

UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO  
CENTRO TECNOLÓGICO  
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA ELÉTRICA

DAVID WILKERSON KÜSTER

DETECÇÃO DE SINAIS ANORMAIS DE EEG POR MEIO DE  
MICROESTADOS E APRENDIZADO DE MÁQUINA

Vitória – ES  
2022



DAVID WILKERSON KÜSTER

# DETECÇÃO DE SINAIS ANORMAIS DE EEG POR MEIO DE MICROESTADOS E APRENDIZADO DE MÁQUINA

Dissertação apresentada ao Programa de Pós Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Universidade Federal do Espírito Santo  
Centro Tecnológico  
Programa de Pós Graduação em Engenharia Elétrica

Orientador: Prof. Dr. Patrick Marques Ciarelli

Vitória – ES  
2022





Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

---

K97d Küster, David Wilkerson, 1989-  
Detecção de sinais anormais de EEG por meio de Microestados e Aprendizado de Máquina / David Wilkerson Küster. - 2022. 83 f. : il.

Orientador: Patrick Marques Ciarelli.  
Dissertação (Mestrado em Engenharia Elétrica) - Universidade Federal do Espírito Santo, Centro Tecnológico.

1. Eletroencefalografia. 2. Aprendizado do computador. 3. Processamento de sinais. 4. Classificação. I. Ciarelli, Patrick Marques. II. Universidade Federal do Espírito Santo. Centro Tecnológico. III. Título.

CDU: 621.3

---

DAVID WILKERSON KÜSTER

## DETECÇÃO DE SINAIS ANORMAIS DE EEG POR MEIO DE MICROESTADOS E APRENDIZADO DE MÁQUINA

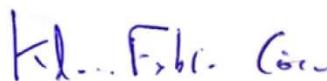
Dissertação apresentada ao Programa de Pós Graduação em Engenharia Elétrica do Centro Tecnológico da Universidade Federal do Espírito Santo, como requisito parcial para obtenção do título de Mestre em Engenharia Elétrica.

Trabalho aprovado. Vitória, 25 de março de 2022.



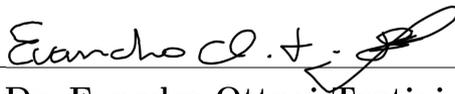
---

**Prof. Dr. Patrick Marques Ciarelli**  
Universidade Federal do Espírito Santo  
Orientador



---

**Prof. Dr. Klaus Fabian Côco**  
Universidade Federal do Espírito Santo  
Examinador



---

**Prof. Dr. Evandro Ottoni Teatini Salles**  
Universidade Federal do Espírito Santo  
Examinador

Vitória – ES  
2022



*Dedicado à minha família.*



## AGRADECIMENTOS

O tempo é engraçado. Cômico até, diria. Ele leva anseios e traz outros novos. Leva amigos, traz amigos. Em uma frequência nem sempre ressonante com nosso desejo. Novos sonhos despertam. Nesse rito de passagem, nasce uma nova era. A prova definitiva de que cada noite mal dormida, cada fim de semana envolto em equações infinitas, valeram muito. Foram o justo pedágio na estrada do conhecimento. Apenas uma palavra pode caber nesse ponto: Gratidão.

Agradeço àquele moleque que, na quarta série, se debruçou sobre o céu noturno estrelado pela primeira vez, e, ainda perplexo, sorriu: era tudo fantástico. Quase mágico. Hoje é tudo tão menos distante, que a beleza do encanto parece desvanecer um pouco. É preciso reaprender, com aquele fascinado garoto, a ver toda a beleza e imensidão com olhos mais humildes e perplexos. Aquele garoto talvez não pudesse compreender todo o fascínio da imagem que via. O garoto perplexo continua, certamente, perplexo. Mas toda a proporção e todas as fronteiras se expandiram mais um pouco! E ainda há muito a se ver. O tempo e a experiência demonstram seu valor.

Agradeço profundamente a todos os colegas e mestres do Departamento de Engenharia Elétrica da Universidade Federal do Espírito Santo (UFES) e, em particular ao Programa de Pós-Graduação em Engenharia Elétrica (PPGEE), agradeço também à Fundação de Amparo à Pesquisa do Espírito Santo (FAPES), que financiou parcialmente este trabalho, através do projeto de número 598/2018. Agradeço com muita alegria aos colegas e profissionais do Instituto Federal do Espírito Santo (IFES).

Agradeço muito ao meu orientador, pelas incontáveis contribuições, discussões e por todo o aprendizado durante esse processo.

Agradeço aos membros da banca por tomarem o valioso tempo de avaliar este trabalho, e por todas as suas contribuições.

