

# APLICAÇÃO DE REDES NEURAIS NA CLASSIFICAÇÃO DE SINAIS DE EEG PARA CONDIÇÕES NEUROPSIQUIÁTRICAS

Mateus Balda Mota<sup>★</sup>  
Alessandro Bof de Oliveira<sup>☆</sup>

## RESUMO

Exames de eletroencefalograma (EEG) são amplamente utilizados para medir a atividade cerebral, sendo uma ferramenta promissora, embora complementar, no diagnóstico de transtornos neuropsiquiátricos, como ansiedade, esquizofrenia e epilepsia, entre outros. Métodos tradicionais, com base no *Manual Diagnóstico e Estatístico de Transtornos Mentais* (DSM-5), apresentam limitações devido à subjetividade, destacando a necessidade de abordagens quantitativas. Este estudo teve como objetivo desenvolver e validar modelos de aprendizado profundo para classificar condições neuropsiquiátricas utilizando sinais de EEG. Os dados, previamente processados com *Power Spectrum Density* (PSD) e *Functional Connectivity* (FC), contêm 1140 características e foram divididos combinando todas as bandas com PSD e FC, com *oversampling* por meio da técnica *Synthetic Minority Oversampling Technique* (SMOTE). Foram avaliadas as arquiteturas ANN, TabNet e FT-Transformer, com métricas de acurácia, sensibilidade e *Area Under the Curve* (AUC). Os modelos apresentaram acurácias médias de 93,62% e 91,93%, com destaque para transtorno de humor (97,20%) usando TabNet beta PSD + FC, e para pânico e estresse pós-traumático (97,30%) com PSD completa e gama PSD + FC usando ANN e FTT. Constatou-se que a variabilidade intersujeito, a limitação no número de amostras e a heterogeneidade entre PSD e FC impactam o desempenho. A classificação binária saudável vs. paciente apresenta melhor desempenho. Sugere-se a aplicação de técnicas de geração de dados sintéticos mais realistas para melhorar a representatividade dos dados, a explicabilidade dos modelos e o uso de *ensemble* dos modelos binários.

**Palavras-chaves:** Eletroencefalografia (EEG). Distúrbios neuropsiquiátricos. Classificação de sinais de EEG. Redes neurais artificiais. Aprendizado profundo.

## ABSTRACT

Electroencephalogram (EEG) exams are widely used to measure brain activity and represent a promising, albeit complementary, tool in the diagnosis of neuropsychiatric disorders such as anxiety, schizophrenia, epilepsy, among others. Traditional methods based on the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-5) have limitations due to their subjectivity, highlighting the need for quantitative approaches. This study aimed to develop and validate deep learning models to classify neuropsychiatric

<sup>★</sup>Aluno do Curso de Engenharia de Software da Universidade Federal do Pampa, Alegrete, Rio Grande do Sul, Brasil  
E-mail: [mateusmota.aluno@unipampa.edu.br](mailto:mateusmota.aluno@unipampa.edu.br)

<sup>☆</sup>Orientador, Professor do Curso de Ciência da Computação da Universidade Federal do Pampa, Alegrete, Rio Grande do Sul, Brasil  
E-mail: [alessandrooliveira@unipampa.edu.br](mailto:alessandrooliveira@unipampa.edu.br)

conditions using EEG signals. The data, preprocessed with *Power Spectrum Density* (PSD) and *Functional Connectivity* (FC), contained 1,140 features and were divided by combining all frequency bands with PSD and FC, using oversampling through the *Synthetic Minority Oversampling Technique* (SMOTE). The ANN, TabNet, and FT-Transformer architectures were evaluated using accuracy, sensitivity, and *Area Under the Curve* (AUC) metrics. The models achieved average accuracies of 93.62% and 91.93%, with notable results for mood disorder (97.20%) using TabNet beta PSD + FC, and for panic and post-traumatic stress disorder (97.30%) using full PSD and gamma PSD + FC with ANN and FTT. It was found that inter-subject variability, the limited number of samples, and the heterogeneity between PSD and FC affect model performance. The binary classification of healthy vs. patient showed better performance. The application of more realistic synthetic data generation techniques is suggested to improve data representativeness, model explainability, and the use of *ensemble* methods based on binary models.

**Keywords:** Electroencephalography (EEG). Neuropsychiatric disorders. EEG signal classification. Artificial neural networks. Deep learning.

## 1. INTRODUÇÃO

### 1.1. CONTEXTO

O eletroencefalograma (EEG) é uma técnica que mede a atividade elétrica do cérebro por meio de eletrodos posicionados no couro cabeludo (SIULY; LI; ZHANG, 2016). Ele capta variações nos sinais elétricos neuronais e permite o monitoramento da atividade cerebral, incluindo diferentes bandas de frequência, como *delta*, *theta*, *alpha*, *beta* e *gamma*. Comumente utilizado na medicina, o EEG é uma ferramenta que pode ser empregada no diagnóstico de condições neurológicas e psiquiátricas (SIULY; LI; ZHANG, 2016; MÜLLER-PUTZ, 2020). (BENICZKY; SCHOMER, 2020) descrevem a aplicação bem-sucedida dessa técnica no diagnóstico de epilepsia. No estudo de (PARK et al., 2021), propõe-se a identificação de transtornos psiquiátricos a partir de sinais provenientes de diferentes grupos de diagnósticos.

Os métodos tradicionais, como os critérios do DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders, 5th Edition*), geralmente dependem da avaliação de sintomas comportamentais e relatórios subjetivos, o que pode levar a diagnósticos imprecisos ou inconsistentes (ASSOCIATION, 2013). Essa subjetividade aumenta a demanda por abordagens mais objetivas e quantitativas, como o QEEG (*Quantitative Electroencephalography*), que possam complementar ou substituir os métodos convencionais, tornando o diagnóstico mais confiável e eficiente.

As condições neuropsiquiátricas incluem uma ampla gama de distúrbios que impactam tanto a saúde mental quanto o funcionamento neurológico. Esta pesquisa, o foco recai sobre os transtornos psiquiátricos, como depressão, ansiedade e esquizofrenia, entre outros. Dados da *Sapein Labs* e da OMS indicam que esses transtornos representam uma carga significativa para a sociedade, com alta prevalência e impacto social (World Health Organization, 2022) (Sapien Labs, 2024).

As redes neurais, inspiradas no funcionamento do cérebro humano (HAYKIN, 2011; RUSSELL; NORVIG, 2022), têm se mostrado ferramentas poderosas na análise

de sinais de EEG (AHMED et al., 2024; AY et al., 2019; JÚNIOR, 2023), especialmente em contextos neuropsiquiátricos. Elas são capazes de identificar padrões complexos e sutis nos dados, superando os métodos convencionais em termos de precisão, rapidez e custo, o que tem melhorado o diagnóstico de condições como transtornos mentais (PARSA et al., 2023).

## 1.2. JUSTIFICATIVA

As condições neuropsiquiátricas afetam diretamente a vida em sociedade. Em 2019, 1 em cada 8 pessoas no mundo vivia com transtorno mental, totalizando 970 milhões (World Health Organization, 2022). Esse número aumentou com a covid-19, com crescimentos de 26% nos casos de ansiedade e 28% de depressão em um ano (World Health Organization, 2022). Com base na classificação da *ICD-11* em 2019 aponta 301 milhões de pessoas com transtornos de ansiedade, 280 milhões com depressão, 40 milhões com transtorno bipolar, 24 milhões com esquizofrenia e 14 milhões com distúrbios alimentares (World Health Organization, 2022). O relatório de 2023 da *Sapient Labs* mostrou que o Quociente de Saúde Mental (MHQ) caiu de 90 pontos em 2019 para 58 pontos em 2021, subindo para 62 em 2022 e 2023 (Sapient Labs, 2024).

Nesse contexto, métodos de diagnóstico, como o EEG, ganham importância. Ele oferece dados sobre a atividade cerebral e pode ser útil na identificação de transtornos mentais, especialmente quando combinado com aprendizado de máquina (SIULY; LI; ZHANG, 2016) (PARK et al., 2021). Pesquisas indicam que o uso de redes neurais, como *multilayer perceptron* e *recurrent neural network*, melhora a acurácia diagnóstica em distúrbios como depressão, esquizofrenia e ansiedade (AY et al., 2019) (AHMED et al., 2024) (JÚNIOR, 2023).

A revisão de (PARSA et al., 2023) analisou estudos que aplicaram redes neurais na classificação de sinais de EEG para diagnosticar distúrbios psiquiátricos, doenças degenerativas e do neurodesenvolvimento, relatando acurácia média de 91,83%. Os resultados indicam que essas redes podem complementar métodos tradicionais, como o DSM-5, desde que treinadas com grandes volumes de dados para garantir boa generalização. Os autores destacam a importância de considerar a alta variabilidade intersujeitos nos sinais de EEG. Além da precisão, estudos ressaltam vantagens práticas das redes neurais, como a rapidez diagnóstica demonstrada por (PEH et al., 2021), cujo modelo analisou 30 minutos de EEG em apenas 4 segundos, e a redução de custos evidenciada por (ULLAH et al., 2018), ao automatizar a detecção de epilepsia e aliviar a carga sobre especialistas.

### 1.2.1. LINHA DE PESQUISA: ENGENHARIA DE SOFTWARE APLICADA

A escolha da linha de pesquisa em engenharia de software aplicada para esta pesquisa justifica-se pela natureza multidisciplinar do problema e pela necessidade de um desenvolvimento estruturado e eficiente de uma solução computacional complexa. O objetivo central é projetar, desenvolver e validar modelos voltada para a classificação de condições neuropsiquiátricas utilizando sinais de EEG. Essa abordagem exige a criação de *softwares*, seguindo os atributos de um bom *software* (SOMMERVILLE, 2011), como manutenibilidade, controle de versões com ferramentas como `GitHub`<sup>1</sup>

---

<sup>1</sup><<https://github.com/>>

e Kaggle<sup>2</sup>, e o uso de metodologia ágil no desenvolvimento, priorizando tarefas por meio de um quadro *Kanban* individual do desenvolvedor. A *pipeline* do desenvolvimento foca em projetar, arquitetar, desenvolver e validar o desempenho dos modelos. Também foi desenvolvido um *software* utilizando o *Streamlit* para visualizações de tabelas e gráficos interativos.

