuroscience*, 10, 184.
- Albizu, A., Fang, R., Indahlastari, A., O’Shea, A., Stolte, S.E., See, K.B., Boutzoukas, E.M., Kraft, J.N., Nissim, N.R., and Woods, A.J. (2020). Machine learning and individual variability in electric field characteristics predict tdcS treatment response. *Brain stimulation*, 13(6), 1753–1764.
- Amorim, E., Van der Stoel, M., Nagaraj, S.B., Ghassemi, M.M., Jing, J., O’Reilly, U.M., Scirica, B.M., Lee, J.W., Cash, S.S., and Westover, M.B. (2019). Quantitative eeg reactivity and machine learning for prognostication in hypoxic-ischemic brain injury. *Clinical Neurophysiology*, 130(10), 1908–1916.
- Andrade, S.M., da Silva Machado, D.G., da Silva-Sauerc, L., Regis, C.T., Mendes, C.K.T.T., de Araújo, J.S.S., de Araújo, K.D.T., Costa, L.P., Queiroz, M.E.B.S., Leitão, M.M., et al. (2022). Effects of multisite anodal transcranial direct current stimulation combined with cognitive stimulation in patients with alzheimer’s disease and its neurophysiological correlates: A double-blind randomized clinical trial. *Neurophysiologie Clinique*.
- Andrade, S.M., de Oliveira, E.A., Alves, N.T., Dos Santos, A.C.G., de Mendonça, C.T.P.L., Sampaio, D.D.A., da Silva, E.E.Q.C., da Fonsêca, É.K.G., de Almeida Rodrigues, E.T., de Lima, G.N.S., et al. (2018). Neurostimulation combined with cognitive intervention in alzheimer’s disease (neuroad): Study protocol of double-blind, randomized, factorial clinical trial. *Frontiers in Aging Neuroscience*, 334.
- Andrade, S.M., Mendonça, C.T.P.L.d., Pereira, T.C.L., Fernandez-Calvo, B., Araújo, R.C.N., and Alves, N.T. (2016). Adjuvant transcranial direct current stimulation for treating alzheimer’s disease: A case study. *Dementia & neuropsychologia*, 10, 156–159.
- Andradóttir, S. (2006). An overview of simulation optimization via random search. *Handbooks in operations research and management science*, 13, 617–631.
- Bentwich, J., Dobronevsky, E., Aichenbaum, S., Shorer, R., Peretz, R., Khaigreht, M., Marton, R.G., and Rabey, J.M. (2011). Beneficial effect of repetitive transcranial magnetic stimulation combined with cognitive training for the treatment of alzheimer’s disease: a proof of concept study. *Journal of neural transmission*, 118(3), 463–471.
- Boyd, L.A., Hayward, K.S., Ward, N.S., Stinear, C.M., Rosso, C., Fisher, R.J., Carter, A.R., Leff, A.P., Copland, D.A., Carey, L.M., et al. (2017). Biomarkers of stroke recovery: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *International Journal of Stroke*, 12(5), 480–493.
- Breiman, L. (2001). Random forests. *Machine learning*, 45(1), 5–32.
- Breiman, L., Friedman, J.H., Olshen, R.A., and Stone, C.J. (2017). *Classification and regression trees*. Routledge.
- Brunoni, A.R., Nitsche, M.A., Bolognini, N., Bikson, M., Wagner, T., Merabet, L., Edwards, D.J., Valero-Cabre, A., Rotenberg, A., Pascual-Leone, A., et al. (2012). Clinical research with transcranial direct current stimulation (tdcs): challenges and future directions. *Brain stimulation*, 5(3), 175–195.
- Buschert, V., Bokde, A.L., and Hampel, H. (2010). Cognitive intervention in alzheimer disease. *Nature Reviews Neurology*, 6(9), 508–517.
- Cutler, A., Cutler, D.R., and Stevens, J.R. (2012). Random forests. In *Ensemble machine learning*, 157–175. Springer.
- de Almeida Rodrigues, E.T., da Silva Machado, D.G., de Mendonça, C.T.P.L., da Rocha Souto, C., Ries, A., Torro, N., de Souza, I.C., Fernández-Calvo, B., and Andrade, S.M. (2020). Multisite transcranial direct current stimulation in two patients with alzheimer’s disease: a 10-month follow-up study. *Neurophysiologie clinique= Clinical neurophysiology*, 50(5), 393–395.
- De Berker, A.O., Bikson, M., and Bestmann, S. (2013). Predicting the behavioral impact of transcranial direct current stimulation: issues and limitations. *Frontiers in human neuroscience*, 7, 613.
- Fregni, F., El-Hagrassy, M.M., Pacheco-Barríos, K., Carvalho, S., Leite, J., Simis, M., Brunelin, J., Nakamura-Palacios, E.M., Marangolo, P., Venkatasubramanian, G., et al. (2021). Evidence-based guidelines and secondary meta-analysis for the use of transcranial direct current stimulation in neurological and psychiatric disorders. *International Journal of Neuropsychopharmacology*, 24(4), 256–313.
- Guerra, A., López-Alonso, V., Cheeran, B., and Suppa, A. (2020). Variability in non-invasive brain stimulation studies: reasons and results. *Neuroscience letters*, 719, 133330.
- Hasanzadeh, F., Mohebbi, M., and Rostami, R. (2019). Prediction of rtms treatment response in major depressive disorder using machine learning techniques and nonlinear features of eeg signal. *Journal of affective disorders*, 256, 132–142.
- Huang, L.K., Chao, S.P., and Hu, C.J. (2020). Clinical trials of new drugs for alzheimer disease. *Journal of biomedical science*, 27(1), 1–13.