Muito obrigado à maior amiga, também conselheira, psicóloga e esposa. À todas as An(n)as, que não são poucas. E um obrigado muito especial àqueles que, mesmo longe, são a parte maior dessa conquista: obrigado irmãs, irmão e sobrinhos. Obrigado, família. Muito obrigado Pai (em sua memória). Muito obrigado Mãe.

Muito obrigado a todos que vieram antes de mim e pavimentaram esse caminho que pude trilhar. Agradeço profundamente a todos que o trilharam a meu lado. Vocês são muitos, logo impossível nominar, mas estou certo que cada um de vocês sabe bem. Grato pelos que ainda estão vindo para sedimentar esse caminho e compartilhar, também, dessa trajetória.

Vocês são a parte **real** nesse mundo **complexo** onde se projetam todos os sonhos que um dia ousei buscar.

*/ˈneɪθən/*

*“Every road is a slippery slope,  
There is always a hand  
that you can hold on to”*

*J. Mraz, 2012*

## RESUMO

A detecção de sinais anormais de eletroencefalograma (EEG) é o primeiro passo para o auxílio na identificação de neuropatologias, tendo o potencial de reduzir consideravelmente o tempo entre a captura do sinal e o laudo médico. Uma técnica que ainda não foi explorada para esta tarefa específica, mas que tem mostrado boa capacidade na detecção de transtornos mentais, devido a sua capacidade de capturar informações espaciais e temporais, é a de análise de microestados de EEG. Neste trabalho é proposta uma metodologia para detecção dos sinais anormais que combina o uso de microestados e uma rede *Learning Vector Quantization* (LVQ) com a intenção de melhorar os protótipos dos microestados obtidos inicialmente pelo método comumente usado *k-means* modificado. Resultados experimentais em uma base de dados pública, sugerem que a análise de microestados, que utiliza as características topográficas do sinal, é promissora para a detecção de sinais anormais de EEG, independentemente de uma neuropatologia especificada *a priori*. Além disso, o uso de microestados com LVQ mostrou ser estatisticamente melhor do que o método tradicional de obtenção dos protótipos de microestados.

Palavras-chave: Microestados de EEG; Otimização LVQ; Processamento de Sinais; Classificação; EEG Anormal.



## ABSTRACT

The detection of abnormal electroencephalogram (EEG) signals is the first step to aid in the identification of neuropathologies, having the potential to considerably reduce the time between signal capture and medical report. A technique that has not yet been explored for this specific task, but has shown good capacity in the detection of mental disorders, due to its ability to capture spatial and temporal information, is the EEG microstate analysis. In this work, a methodology for detecting abnormal EEG signals is proposed, combining the use of microstate analysis and a Learning Vector Quantization (LVQ) network with the intention of improving the prototypes of the microstates obtained initially by the commonly used modified k-means clustering method. Experimental results in a public dataset suggest that microstate analysis, which uses the topographic characteristics of the EEG signal, is promising for the detection of abnormal EEG signals, regardless of an *a priori* specified neuropathology. Furthermore, the use of microstates with LVQ proved to be statistically better than the traditional method of obtaining microstate prototypes.

Keywords: EEG microstates; LVQ optimization; Signal Processing; Classification; Abnormal EEG.



*“Mas ao menos fica da amargura do que nunca serei  
A caligrafia rápida destes versos,  
Pórtico partido para o Impossível.”  
F. Pessoa, 1928.*



## LISTA DE ILUSTRAÇÕES

Figura 1. Árvore de decisão – inspeção visual por neurologista.....	21
Figura 2. Captura do EEG - Sistema Internacional 10-20 .....	26
Figura 3. Regiões do Cérebro .....	28
Figura 4. Microestados canônicos (A, B, C e D).....	30
Figura 5. GEV vs Número de Protótipos de Microestado.....	31
Figura 6. Microestados extraídos do sinal de EEG .....	36
Figura 7. Arquitetura de uma Rede Neural Competitiva .....	40
Figura 8. Exemplo de um classificador SVM .....	45
Figura 9. Exemplo de um classificador RF .....	46
Figura 10. Exemplo de partição <i>k-fold</i> para validação cruzada .....	48
Figura 11. Distribuição de idade por sessão .....	52
Figura 12. Diagrama de fluxo de dados.....	57
Figura 13. Etapas da Metodologia e Implementação .....	59
Figura 14. t-SNE bidimensional, usando 276 sessões de teste, aplicado sobre o vetor de características de cada amostra.....	61
Figura 15. Protótipos de microestado gerados somente com <i>k-means</i> modificado (acima) e gerados por <i>k-means</i> modificado e ajustados com LVQ (abaixo) .....	62
Figura 16. Protótipos de microestado gerados apenas com a rede LVQ para $k = 8$ e 16 .....	63
Figura 17. <i>Boxplot</i> do erro de classificação para 500 sessões de treino e agrupamento com <i>k-means</i> e ajuste LVQ, para $k = 8$ microestados .....	66
Figura 18. <i>Boxplot</i> do erro de classificação para 500 sessões de treino e agrupamento com <i>k-means</i> e ajuste LVQ, para $k = 16$ microestados.....	66
Figura 19. <i>Boxplot</i> do erro de classificação para 2717 sessões de treino e agrupamento com <i>k-means</i> e LVQ, para $k = 8$ microestados .....	68