### 1.3. OBJETIVOS

O objetivo geral desta pesquisa é desenvolver e validar modelos de aprendizado profundo para a classificação de condições neuropsiquiátricas, utilizando sinais de EEG, com o intuito de identificar padrões relevantes que possam auxiliar no diagnóstico clínico. Para isso, foram definidos os seguintes objetivos específicos:

- Dividir o conjunto de dados em subconjuntos, variando as combinações de bandas de frequência e tipos de parâmetros, com o objetivo de avaliar como essas variações influenciam o desempenho dos modelos.
- Balancear as classes minoritárias em relação à classe majoritária utilizando técnicas de *oversampling*, visando reduzir o desbalanceamento e melhorar o desempenho dos modelos na classificação.
- Implementar e comparar o desempenho de diferentes modelos, para avaliar a eficácia de cada abordagem no contexto do problema.
- Avaliar o desempenho dos modelos por meio de métricas como *accuracy*, *precision*, *recall* e AUC (Área sob a Curva ROC – *Receiver Operating Characteristic*), e comparar os resultados com as abordagens apresentadas por (PARK et al., 2021), que utilizaram *SVM (Support Vector Machines)*, *random forest* e *logistic regression*.

### 1.4. ORGANIZAÇÃO

Este trabalho está organizado em quatro seções. A Seção 2 aborda condições neuropsiquiátricas, eletroencefalografia, aprendizado de máquina e profundo, arquiteturas TabNet e FT-Transformer, bibliotecas e trabalhos relacionados. A Seção 3 descreve os métodos e materiais. Na Seção 4, são apresentados os resultados e discussões. Por fim, a Seção 5 traz as considerações finais e propostas de trabalhos futuros.

## 2. FUNDAMENTAÇÃO TEÓRICA

Esta seção explora os principais conceitos que fundamentam a pesquisa, abordando as seguintes subseções: condições neuropsiquiátricas, eletroencefalografia, aprendizado de máquina, aprendizado profundo, TabNet e FT-transformer. Na subseção de bibliotecas e ferramentas, são apresentados conceitos tecnológicos. A subseção de trabalhos relacionados discute pesquisas pertinentes ao estudo.

---

<sup>2</sup><<https://www.kaggle.com/>>

## 2.1. CONDIÇÕES NEUROPSIQUIÁTRICAS

O diagnóstico em psiquiatria segue quatro etapas: 1) pré-sindrômica, com base na confirmação ou exclusão de um sintoma; 2) sindrômica, identificando um conjunto de sintomas que formam uma síndrome; 3) fisiopatológica, reconhecendo uma fisiopatologia subjacente; e 4) etiológica, que investiga a origem das doenças, pouco manejável (NARDI; SILVA; QUEVEDO, 2021).

Nos últimos anos, o diagnóstico psiquiátrico depende de entrevistas clínicas, que identificam sinais e sintomas de doenças específicas, sendo um processo indutivo. Com o avanço na identificação de marcadores biológicos (neuroimagem, inflamatórios, epigenéticos, neurotransmissores, entre outros), a etiologia se tornou o centro da classificação, embora a ausência de mecanismos claros comprometa a validade diagnóstica (NARDI; SILVA; QUEVEDO, 2021).

Transtornos como a depressão apresentam grande variação fenotípica, dificultando previsões precisas. Classificações como CID-10 (*International Statistical Classification of Diseases and Related Health Problems 10th Revision*), CID-11 e DSM-5 (*Diagnostic and Statistical Manual of Mental Disorders, 5th edition*) são essenciais para orientar o diagnóstico clínico, estabelecendo categorias e fornecendo informações sobre prognóstico e sintomas. O DSM-5 adota uma abordagem dimensional, classificando sintomas por gravidade. A CID-10 é descritiva, utilizando etiologia apenas em transtornos específicos, como os orgânicos e os relacionados ao uso de substâncias (ASSOCIATION, 2013) (NARDI; SILVA; QUEVEDO, 2021).

A CID-10 agrupa os transtornos mentais em dez grandes categorias, identificadas por códigos de F0 a F9, abrangendo desde condições orgânicas e relacionadas ao uso de substâncias até transtornos de humor, personalidade, desenvolvimento e infância (ORGANIZATION, 2004).

Transtornos mentais, como depressão e ansiedade, resultam de disfunções cerebrais que podem ser identificadas por exames de neuroimagem e neurofisiologia. Sistemas como o de valência negativa (ligado ao medo e ansiedade) e o de valência positiva (associado à impulsividade) ajudam a compreender diferentes transtornos. Já os sistemas regulatórios, como o ciclo sono-vigília, estão diretamente relacionados a condições como a depressão, afetando energia, apetite e sono (NARDI; SILVA; QUEVEDO, 2021).

## 2.2. ELETROENCEFALOGRAFIA

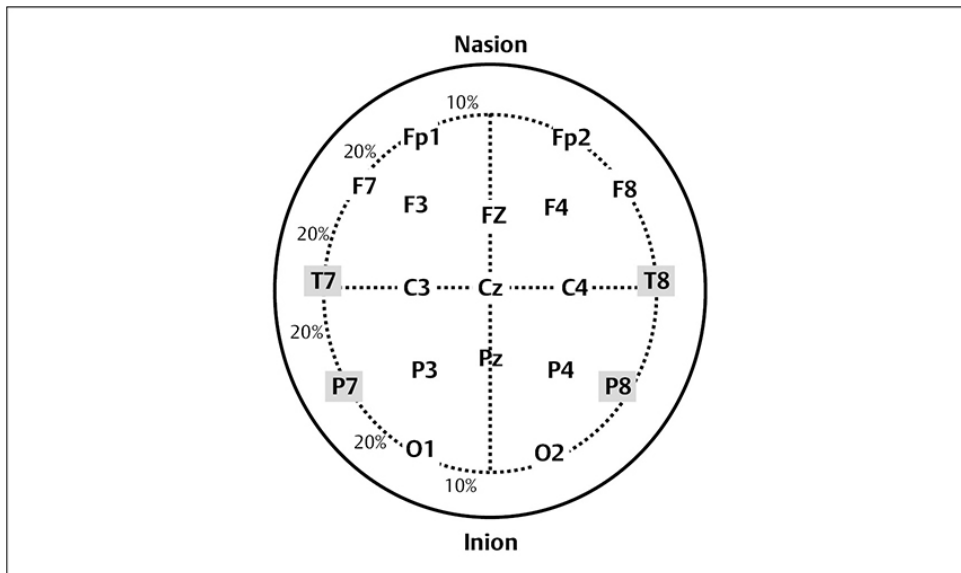
A integração sináptica no sistema nervoso ocorre por dois meios: elétrico e químico. As sinapses elétricas no sistema nervoso são responsáveis pela despolarização. Nessas sinapses, o fluxo elétrico ocorre através da membrana pré-sináptica e de canais que conectam as células pré e pós-sinápticas (MONTENEGRO et al., 2022). A atividade elétrica cerebral é gerada pelos potenciais de ação da membrana neural, com potenciais pós-sinápticos excitatórios e inibitórios (BENICZKY; SCHOMER, 2020). Para que ocorra o movimento de elétrons, é necessária uma diferença de potencial entre as extremidades do condutor. A soma desses potenciais resulta na atividade elétrica cerebral registrada no escalpo (MONTENEGRO et al., 2022).

A eletroencefalografia monitora sinais eletrofisiológicos registrados através do couro cabeludo por meio de eletrodos, captando a atividade elétrica cerebral a partir da

soma dos potenciais pós-sinápticos de neurônios localizados nas camadas do córtex cerebral (MÜLLER-PUTZ, 2020) (BENICZKY; SCHOMER, 2020). O sistema internacional 10-20 (PEDLEY; DALY, 1997) é baseado em pontos anatômicos específicos que servem de referência para determinar a localização de cada eletrodo no escalpo. O sistema se refere à distância e à distribuição entre os eletrodos, que é de 10% ou 20% da distância total entre dois pontos de referência. Essa técnica possibilita que o intervalo entre os eletrodos seja proporcional e equivalente, independentemente do tamanho da cabeça do paciente. A colocação de 21 eletrodos é feita com base em pontos de referência como o *nasion* e o *inion*, com o objetivo de cobrir todas as áreas do escalpo: fronto polar, frontal, parietal, temporal, central e occipital (MONTENEGRO et al., 2022).

A nomenclatura adotada para descrever a posição dos eletrodos segue um padrão. Cada eletrodo recebe uma letra correspondente à área do cérebro onde está posicionado: **Fp** para fronto polar, **F** para frontal, **T** para temporal, **C** para central, **P** para parietal e **O** para occipital. Os eletrodos localizados na linha média são denominados com a letra **z**. Os números ímpares indicam locais no hemisfério esquerdo, e os números pares indicam locais no hemisfério direito. Dessa forma, a nomenclatura é padronizada, conforme apresentado na figura 1, com os eletrodos numerados para distinguir os hemisférios e a localização em relação à linha média (MONTENEGRO et al., 2022).