- Iannone, A., Cruz, A.P.d.M., Brasil-Neto, J.P., and Boechat-Barros, R. (2016). Transcranial magnetic stimulation and transcranial direct current stimulation appear to be safe neuromodulatory techniques useful in the treatment of anxiety disorders and other neuropsychy-

- chiatric disorders. *Arquivos de Neuro-Psiquiatria*, 74, 829–835.
- Ilić, N.V., Dubljanin-Raspopović, E., Nedeljković, U., Tomanović-Vujadinović, S., Milanović, S.D., Petronić-Marković, I., and Ilić, T.V. (2016). Effects of anodal tdcS and occupational therapy on fine motor skill deficits in patients with chronic stroke. *Restorative neurology and neuroscience*, 34(6), 935–945.
- Jain, A.K., Murty, M.N., and Flynn, P.J. (1999). Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3), 264–323.
- Jamal, A., Handayani, A., Septiandri, A.A., Ripmiatin, E., and Effendi, Y. (2018). Dimensionality reduction using pca and k-means clustering for breast cancer prediction. *Lontar Komputer: Jurnal Ilmiah Teknologi Informatika*, 192–201.
- Kambeitz, J., Goerigk, S., Gattaz, W., Falkai, P., Bensenor, I.M., Lotufo, P.A., Buehner, M., Koutsouleris, N., Padberg, F., and Brunoni, A.R. (2020). Clinical patterns differentially predict response to transcranial direct current stimulation (tdcs) and escitalopram in major depression: a machine learning analysis of the elect-tdcs study. *Journal of affective disorders*, 265, 460–467.
- Khedr, E.M., Gamal, N.F.E., El-Fetoh, N.A., Khalifa, H., Ahmed, E.M., Ali, A.M., Noaman, M., El-Baki, A.A., and Karim, A.A. (2014). A double-blind randomized clinical trial on the efficacy of cortical direct current stimulation for the treatment of alzheimer's disease. *Frontiers in aging neuroscience*, 6, 275.
- Khodayari-Rostamabad, A., Reilly, J.P., Hasey, G.M., de Bruin, H., and MacCrimmon, D.J. (2013). A machine learning approach using eeg data to predict response to ssri treatment for major depressive disorder. *Clinical Neurophysiology*, 124(10), 1975–1985.
- Knopman, D.S., Amieva, H., Petersen, R.C., Chételat, G., Holtzman, D.M., Hyman, B.T., Nixon, R.A., and Jones, D.T. (2021). Alzheimer disease. *Nature reviews Disease primers*, 7(1), 1–21.
- Lefaucheur, J.P., Antal, A., Ayache, S.S., Benninger, D.H., Brunelin, J., Cogiamanian, F., Cotelli, M., De Ridder, D., Ferrucci, R., Langguth, B., et al. (2017). Evidence-based guidelines on the therapeutic use of transcranial direct current stimulation (tdcs). *Clinical Neurophysiology*, 128(1), 56–92.
- Marceglia, S., Mrakic-Sposta, S., Rosa, M., Ferrucci, R., Mameli, F., Vergari, M., Arlotti, M., Ruggiero, F., Scarpini, E., Galimberti, D., et al. (2016). Transcranial direct current stimulation modulates cortical neuronal activity in alzheimer's disease. *Frontiers in neuroscience*, 10, 134.
- Myles, A.J., Feudale, R.N., Liu, Y., Woody, N.A., and Brown, S.D. (2004). An introduction to decision tree modeling. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 18(6), 275–285.
- Nascimento, W.K., Queiroz, M.E.B.S., Costa, L.P., de Oliveira, E.A., Fernández-Calvo, B., and Andrade, S.M. (2020). Estimulação transcraniana por corrente contínua na doença de alzheimer: uma revisão sistemática. *Research, Society and Development*, 9(12), e43291210459–e43291210459.
- Nguyen, J.P., Suarez, A., Kemoun, G., Meignier, M., Le Saout, E., Damier, P., Nizard, J., and Lefaucheur, J.P. (2017). Repetitive transcranial magnetic stimulation combined with cognitive training for the treatment of alzheimer's disease. *Neurophysiologie Clinique/Clinical Neurophysiology*, 47(1), 47–53.
- Probst, P., Wright, M.N., and Boulesteix, A.L. (2019). Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery*, 9(3), e1301.
- Products, B. (2018). Brain vision actichamp. <https://brainvision.com/products/actichamp-plus/>. Acessado em: 2022-04-25.
- Rabey, J.M., Dobronevsky, E., Aichenbaum, S., Gonen, O., Marton, R.G., and Khaigreht, M. (2013). Repetitive transcranial magnetic stimulation combined with cognitive training is a safe and effective modality for the treatment of alzheimer's disease: a randomized, double-blind study. *Journal of Neural Transmission*, 120(5), 813–819.
- Sá, C.d.C., Silva, D.F.d., Bigongiari, A., and Machado-Lima, A. (2019). Eficácia da reabilitação cognitiva na melhoria e manutenção das atividades de vida diária em pacientes com doença de alzheimer: uma revisão sistemática da literatura. *Jornal Brasileiro de Psiquiatria*, 68, 153–160.
- Zhao, H., Qiao, L., Fan, D., Zhang, S., Turel, O., Li, Y., Li, J., Xue, G., Chen, A., and He, Q. (2017). Modulation of brain activity with noninvasive transcranial direct current stimulation (tdcs): clinical applications and safety concerns. *Frontiers in psychology*, 8, 685.