Figura 20. <i>Boxplot</i> do erro de classificação para 2717 sessões de treino e agrupamento com <i>k-means</i> e LVQ, para $k = 16$ microestados .....	68
Figura 21. <i>Boxplot</i> do erro de classificação do melhor resultado deste trabalho e erro do melhor resultado de Lopez (2017) .....	72

## LISTA DE TABELAS

Tabela 1. Levantamento das alterações dos microestados canônicos.....	33
Tabela 2. Conjunto de Pacientes TUAB.....	51
Tabela 3. Conjunto de Sessões TUAB .....	51
Tabela 4. Distribuição de gênero por sessão (TUAB) .....	52
Tabela 5. Comparativo do classificador SVM com e sem SFS.....	61
Tabela 6. Comparação da taxa de erro para diferentes tipos de agrupamento com 500 sessões de treino, 8 e 16 microestados .....	65
Tabela 7. Comparação da taxa de erro para 2717 sessões de treino, 8 e 16 microestados, agrupamento com <i>k-means</i> e ajuste LVQ .....	68
Tabela 8. Comparativo de resultados para 162 sessões de treino ( <i>short set</i> ).....	69
Tabela 9. Sumário de taxa de erro de classificação para todos os dados da base TUAB .....	71



## LISTA DE ABREVIATURAS E SIGLAS

A, B, C, D	Microestados Canônicos
AVC (esq, dir)	Acidente Vascular Cerebral (esquerdo, direito)
BOLD	<i>Blood-Oxygen-Level-Dependant</i>
CNN	<i>Convolutional Neural Network</i>
Cov	Covariância
CT	<i>Computerized Tomography</i>
DT	<i>Decision Tree</i>
Dur	Duração
EEG	Eletroencefalograma
FMRI	<i>functional Magnetic Resonance Imaging</i>
GEV	<i>Global Explained Variance</i>
GFP	<i>Global Field Power</i>
GMD	<i>Global Map Dissimilarity</i>
GMM	<i>Gaussian Mixture Models</i>
HMM	<i>Hidden Markov Models</i>
Iz	<i>Inion</i>
KNN	<i>k Nearest Neighbors</i>
LVQ	<i>Learning Vector Quantization</i>
MFCC	<i>Mel Frequency Cepstral Coefficients</i>
MLP	<i>Multilayer Perceptron</i>
Nz	<i>Nasion</i>
Occ	Ocorrência
PCA	<i>Principal Component Analysis</i>
PDF	<i>Probability Density Function</i>
PDR	<i>Posterior Dominant Rhythm</i>

RF	<i>Random Forest</i>
RSN	<i>Rest State Networks</i>
SdA	<i>Stacked denoising Autoencoder</i>
SFS	<i>Sequential Forward Selection</i>
SNR	<i>Signal-to-noise Ratio</i>
SVM	<i>Support Vector Machine</i>
TAAHC	<i>Topographic Atomize and Agglomerate Hierarchical Clustering</i>
TCP/AR	<i>Temporal Central Parassagital / Average Referenced</i>
TDAH	Transtorno de Déficit de Atenção e Hiperatividade
Topo	Topografia
t-SNE	<i>t-distributed Stochastic Neighbor Embedding</i>
TUAB	<i>The TUH Abnormal EEG Corpus</i>
TUEG	<i>The TUH EEG Corpus</i>
TUH	<i>Temple University Hospital</i>