Figura 1 - Sistema 10-20 utilizado para definir a nomenclatura e a posição dos eletrodos



Fonte: Figura extraída do Capítulo 3 - Montagens e Polaridade (MONTENEGRO et al., 2022)

Os mecanismos de ativação servem para modular a atividade elétrica cerebral, com o intuito de identificar respostas anormais. Os mais utilizados são: abertura e fechamento ocular, hiperventilação, fotoestimulação intermitente e estímulos sensitivos (sonoros, táteis e dolorosos) (MONTENEGRO et al., 2022). Aconselha-se realizar, além dos registros eletroencefalográficos em vigília, também durante o sono, para que se possam detectar anormalidades que podem estar ausentes nos traçados de vigília; para isso, é indicado que o paciente passe por privação de sono (MONTENEGRO et al., 2022).

Os sinais são captados no escalpo e transmitidos através das camadas corticais (BENICZKY; SCHOMER, 2020), apresentando oscilações rítmicas. As principais ondas cerebrais estão na faixa de frequência de 0.5 Hz a 40 Hz (KOUDELKOVÁ; STRMISKA, 2018). (BENICZKY; SCHOMER, 2020) indica uma faixa de frequência de 0.1 a 70 Hz para interpretações clínicas relevantes. Em (PARK et al., 2021), foram utilizadas 6 bandas de frequências, uma a mais do que o usual, incluindo a frequência *high beta*, intermediária entre *beta* e *gamma*, que varia de 25 a 30 Hz.

O intervalo de frequência das ondas cerebrais é subdividido em faixas menores, conhecidas como bandas de frequência (KOUDELKOVÁ; STRMISKA, 2018). Cada uma dessas bandas possui características específicas, incluindo frequências, amplitudes e significados, como ilustrado na figura 2

- delta  $\delta$  (0.5-4 Hz): ocorre durante a meditação, sono profundo ou coma, e pode estar presente em pessoas com dificuldades de aprendizagem ou problemas de consciência.
- theta  $\theta$  (4-8 Hz): relacionada ao sono, à intuição e a tarefas automáticas.
- alpha  $\alpha$  (8-13 Hz): envolve o pensamento consciente e a mente subconsciente, proporcionando uma sensação de relaxamento.
- beta  $\beta$  (13-30 Hz): ativa em estados de vigília, essa frequência é frequentemente associada ao raciocínio lógico-analítico e à resolução de problemas.
- gamma  $\gamma$  (>30 Hz): relacionada ao aprendizado, à memória e à assimilação de informações.

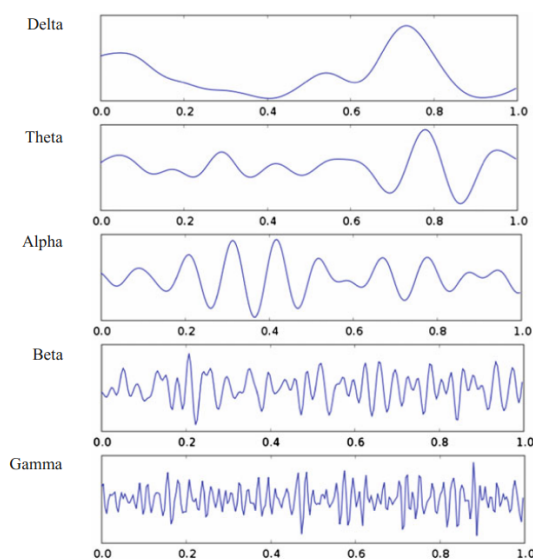


Figura 2 - Ritmos de EEG em diferentes frequências

Fonte: (SIJLY; LI; ZHANG, 2016).

### 2.3. APRENDIZADO DE MÁQUINA (*Machine learning*)

A *artificial intelligence* (AI) pode ser definida tanto por semelhança ao desempenho humano quanto por abordagens formais baseadas na racionalidade. Ela abrange subáreas como aprendizado, raciocínio e percepção, sendo frequentemente descrita como um agente racional, um sistema que toma decisões autônomas para atingir objetivos da melhor forma possível. Apesar de eficaz em contextos com metas bem definidas, esse modelo encontra limitações em cenários incertos, como na visão computacional de carros autônomos (RUSSELL; NORVIG, 2022).

Na aprendizagem de máquina, um agente aprende a melhorar seu desempenho a partir de observações sobre determinado objeto ou situação. Por meio do aprendizado de máquina, qualquer componente de um agente pode ser aprimorado. Por exemplo, um agente de carro autônomo aprende observando um motorista. Ao observar o

motorista frear, o agente desenvolve uma regra de condição-ação (componente 1). Além disso, ao analisar imagens de ônibus, ele aprende a reconhecê-los (componente 2), aprimorando diferentes componentes de sua funcionalidade (RUSSELL; NORVIG, 2022). Segundo (RUSSELL; NORVIG, 2022), quando a saída do modelo é composta por um conjunto finito de valores, como verdadeiro/falso, o problema de aprendizagem é classificado como uma tarefa de **classificação**.

O *feedback* de aprendizado de máquina consiste em três tipos que podem acompanhar as entradas: **aprendizado supervisionado**, no qual o agente observa dados em pares de entrada e saída, aprendendo a partir de rótulos. Utiliza uma função que realiza o mapeamento mútuo entre entrada e saída. Este é o tipo de aprendizado que será utilizado neste trabalho, uma vez que dispomos de dados rotulados para a tarefa de classificação; aprendizado não supervisionado, em que o agente busca aprender padrões nas entradas sem o fornecimento de rótulos explícitos. Uma tarefa comumente utilizada na aprendizagem não supervisionada é o agrupamento; aprendizado por reforço, no qual o agente aprende por meio de uma série de reforços, recebendo recompensas e punições (RUSSELL; NORVIG, 2022).

## APRENDIZADO SUPERVISIONADO

O aprendizado supervisionado envolve um conjunto de treino com  $N$  pares de observações de entrada e saída,  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ . Uma função desconhecida  $y = f(x)$  gera cada par. O objetivo é descobrir a função  $h$ , que é a hipótese aproximada da função verdadeira  $f$ . O conjunto de hipóteses pode incluir funções lineares, como  $h(x) = w_1x + w_0$ , funções senoidais, como  $h(x) = w_1x + \sin(w_0x)$ , e funções polinomiais, como  $h(x) = \sum_{i=0}^{12} w_i x^i$  (RUSSELL; NORVIG, 2022).

Uma das formas de analisar o conjunto de hipóteses é através do viés. O *underfitting* acontece quando não se consegue encontrar um padrão nos dados, enquanto o *overfitting* ocorre quando a função está sobreajustada aos dados, foca demais no conjunto de dados de treinamento, o que prejudica seu desempenho em dados novos (RUSSELL; NORVIG, 2022).

Os modelos para aprendizado supervisionado incluem árvores de decisão, regressão linear, regressão logística, modelos não paramétricos como a regressão de vizinhos mais próximos, máquinas de vetores de suporte, e modelos probabilísticos, como a aprendizagem *bayesiana* (RUSSELL; NORVIG, 2022).

### 2.4. APRENDIZADO PROFUNDO (*Deep Learning*)

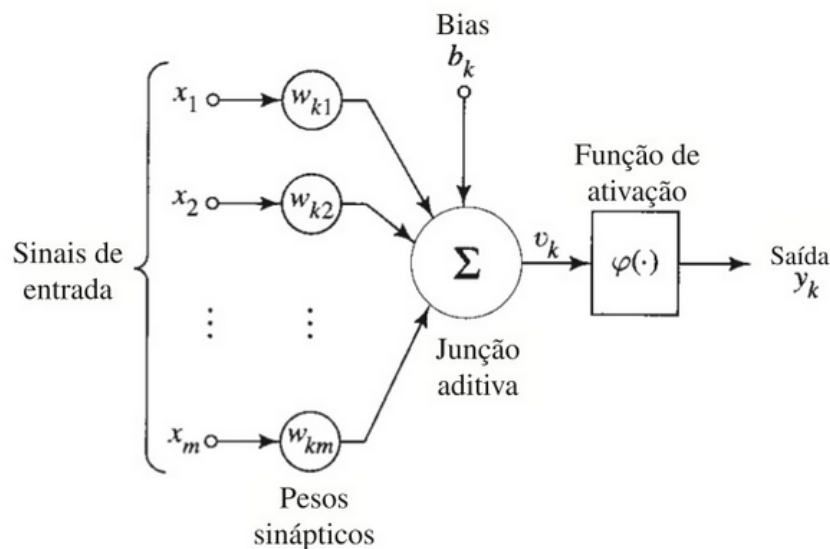
O sistema nervoso humano opera em três estágios: receptores, rede neural e atuadores. Receptores convertem estímulos externos em impulsos elétricos, que são processados pelo cérebro. Atuadores geram respostas a partir dessas informações. Há realimentação em todos os estágios (HAYKIN, 2011).

Redes neurais artificiais se inspiram nesse funcionamento. O cérebro processa informações de forma complexa, não linear e paralela, realizando tarefas como reconhecimento de padrões e controle motor. Por exemplo, o reconhecimento de um rosto familiar ocorre em 100 a 200 ms (HAYKIN, 2011). Segundo Haykin (2011, p. 28), “uma rede neural é um processador maciçamente paralelo distribuído constituído de unidades de processamento simples, que têm a propensão natural para armazenar conhecimento experimental e torná-lo disponível para o uso”.

Uma rede neural profunda é composta por múltiplas camadas que aplicam mapeamentos não lineares entre entrada e saída, com a maioria das camadas sujeitas a aprendizado. Essas redes possuem propriedades como adaptabilidade, resposta a evidências e consideração de informações contextuais. Com profundidades variando de 5 a 20 camadas, cada camada transforma a entrada, aprimorando a seletividade e invariância das representações (HAYKIN, 2011; LECUN; BENGIO; HINTON, 2015).

Na figura abaixo 3, é apresentado o modelo de um neurônio que serve como base para uma rede neural, composto por três elementos básicos: pesos sinápticos, junção aditiva e função de ativação.

Figura 3 - Modelo não-linear de um neurônio



Fonte:(HAYKIN, 2011).

**Feedforward (propagação direta):** A forma mais fundamental da rede neural artificial é a rede de propagação direta, onde os dados fluem em uma única direção: das camadas de entrada, passando por uma ou mais camadas ocultas, até a camada de saída, sem ciclos recorrentes (GOODFELLOW, 2016). Em cada neurônio de uma camada, é realizada uma operação linear sobre as entradas, seguida por uma função de ativação não linear:

$$y_k = \varphi \left( \sum_{i=1}^m w_{ki}x_i + b_k \right) \quad (1)$$

onde  $x_i$  representa a entrada  $i$ ,  $w_{ki}$  é o peso sináptico entre  $x_i$  e o neurônio  $k$ ,  $b_k$  é o viés adicionado ao neurônio e  $\varphi$  é a função de ativação (como softmax, sigmoide, entre outras) (HAYKIN, 2011).

**Backpropagation (retropropagação):** Responsável pela minimização da função de custo (*loss function*), como a *Binary Cross-Entropy* ou a *Categorical Cross-Entropy*, ajustando e atualizando iterativamente os pesos sinápticos. Fundamentado no cálculo do gradiente descendente estocástico, ele aplica a regra da cadeia para derivadas (LECUN; BENGIO; HINTON, 2015). Após a ativação, calcula-se o erro de predição em relação ao valor real, computando-se o gradiente do erro, que é transmitido de volta pela rede, camada por camada (da saída ao início) (RUMELHART; HINTON;

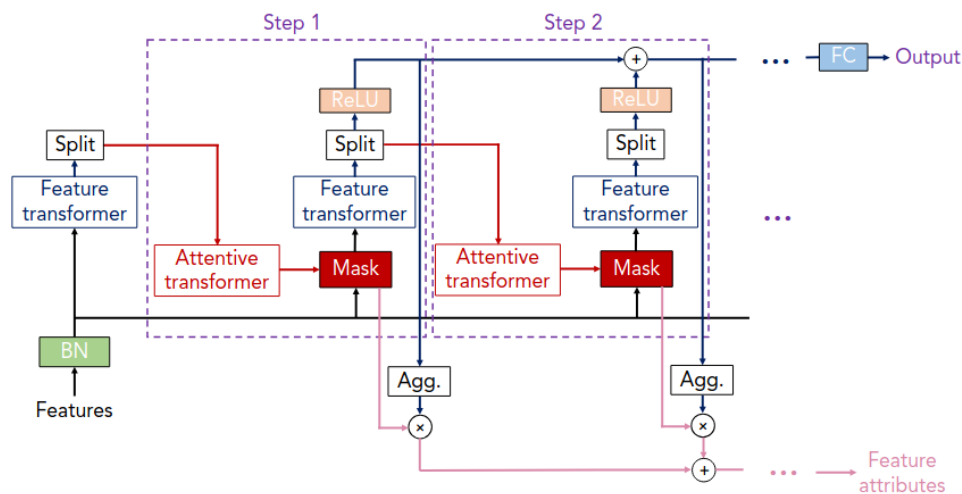
WILLIAMS, 1986).

As arquiteturas de redes neurais podem ser classificadas em três tipos: redes de camada única, com conexão direta entre entrada e saída; redes multicamadas, que incluem camadas ocultas para processar gradualmente as informações; e redes recorrentes, caracterizadas por laços de realimentação, nos quais as saídas retornam como entradas, permitindo o processamento de dados sequenciais (HAYKIN, 2011).

## 2.5. TABNET E FT-TRANSFORMER

O TabNet é um modelo de rede neural desenvolvido especificamente para dados tabulares, proposto por (ARIK; PFISTER, 2021). Diferentemente das abordagens tradicionais, ele utiliza uma arquitetura baseada em atenção sequencial e aprendizado supervisionado ou auto-supervisionado 4, permitindo que o modelo selecione dinamicamente quais atributos devem ser considerados em cada etapa da decisão. Essa estrutura melhora a interpretabilidade, pois possibilita visualizar quais características foram mais relevantes na predição. O TabNet também incorpora mecanismos de compressão de informações e sparsidade induzida, o que contribui para maior eficiência computacional e capacidade de generalização.

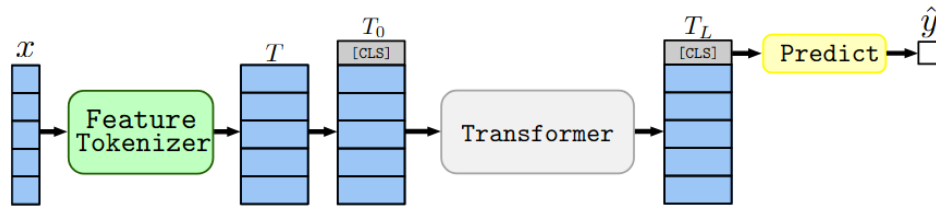
Figura 4 - Arquitetura *encoder* TabNet



Fonte:(ARIK; PFISTER, 2021).

O FT-Transformer (*Feature Tokenizer Transformer*) é uma variação dos modelos baseados em *Transformer* para lidar com dados tabulares, proposto por (GORISHNIY et al., 2021). Combina a tokenização de atributos com o mecanismo de autoatenção multi-cabeças 5, originalmente popularizado em tarefas de linguagem natural, mas adaptado para capturar relações entre características numéricas e categóricas. Ao tratar cada *feature* como um *token* independente, o *FT-Transformer* consegue modelar interações complexas entre variáveis, respeitando a estrutura tabular dos dados. Além disso, o uso de *embeddings* e camadas normalizadas favorece o treinamento estável em *datasets* com alta dimensionalidade, como os provenientes de sinais biomédicos ou neurofisiológicos.

Figura 5 - Arquitetura FT-Transformer



Fonte:(GORISHNIY et al., 2021).

## 2.6. BIBLIOTECAS E FERRAMENTAS

Este estudo utiliza diversas bibliotecas para pré-processamento, análise e modelagem, com o Python versão 3.10.\*. Para manipulação de dados, são empregadas as bibliotecas *NumPy* e *Pandas*, que auxiliam em operações numéricas e organização de dados em tabelas. As bibliotecas *Matplotlib* e *Seaborn* são utilizadas para visualização de gráficos e resultados dos modelos. O *Scikit-Learn* oferece ferramentas para padronização, divisão de atributos e avaliação de desempenho. Por fim, o *PyTorch* é usado para construir e treinar redes neurais.<sup>3</sup>

## 2.7. TRABALHOS RELACIONADOS

Esta subseção apresenta uma revisão de trabalhos relacionados, destacando pesquisas que se correlacionam com o estudo e suas principais contribuições.

### ***Automated depression detection using deep representation and sequence learning with eeg signals***

(AY et al., 2019) propôs um método automatizado para detectar depressão usando sinais de EEG brutos, baseado em um modelo híbrido CNN-LSTM, sem necessidade de extração ou seleção de características. Utilizando sinais dos hemisférios cerebrais, o processo aplica os sinais ao CNN, cujas características alimentam o LSTM para classificar os sinais como **normal** ou **deprimido**. Os dados foram coletados de 30 indivíduos (15 saudáveis e 15 deprimidos), com um filtro de 50 Hz para ruídos. Usou-se validação cruzada ( $k = 10$ ), com 70% dos dados para treinamento, 15% para validação e 15% para teste. Após 15 épocas, o modelo obteve precisões de 0.9934 (treinamento) e 0.9898 (validação) no hemisfério direito, e 0.9837 e 0.9868 no esquerdo. No teste, o hemisfério direito alcançou 0.9912 de precisão, com sensibilidade de 0.9855 para **deprimido** e 0.9970 para **normal**; o esquerdo, 0.9766, com 0.9703 para **deprimido** e 0.9827 para **normal**. A validação cruzada obteve médias de 0.9895 (direito) e 0.9776 (esquerdo). Apesar do desempenho promissor para ambientes clínicos, a principal limitação foi o pequeno conjunto de dados.

### ***Automated Rest EEG-Based Diagnosis of Depression and Schizophrenia Using a Deep Convolutional Neural Network***

(WANG et al., 2022) a pesquisa propôs a rede neural MUCHf-Net para classificar EEGs de indivíduos com depressão (DP), esquizofrenia (SCZ) e normais. Os resultados mostraram que a rede obteve 0.9112 de precisão na distinção entre EEGs normais

<sup>3</sup><<https://www.python.org/>>, <<https://numpy.org/>>, <<https://pandas.pydata.org/>>, <<https://matplotlib.org/>>, <<https://seaborn.pydata.org/>>, <<https://scikit-learn.org/stable/>>, <<https://pytorch.org/>>

e patológicos, com maior contribuição das bandas de baixa frequência e das regiões frontais e parietais do cérebro. A análise também indicou que o modelo teve mais dificuldade em diferenciar DP de SCZ, destaca-se o potencial do aprendizado profundo para auxiliar no diagnóstico de transtornos psiquiátricos.

### ***ETSNet: A deep neural network for EEG-based temporal–spatial pattern recognition in psychiatric disorder and emotional distress classification***

(SHAH et al., 2023) a pesquisa categoriza indivíduos com esquizofrenia (SZ), parentes biológicos (REL) e controles saudáveis (HC) usando sinais de EEG em repouso de 78 regiões corticais. A rede neural profunda ETSNet obteve 0.9957 de precisão para classificar SZ, REL e HC com olhos abertos (EO), e 0.9315 para olhos fechados (EC). Os dados, pré-processados com filtros passa-alta e passa-baixa, foram testados em dois conjuntos e treinados em três bases: uma não pública, outra com transtornos intelectuais e uma sobre emoções de 15 sujeitos. O treinamento usou uma 1D-CNN, com três grupos (SZ, REL, HC), dois estados (EO, EC), validação k-fold (k=5) e até 150 épocas. O modelo demonstra alta precisão e é útil para distinguir condições de saúde mental, incluindo esquizofrenia e alto risco de desenvolvimento do transtorno.

### **ALGORITMOS GENÉTICOS E APRENDIZADO PROFUNDO BASEADO EM REDES NEURAIS RECORRENTES DO TIPO LSTM PARA AUXÍLIO AO DIAGNÓSTICO MÉDICO**

(JÚNIOR, 2023) explora o uso de EEG para predição de tratamento em casos de AVC (Acidente Vascular Cerebral) e coma, utilizando redes LSTM (*long short-term memory*). Características extraídas do EEG são inseridas na camada intermediária, e um algoritmo genético (AG) é usado para seleção de características. O modelo combina redes CNN (*convolutional neural network*) e LSTM, processando características no domínio da frequência e do tempo e sinais brutos de EEG até a camada de saída. O AG também otimiza hiper-parâmetros. Na base de dados de coma (80% treinamento, 20% teste), o modelo CNN-LSTM combinado com características obteve acurácia de  $0.833 \pm 0.091$  e *f1-score* de  $0.826 \pm 0.092$ , enquanto o AG alcançou acurácia de  $0.933 \pm 0.082$  e *f1-score* de  $0.938 \pm 0.076$ . Na etiologia, o CNN-LSTM teve acurácia de  $0.617 \pm 0.125$  e *f1-score* de  $0.576 \pm 0.102$ . Para AVC, o AG obteve acurácia de 0.763 e *f1-score* de 0.746. As contribuições incluem: (1) um modelo de *deep learning* para análise de EEG e prognóstico em coma e AVC; (2) um modelo híbrido CNN-LSTM, aplicável a *Alzheimer*, distúrbios do sono e epilepsia, melhorando diagnóstico e tratamento.

### ***Psychiatric disorders from EEG signals through deep learning models***

(AHMED et al., 2024) analisa sinais de EEG com modelos de *deep learning* (ANN, KNN, LSTM, Bi-LSTM, CNN-LSTM) para compreender transtornos psiquiátricos. Os dados, previamente processados, passaram por extração de características via PSD e FC em diferentes bandas de frequência, com eliminação de atributos irrelevantes e discretização. Aplicaram-se cinco modelos com validação cruzada e 100 épocas. A ANN atingiu 0.9683 de precisão para transtorno obsessivo-compulsivo com todas as bandas. O CNN-LSTM obteve desempenho equivalente para transtorno de ajustamento. KNN e LSTM alcançaram 0.9894 para transtorno de estresse agudo, e KNN com Bi-LSTM, 0.9788 para transtorno obsessivo-compulsivo. As limitações incluem

número reduzido de sujeitos e pequeno tamanho amostral. Apesar da alta precisão, os modelos podem ser limitados para prognósticos de longo prazo. A principal contribuição foi superar a precisão de estudos anteriores utilizando o *dataset* de (PARK et al., 2021).

### ***Impact of Feature Selection Techniques on the Performance of Machine Learning Models for Depression Detection Using EEG Data***

(HASSAN; KAABOUCH, 2024) investiga como técnicas de seleção de características influenciam o desempenho de modelos de aprendizado de máquina na detecção de depressão com dados de EEG. Foram testados seis métodos de seleção, como Elastic Net, Mutual Information, SVM com eliminação recursiva de atributos e mRMR, combinados com seis classificadores, incluindo SVM, Random Forest e XGBoost. A combinação de SVM com eliminação recursiva obteve os melhores resultados, com 93,54% de acurácia e 95,29% de F1. O estudo reforça a importância da escolha criteriosa de atributos para otimizar a detecção de depressão.

### ***Mental Disorder Assessment in IoT-Enabled WBAN Systems with Dimensionality Reduction and Deep Learning***

(OLATINWO; ABU-MAHFOUZ; MYBURGH, 2025) propõe um modelo híbrido de aprendizado profundo que combina CNN e BiLSTM para avaliar transtornos mentais com dados de EEG. Para melhorar o desempenho, os autores utilizam imputação k-NN para lidar com dados ausentes, SMOTE para balancear classes e PCA para reduzir a dimensionalidade. O modelo incorpora o método LIME para tornar as previsões mais interpretáveis e confiáveis. Os resultados mostram que o modelo supera os métodos de *baseline*, alcançando 76% de acurácia na classificação de transtornos principais (vs. 52–74% dos outros modelos) e 80% para transtornos específicos (vs. 65–78%), destacando seu potencial diagnóstico.

## **3. METODOLOGIA**

Esta seção descreve os materiais e métodos utilizados, detalhando as etapas seguidas durante o desenvolvimento do trabalho, as técnicas aplicadas e os processos implementados para atingir os objetivos propostos.

### **3.1. Dataset**

Conforme o *dataset* apresentado por (PARK, 2021), os dados foram coletados retrospectivamente a partir de registros médicos de pacientes em estado de repouso entre janeiro de 2011 e dezembro de 2018. O conjunto inclui avaliações psicológicas e EEG quantitativo (QEEG) do *SMG-SNU Boramae Medical Center*, com diagnósticos baseados nos critérios do DSM-5 e entrevistas neuropsiquiátricas. A confirmação clínica final foi realizada por dois psiquiatras e dois psicólogos entre março e agosto de 2019.

O conjunto de dados inclui 945 pacientes, com idades entre 18 e 70 anos, média de  $30.59 \pm 11.78$  anos, escolaridade de  $13.43 \pm 2.55$  anos e QI médio de  $101.58 \pm 17.02$ . As categorias de diagnóstico incluem 6 classes em nível principal (*main disorder*) e 9 classes em nível específico (*specific disorder*). O EEG foi registrado com 19 canais

conforme o sistema internacional 10-20, utilizando os eletrodos: FP1, FP2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, O2.

O pré-processamento descrito por (PARK et al., 2021) utilizou a Transformada Rápida de *Fourier* (FFT) para converter os sinais para o domínio da frequência (0,5 a 40 Hz), com resolução de 0,5 Hz e janela cônica de cosseno. A Densidade Espectral de Potência (PSD) foi calculada nas bandas *delta*, *theta*, *alpha*, *beta*, *high beta* e *gamma*, enquanto a Coerência Funcional (FC) avaliou a sincronização entre regiões cerebrais. Os sinais foram segmentados em épocas de 2 segundos com 75% de sobreposição, permitindo análise contínua das interações neurais.

A extração de características resultou em 114 atributos via PSD e 1026 via FC, totalizando 1140. Foram também incluídas variáveis quantitativas (*age*, *education*, *IQ*), qualitativas (*main.disorder*, *specific.disorder*, *sex*) e a data do EEG (*eeg.date*), compondo um *dataset* final com dimensões 945 x 1148.

O estudo foi aprovado pelo Comitê de Ética em Pesquisa do *SMG-SNU Boramae Medical Center*, em Seul, Coreia do Sul.

### 3.2. ANÁLISE EXPLORATÓRIA

A tabela 1 resume características demográficas e cognitivas por transtorno. Transtornos aditivos e de ansiedade social têm maioria masculina, enquanto os relacionados a trauma e estresse apresentam maior proporção de mulheres. A maior idade média ocorre em estresse pós-traumático (42,7 anos), e o maior QI e escolaridade estão nos controles saudáveis (116,2 e 14,9 anos). Já a esquizofrenia apresenta os menores valores médios de QI (89,6) e escolaridade (12,8), indicando possível prejuízo cognitivo. O número de amostras por grupo indica um desbalanceamento entre as classes. Os gráficos complementares estão disponíveis no material suplementar (*GitHub*: <https://gitlab.com/matt-balda/tcc-eeg-signals-neuropsychiatrics>).

Tabela 1 - Estatísticas descritivas por transtorno principal e específico.

Transtorno principal	Transtorno específico	N	Sexo (%H/%M)	Id (M ± DP)	min-máx	Esc (M ± DP)	min-máx	QI (M ± DP)	min-máx
Dependência	Total	186	88.2 / 11.8	29.6 ± 10.9	18-68	13.2 ± 2.5	1-18	103.9 ± 16.2	56-142
	Uso de álcool	93	80.6 / 19.4	34.2 ± 11.9	19-64	13.3 ± 3.1	1-18	103.4 ± 13.6	68-130
	Dependência comportamental	93	95.7 / 4.3	25.1 ± 7.5	18-68	13.2 ± 1.9	7-18	104.4 ± 18.5	56-142
Ansiedade	Total	107	73.8 / 26.2	29.0 ± 10.6	19-61	13.1 ± 2.4	0-18	98.3 ± 16.3	49-142
	Pânico	59	64.4 / 35.6	31.0 ± 11.3	19-61	13.4 ± 2.9	0-18	100.3 ± 14.8	72-142
	Ansiedade social	48	85.4 / 14.6	26.5 ± 9.1	19-58	12.8 ± 1.6	10-16	95.9 ± 17.9	49-128
Controle saudável	Total								
	Controle saudável	95	63.2 / 36.8	25.7 ± 4.5	18-42	14.9 ± 2.1	11-20	116.2 ± 10.9	88-141
Humor	Total	266	56.8 / 43.2	30.9 ± 12.7	18-70	13.3 ± 2.5	1-19	101.6 ± 15.7	60-145
	Bipolar	67	62.7 / 37.3	29.7 ± 11.0	18-62	14.1 ± 2.2	6-19	100.8 ± 17.0	60-134
	Depressivo	199	54.8 / 45.2	31.3 ± 13.2	18-70	13.0 ± 2.5	1-18	101.8 ± 15.3	61-145
Obsessivo-compulsivo	Total								
	Obsessivo-compulsivo	46	82.6 / 17.4	28.5 ± 9.8	18-70	13.9 ± 2.3	5-18	107.8 ± 15.2	76-138
Esquizofrenia	Total								
	Esquizofrenia	117	55.6 / 44.4	31.7 ± 12.1	19-72	12.8 ± 2.9	0-18	89.6 ± 17.5	50-131
Trauma e estresse relacionados	Total	128	34.4 / 65.6	36.1 ± 13.8	18-68	13.6 ± 2.4	6-20	98.9 ± 15.9	53-137
	Estresse agudo	38	7.9 / 92.1	28.9 ± 9.0	19-55	14.3 ± 2.3	8-20	104.1 ± 15.4	76-134
	Adaptação	38	71.1 / 28.9	34.2 ± 14.9	18-67	13.2 ± 2.4	6-18	94.2 ± 15.4	60-122
	Estresse pós-traumático	52	26.9 / 73.1	42.7 ± 13.0	20-68	13.4 ± 2.5	6-18	98.9 ± 15.7	53-137

Nota. N: amostras; H: homens; M: mulheres; Id: idade; Esc: anos de escolaridade; QI: quociente de inteligência; M: média; DP: desvio-padrão.

### 3.3. PRÉ-PROCESSAMENTO

Foram identificados valores *NaN* (*Not a Number*),  $n = 28$ , nas variáveis *education* e *IQ*. Para imputar e substituir esses valores, foi aplicado o *KNNImputer*<sup>4</sup>, com  $n\_neighbors = 5$ . Essa técnica utiliza o cálculo da distância euclidiana, configurando a métrica como *nan\_euclidean*.

Para a identificação de *outliers*, foi utilizado o IQR (*Interquartile Range*), com factor = 1.5, e a winsorização foi realizada por meio da função *clip* do NumPy<sup>5</sup>, uma vez que essa função delimita um intervalo fechado, recortando os *outliers* e substituindo-os pelos limites. Foram identificados mais de 26 mil *outliers*. Considerando o *shape* de  $(945 \times 1143)$ , com três variáveis quantitativas (idade, escolaridade e QI), temos mais de 1 milhão de campos na matriz, representando aproximadamente 2,44% de *outliers* em relação ao conjunto total. A remoção ou substituição desses *outliers* é opcional, uma vez que não afetaria significativamente a convergência durante o treinamento.

### 3.4. PREPARAÇÃO DOS DADOS

As classes 6 e 9 das variáveis *main.disorder* e *specific.disorder* foram reagrupadas em um problema de classificação binária: **controle saudável vs. transtorno**, com os rótulos codificados como  $[0, 1]$ . Como mostrado na Tabela 1, há um desbalanceamento entre essas classes. Para corrigir esse desequilíbrio, foi aplicada a técnica de *oversampling*, utilizando como referência a classe com maior número de amostras. O método escolhido foi o **SMOTE** (*Synthetic Minority Oversampling Technique*)<sup>6</sup>, que gera exemplos sintéticos para a classe minoritária por meio de interpolação entre amostras vizinhas no espaço de características.

Após o balanceamento, o conjunto de dados foi dividido em subconjuntos baseados em três grupos de atributos: **PSD**, **FC** e **PSD + FC**. Cada grupo foi combinado separadamente para seis bandas de frequência (*delta*, *theta*, *alpha*, *beta*, *high beta* e *gamma*) e também considerando todas as bandas combinadas, totalizando **21 combinações por transtorno**. Dessa forma, foram definidos os subconjuntos  $X$  (atributos) e  $y$  (rótulos binários) correspondentes a cada cenário de combinado.

No uso do *PyTorch*<sup>7</sup>, é necessário converter os dados do formato *array* para *tensor*, estrutura que generaliza os conceitos de vetor e matriz, podendo ter uma ou mais dimensões (1D, 2D, 3D, ..., nD). Essa conversão é essencial para que os dados sejam compatíveis com as redes neurais do *framework*. Em seguida, os dados devem ser divididos em conjuntos de treino e teste. Neste caso, utilizou-se uma divisão de 80% para treinamento e 20% para teste.

O *DataLoader*<sup>8</sup> é uma classe do *PyTorch* que facilita o carregamento e a iteração sobre grandes conjuntos de dados, organizando-os em pequenos lotes (*batches*), o que aumenta a eficiência durante o treinamento. O parâmetro *batch size* define quantas amostras são processadas por vez (neste caso, 32), enquanto o parâmetro *shuffle* indica se os dados devem ser embaralhados antes de cada época, *True* no treino e *False* no teste.

<sup>4</sup><<https://scikit-learn.org/1.5/modules/generated/sklearn.impute.KNNImputer.html>>

<sup>5</sup><<https://numpy.org/doc/stable/reference/generated/numpy.clip.html>>

<sup>6</sup><[https://imbalanced-learn.org/stable/references/generated/imblearn.over\\_sampling.SMOTE.html](https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTE.html)>

<sup>7</sup><<https://pytorch.org/>>

<sup>8</sup><[https://pytorch.org/tutorials/beginner/basics/data\\_tutorial.html](https://pytorch.org/tutorials/beginner/basics/data_tutorial.html)>

## 3.5. IMPLEMENTAÇÃO DOS MODELOS

### 3.5.1. *Perceptron* MULTICAMADA

Foi utilizado um *perceptron* multicamada com uma camada de entrada, quatro camadas ocultas contendo 128, 64, 32 e 16 neurônios, respectivamente, e uma camada de saída com uma unidade. Cada camada oculta é seguida por *Batch Normalization*, ativação GELU e *dropout* de 10% para mitigar o *overfitting*, com inicialização dos pesos pelo método *Xavier Uniform*. A função de perda adotada foi a *BCEWithLogitsLoss*, adequada para problemas de classificação binária, e a otimização foi realizada com o algoritmo *Adam*, utilizando taxa de aprendizado de  $1 \times 10^{-3}$ , regularização L2 (*weight decay*) de  $1 \times 10^{-5}$  e ajuste automático da taxa via *ReduceLROnPlateau*, que reduz o valor em 10% após 3 épocas sem melhora, até um mínimo de  $1 \times 10^{-6}$ . O modelo foi treinado por até 200 épocas com *early stopping* configurado com paciência de 20 épocas, monitorando a perda do treinamento a cada época para possível interrupção do treinamento e atualização do agendador.

### 3.5.2. TABNET

O treinamento foi realizado com o otimizador *Adam*, conhecido por sua eficiência em problemas de alta dimensionalidade e dados ruidosos. Além disso, empregou-se um agendador de taxa de aprendizado do tipo StepLR, com parâmetros `step_size=30` e `gamma=0.9`, o que implica na redução da taxa de aprendizado em 10% a cada 30 épocas, favorecendo a convergência estável do modelo.

### 3.5.3. FT-TRANSFORMER

O modelo foi configurado com entrada exclusivamente numérica, desconsiderando variáveis categóricas (`cat_cardinalities = None`). O número de atributos de entrada foi ajustado de acordo com o número de colunas do conjunto de treinamento (`n_num_features = X_train_array.shape[1]`). Para a tarefa de classificação binária, a saída foi definida como um único neurônio (`d_out = 1`). Além disso, utilizou-se o parâmetro `last_layer_query_idx = [-1]`, permitindo que a saída final seja baseada exclusivamente no último *token* do transformador, conforme recomendado para problemas de classificação.

## 3.6. MÉTRICAS DE AVALIAÇÃO

As métricas utilizadas para avaliar o desempenho dos modelos de classificação foram: *accuracy* (proporção de previsões corretas em relação ao total de previsões), *Recall* (ou *True Positive Rate*, *Sensitivity*) 3, *Specificity* 2 e *False Positive Rate (FPR)* 4, conforme (BROWNLEE, 2021).

Onde: *TP*: Verdadeiros positivos, *TN*: Verdadeiros negativos, *FP*: Falsos positivos e *FN*: Falsos negativos.

### **Specificity**

Proporção de negativos corretamente classificados:

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (2)$$

### ***TPR / Recall / Sensitivity***

Proporção de positivos corretamente classificados:

$$TPR = \frac{TP}{TP + FN} \quad (3)$$

### ***FPR***

Proporção de negativos classificados como positivos:

$$FPR = \frac{FP}{FP + TN} = 1 - Specificity \quad (4)$$

## **4. RESULTADO E DISCUSSÃO**

Nesta seção, são apresentados os resultados obtidos para as classes *main disorder* e *specific disorder*. As tabelas 2 e 3 mostram a acurácia, a AUC e a sensibilidade associadas às melhores configurações de características, selecionadas com base nas acurácias globais dos três modelos (ANN, FTT e TabNet). As Figuras 6 e 7 exibem os valores de acurácia e AUC correspondentes às melhores acurácias obtidas por transtorno. Por fim, a tabela 4 apresenta a acurácia, a AUC e a sensibilidade do conjunto *entire* (que inclui todas as características, como PSD e FC), também com base nas melhores acurácias por classe.

### **4.1. *Main disorder* (TRANSTORNO PRINCIPAL)**

De modo geral, o modelo TabNet obteve os melhores resultados em termos de acurácia, AUC e sensibilidade na maioria das classes, com destaque para os transtornos de dependência (96,00% de acurácia, 95,45% de AUC e 97,00% de sensibilidade) e humor (97,20% de acurácia, 96,72% de AUC e 94,00% de sensibilidade). O ANN também demonstrou desempenho competitivo, especialmente nos transtornos de humor e trauma, com AUC superiores a 97%. Em contraste, o FTT apresentou desempenho mais modesto, embora tenha se destacado na classificação de esquizofrenia, alcançando AUC de 93,84% e sensibilidade de 78,26%. Esses resultados sugerem que a escolha do modelo e das características extraídas (PSD e FC) impactam significativamente a performance, sendo a TabNet particularmente eficaz na generalização entre classes.

Tabela 2 - Desempenho, em porcentagem, dos modelos na classificação da classe *main disorder* entre controles saudáveis e transtornos.

Transtorno	Melhor carac.	Acurácia (%)			AUC (%)			Sensibilidade (%)		
		ANN	FTT	TabNet	ANN	FTT	TabNet	ANN	FTT	TabNet
Dependência	Beta PSD + FC	84,00	78,67	<b>96,00</b>	92,89	80,65	<b>95,45</b>	83,78	81,08	<b>97,00</b>
Ansiedade	Delta PSD	79,07	76,74	<b>90,70</b>	85,71	85,93	<b>89,18</b>	71,43	66,67	<b>81,00</b>
Humor	Beta PSD + FC	95,33	87,85	<b>97,20</b>	<b>99,20</b>	93,92	96,72	92,45	86,79	<b>94,00</b>
Obsessivo-compulsivo	Delta PSD	<b>92,11</b>	81,58	76,32	<b>98,89</b>	87,81	75,90	<b>94,74</b>	73,68	74,00
Esquizofrenia	Beta PSD	80,85	87,23	<b>91,49</b>	86,41	93,84	<b>95,83</b>	69,57	78,26	<b>91,00</b>
Trauma e estresse	PSD completa	<b>94,23</b>	90,38	88,46	<b>97,63</b>	94,90	93,05	88,46	<b>92,31</b>	85,00

Nota: Melhor conjunto de características (**Melhor carac.**) por transtorno. AUC = área sob a curva ROC; FC = conectividade funcional e PSD = densidade espectral de potência.

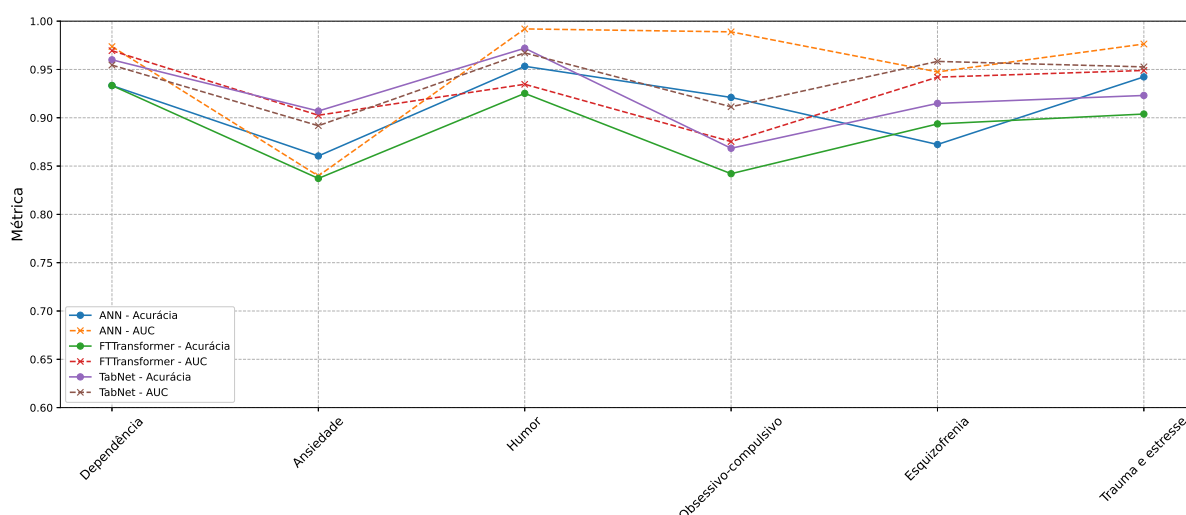


Figura 6 - Comparação da acurácia máxima e da AUC entre modelos para a classificação de transtornos principais.

#### 4.1.1. *Specific disorder* (TRANSTORNO ESPECÍFICO)

Os resultados revelam que não há um modelo universalmente superior, pois o desempenho varia conforme o transtorno. O modelo TabNet destacou-se nos transtornos depressivo (96,20% de acurácia e 96,60% de AUC) e bipolar (89,40% de acurácia e 90,50% de AUC), enquanto o ANN foi mais eficaz na detecção de transtornos como pânico (97,30% de acurácia, 99,40% de AUC e 100% de sensibilidade) e adaptação (94,70% de acurácia). Já o FTT obteve os melhores resultados na classificação do transtorno de estresse pós-traumático (97,30% de acurácia e 100% de sensibilidade) e no uso de álcool (92,10% de acurácia e 91,10% de AUC). A TabNet foi escolhida como referência por apresentar desempenho consistente em diferentes condições, como nos transtornos depressivo, bipolar e de dependência.

Tabela 3 - Desempenho, em porcentagem, dos modelos na classificação da classe *specific disorder* entre controles saudáveis e transtornos.

Transtorno	Melhor carac.	Acurácia (%)			AUC (%)			Sensibilidade (%)		
		ANN	FTT	TabNet	ANN	FTT	TabNet	ANN	FTT	TabNet
Estresse agudo	Delta PSD	<b>92,10</b>	86,80	92,10	95,50	92,20	<b>97,20</b>	94,70	<b>100,00</b>	84,00
Adaptação	Beta PSD	<b>94,70</b>	84,20	86,80	<b>95,50</b>	89,10	89,40	<b>89,40</b>	74,00	79,00
Uso de álcool	Beta alta FC	78,90	<b>92,10</b>	65,70	85,80	<b>91,10</b>	58,70	78,90	89,00	<b>95,00</b>
Depen. comportamental	Theta PSD + FC	73,60	73,60	<b>81,50</b>	76,10	<b>79,70</b>	79,70	73,60	74,00	<b>79,00</b>
Bipolar	Beta PSD	78,90	78,90	<b>89,40</b>	86,40	87,50	<b>90,50</b>	68,40	79,00	<b>89,00</b>
Depressivo	Beta FC	92,50	87,50	<b>96,20</b>	96,10	93,80	<b>96,60</b>	<b>95,00</b>	85,00	93,00
Pânico	PSD completa	<b>97,30</b>	89,40	78,90	<b>99,40</b>	94,70	83,30	<b>100,00</b>	84,00	95,00
Estresse pós-traumático	Gama PSD + FC	89,40	<b>97,30</b>	73,60	<b>97,50</b>	94,70	77,80	<b>100,00</b>	100,00	79,00
Ansiedade social	FC completa	<b>86,80</b>	76,30	78,90	<b>94,40</b>	85,50	85,80	<b>84,20</b>	79,00	84,00

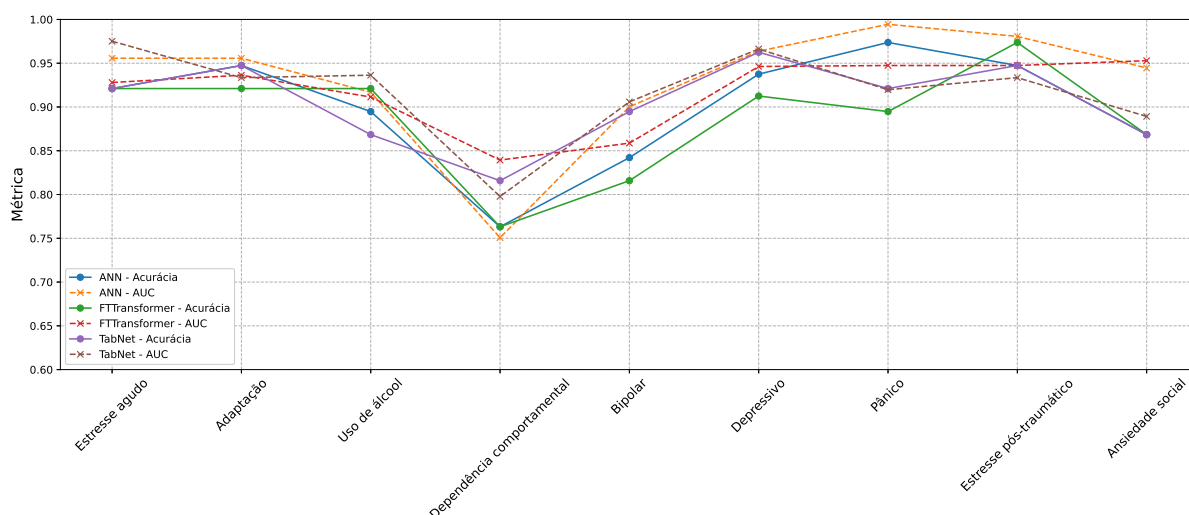


Figura 7 - Comparação da acurácia máxima e da AUC entre modelos para a classificação de transtornos específicos.

#### 4.2. Main / *specific disorder* entire

O TabNet obteve os melhores resultados para transtornos de humor (93,50% de acurácia e 93,19% de AUC) e trauma e estresse (89,35% de AUC), enquanto o ANN se destacou na esquizofrenia (94,38% de AUC) e em transtornos depressivos e bipolares. Já o FTTransformer apresentou desempenho expressivo em subclasses específicas, como estresse agudo e transtorno de adaptação, com AUC superiores a 90%. Para a classe *main disorder*, qualquer um dos modelos pode ser utilizado, enquanto para a classe *specific disorder* o modelo FTTransformer.

Tabela 4 - Desempenho do (conjunto completo) dos modelos TabNet, ANN e FTTransformer para as classes *main* e *specific disorder*.

Tipo de classe	Transtorno	Modelo	Acurácia (%)	AUC (%)	Sensibilidade (%)
Main Disorder	Transtorno dependência	ANN	84,00	92,39	84,00
	Transtorno de ansiedade	FTTransformer	76,70	72,73	61,90
	Transtorno de humor	TabNet	93,50	93,19	89,00
	Transtorno obsessivo-compulsivo	FTTransformer	84,20	87,53	89,50
	Esquizofrenia	ANN	87,20	94,38	87,00
	Trauma e estresse	TabNet	86,50	89,35	88,00
Specific Disorder	Estresse agudo	FTTransformer	86,80	90,90	100,00
	Transtorno de adaptação	FTTransformer	92,10	93,60	95,00
	Transtorno por uso de álcool	ANN	81,60	89,80	79,00
	Dependência comportamental	FTTransformer	71,10	73,70	84,00
	Transtorno bipolar	ANN	81,60	88,10	79,00
	Transtorno depressivo	ANN	86,30	89,90	83,00
	Transtorno de pânico	FTTransformer	78,90	85,00	84,00
	Estresse pós-traumático	FTTransformer	84,20	97,20	95,00
	Ansiedade social	ANN	81,60	90,60	84,00

Nota: AUC = área sob a curva ROC.

## 5. CONSIDERAÇÕES FINAIS E TRABALHOS FUTUROS

Este trabalho teve como objetivo aplicar abordagens de aprendizado profundo para classificar condições neuropsiquiátricas a partir de sinais de EEG, motivado pela complexidade dos dados neurofisiológicos e pela importância clínica do diagnóstico precoce.

Foram avaliadas diferentes arquiteturas, como ANN, TabNet e FT-Transformer, combinadas a pré-processamentos com parâmetros de densidade espectral de potência (PSD) e conectividade funcional (FC), visando à classificação binária entre controles saudáveis e pacientes. Em comparação com os resultados de (PARK et al., 2021), os modelos propostos apresentaram desempenho superior nas classificações *main disorder* e *specific disorder*, com acurácias médias de 93,62% e 91,93%, respectivamente. Trabalhos anteriores (MOTA et al., 2025; OLATINWO; ABU-MAHFOUZ; MYBURGH, 2025) utilizaram classificações multiclasse, mas enfrentaram dificuldades de convergência, com acurácias entre 60% e 80%.

Apesar de limitações como o número reduzido de amostras, o desbalanceamento entre classes, a heterogeneidade entre PSD e FC e a variabilidade entre sujeitos, os resultados indicam que modelos de aprendizado profundo, quando bem ajustados, superam os métodos tradicionais de aprendizado de máquina. Abrem-se, assim, possibilidades para estudos futuros envolvendo a geração de dados sintéticos mais realistas, a explicabilidade dos modelos, a combinação de classificadores binários e a agregação de dados multimodais de naturezas distintas.

Reforça-se o papel da inteligência artificial na saúde mental, mostrando o potencial da integração entre neurociência e aprendizado profundo na detecção de transtornos neuropsiquiátricos.

## REFERÊNCIAS

AHMED, Z. et al. Psychiatric disorders from eeg signals through deep learning models. **IBRO Neuroscience Reports**, Elsevier, 2024.

ARIK, S. Ö.; PFISTER, T. Tabnet: Attentive interpretable tabular learning. In: **Proceedings of the AAAI conference on artificial intelligence**. [S.l.: s.n.], 2021. v. 35, n. 8, p. 6679–6687.

ASSOCIATION, A. P. **Diagnostic and statistical manual of mental disorders: DSM-5**. [S.l.]: American psychiatric association Washington, DC, 2013. v. 5.

AY, B. et al. Automated depression detection using deep representation and sequence learning with eeg signals. **Journal of Medical Systems**, Springer New York LLC, v. 43, p. 1–12, 7 2019. ISSN 1573689X. Disponível em: <<https://link.springer.com/article/10.1007/s10916-019-1345-y>>.

BENICZKY, S.; SCHOMER, D. L. Electroencephalography: basic biophysical and technological aspects important for clinical applications. **Epileptic Disorders**, JLE Editions, v. 22, p. 697–715, 12 2020. ISSN 1294-9361. Disponível em: <<https://stm.cairn.info/revue-epileptic-disorders-2020-6-page-697?lang=en&tab=texte-integralhttps://stm.cairn.info/revue-epileptic-disorders-2020-6-page-697?lang=en&tab=resumehttps://stm.cairn.info/revue-epileptic-disorders-2020-6-page-697>>.

BROWNLEE, J. **Tour of Evaluation Metrics for Imbalanced Classification**. 2021. Acessado em: 27 nov. 2024. Disponível em: <<https://machinelearningmastery.com/tour-of-evaluation-metrics-for-imbalanced-classification/>>.

GOODFELLOW, I. **Deep learning**. [S.l.]: MIT press, 2016.

GORISHNIY, Y. et al. Revisiting deep learning models for tabular data. **Advances in neural information processing systems**, v. 34, p. 18932–18943, 2021.

HASSAN, M.; KAABOUCH, N. Impact of feature selection techniques on the performance of machine learning models for depression detection using eeg data. **Applied Sciences**, v. 14, n. 22, 2024. ISSN 2076-3417. Disponível em: <<https://www.mdpi.com/2076-3417/14/22/10532>>.

HAYKIN, S. Neural networks: principles and practice. **Bookman**, v. 11, n. 900, 2011.

JÚNIOR, S. B. Algoritmos genéticos e aprendizado profundo baseado em redes neurais recorrentes do tipo lstm para auxílio ao diagnóstico médico. 2023.

KOUDELKOVÁ, Z.; STRMISKA, M. Introduction to the identification of brain waves based on their frequency. **MATEC Web of Conferences**, EDP Sciences, v. 210, p. 05012, 10 2018. ISSN 2261-236X. Disponível em: <<https://www.matec-conferences.org/10.1051/matecconf/201821005012>>.

LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.

MONTENEGRO, M. A. et al. **EEG na Prática Clínica**. 4th. ed. [S.l.]: Thieme Revinter, 2022. ISBN 9786555721607.

MOTA, M. B. et al. Neuropsychiatric disorders classification using eeg signal and deep neural networks. In: SBC. **Simpósio Brasileiro de Computação Aplicada à Saúde (SBCAS)**. [S.l.], 2025. p. 287–292.

MÜLLER-PUTZ, G. R. Electroencephalography. **Handbook of Clinical Neurology**, Elsevier, v. 168, p. 249–262, 1 2020. ISSN 0072-9752.

NARDI, A. E.; SILVA, A. G. da; QUEVEDO, J. **Tratado de psiquiatria da Associação Brasileira de Psiquiatria**. [S.l.]: Artmed Editora, 2021. ISBN 9786558820345.

OLATINWO, D.; ABU-MAHFOUZ, A.; MYBURGH, H. Mental disorder assessment in iot-enabled wban systems with dimensionality reduction and deep learning. **Journal of Sensor and Actuator Networks**, v. 14, n. 3, 2025. ISSN 2224-2708. Disponível em: <<https://www.mdpi.com/2224-2708/14/3/49>>.

ORGANIZATION, W. H. **International Statistical Classification of Diseases and related health problems: Alphabetical index**. [S.l.]: World Health Organization, 2004. v. 3.

PARK, S. M. **EEG machine learning**. 2021. Retrieved August 16, 2021, from <<https://osf.io/8bsvr>>.

PARK, S. M. et al. Identification of major psychiatric disorders from resting-state electroencephalography using a machine learning approach. **Frontiers in Psychiatry**, Frontiers Media S.A., v. 12, p. 707581, 8 2021. ISSN 16640640. Disponível em: <[www.frontiersin.org](http://www.frontiersin.org)>.

PARSA, M. et al. Eeg-based classification of individuals with neuropsychiatric disorders using deep neural networks: A systematic review of current status and future directions. **Computer Methods and Programs in Biomedicine**, Elsevier, v. 240, p. 107683, 2023.

PEDLEY, T. A.; DALY, D. D. **Current practice of clinical electroencephalography**. [S.l.]: Lippincott-Raven, 1997.

PEH, W. Y. et al. Multi-center validation study of automated classification of pathological slowing in adult scalp electroencephalograms via frequency features. **International Journal of Neural Systems**, World Scientific, v. 31, n. 06, p. 2150016, 2021.

RUMELHART, D. E.; HINTON, G. E.; WILLIAMS, R. J. Learning representations by back-propagating errors. **nature**, Nature Publishing Group UK London, v. 323, n. 6088, p. 533–536, 1986.

RUSSELL, S. J.; NORVIG, P. **Artificial intelligence: a modern approach**. [S.l.]: Pearson, 2022.

Sapien Labs. **The Mental State of the World Report 2023**. 2024. Accessed: 2024-10-10. Disponível em: <[https://mentalstateoftheworld.report/2023\\_read/](https://mentalstateoftheworld.report/2023_read/)>.

SHAH, S. J. H. et al. Etsnet: A deep neural network for eeg-based temporal–spatial pattern recognition in psychiatric disorder and emotional distress classification. **Computers in Biology and Medicine**, Elsevier, v. 158, p. 106857, 2023.

SIJLY, S.; LI, Y.; ZHANG, Y. Eeg signal analysis and classification. **IEEE Trans Neural Syst Rehabil Eng**, Springer, v. 11, p. 141–144, 2016.

SOMMERVILLE, I. Engenharia de software, 9a. **São Palo, SP, Brasil**, p. 63, 2011.

ULLAH, I. et al. An automated system for epilepsy detection using eeg brain signals based on deep learning approach. **Expert Systems with Applications**, Elsevier, v. 107, p. 61–71, 2018.

WANG, Z. et al. Automated rest eeg-based diagnosis of depression and schizophrenia using a deep convolutional neural network. **IEEE Access**, IEEE, v. 10, p. 104472–104485, 2022.

World Health Organization. **Mental disorders**. 2022. Accessed: 2024-10-10. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/mental-disorders>>.