## LISTA DE SÍMBOLOS

$\mathbf{a}_k$	Vetor contendo o $k$ -ésimo protótipo de microestado
$\mathbf{a}_{ln}$	Vetor do microestado atribuído ao $n$ -ésimo ponto do sinal de EEG
$\alpha(t)$	Taxa de aprendizagem
$C$	Quantidade de canais (eletrodos)
$Corr(\mathbf{x}_n, \mathbf{a}_{ln})^2$	Quadrado da correlação entre $\mathbf{x}_n$ e $\mathbf{a}_{ln}$
$dist(\mathbf{u}, \mathbf{w}_i)$	Distância entre os vetores $\mathbf{u}$ e $\mathbf{w}_i$
$H_0$	Hipótese nula em um teste estatístico
$H_1$	Hipótese alternativa em um teste estatístico
$i$	$i$ -ésimo neurônio da rede LVQ
$k$	Quantidade de <i>clusters</i> ou protótipos de microestado
$K(\mathbf{x}, \mathbf{y})$	<i>Kernel</i> de transformação de um classificador SVM
$l_n$	<i>Label</i> (etiqueta) assinalado ao $n$ -ésimo ponto do sinal de EEG
$\mu_E$	Média da taxa de erro de classificação
$N_k$	Quantidade de pontos do sinal de EEG etiquetados com o protótipo $k$
$N$	Quantidade total de pontos do sinal de EEG
$n$	$n$ -ésimo ponto do sinal de EEG
$N_A$	Número de sessões classificadas corretamente
$N_T$	Número total de sessões classificadas
$N_t$	Número de árvores em um classificador RF
$O(n)$	Ordem da complexidade computacional
$p(a)$	Função de probabilidade de $a$
$\sigma_E$	Desvio padrão da taxa de erro de classificação
$\mathbf{u}$	Vetor de entrada da rede LVQ
$\mathbf{v}$	Vetor de características ( <i>features</i> ) extraídas de uma dada amostra
$V_i[n]$	Potencial do $i$ -ésimo canal (eletrodo) no instante $n$

$V_{méd}[n]$	Potencial médio de todos os canais (eletrodos) no instante $n$
$\mathbf{w}_i$	Centroide do $i$ -ésimo <i>cluster</i> em uma rede LVQ
$\mathbf{x}_n$	Vetor contendo o $n$ -ésimo ponto do sinal de EEG
$y$	Classe atribuída a uma dada sessão de EEG

# SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	<b>19</b>
1.1. Contextualização .....	19
1.2. Análise visual do sinal de EEG .....	20
1.3. Principal Trabalho Relacionado .....	21
1.4. Objetivos .....	23
1.4.1. Objetivo Geral .....	23
1.4.2. Objetivos Específicos .....	24
1.5. Hipóteses .....	24
1.6. Estrutura do Trabalho .....	24
<b>2. MATERIAL E MÉTODOS</b> .....	<b>26</b>
2.1. Análise da Informação contida no Sinal de EEG .....	26
2.1.1. Captura do Sinal .....	26
2.1.2. Características Temporais e Espectrais do Sinal de EEG .....	27
2.1.3. Características Topológicas do sinal de EEG .....	29
2.1.4. Aspectos normais e anormais dos microestados em sinais de EEG .....	32
2.2. Microestados de EEG .....	34
2.2.1. Backfit e Temporal Smoothing .....	36
2.2.2. Características dos Microestados .....	38
2.3. Learning Vector Quantization .....	39
2.3.1. Implementação da Rede LVQ .....	40
2.3.2. Treinamento LVQ1 .....	41
2.3.3. Early stopping .....	43
2.4. Seleção Sequencial de Características e Visualização .....	43
2.5. Classificadores .....	44
2.5.1. Support Vector Machine .....	44

2.5.2. Random Forest.....	45
2.5.3. k Nearest Neighbors .....	46
2.5.4. Otimização Bayesiana .....	47
2.5.5. Holdout e cross validation.....	47
2.5.6. Métrica dos classificadores .....	49
2.5.7. Teste de Hipóteses.....	49
2.6. Base de Dados.....	50
2.6.1. Base TUAB versão 2.0.0 .....	51
2.6.2. Pré-processamento.....	53
<b>3. METODOLOGIA E IMPLEMENTAÇÃO .....</b>	<b>55</b>
3.1. Seleção de dados e pré-processamento .....	55
3.2. Obtenção dos Microestados e pós-processamento .....	55
3.3. Extração e Seleção de Características, Treino e Classificação .....	56
<b>4. RESULTADOS E DISCUSSÃO .....</b>	<b>60</b>
4.1. Resultados da Seleção de Características e Visualização das Classes .....	60
4.2. Análise dos Algoritmos de Agrupamento.....	62
4.3. Comparativo dos Algoritmos de Agrupamento e dos Classificadores .....	64
4.4. Comparativo de classificação para 8 e 16 microestados.....	67
4.5. Comparativo com os resultados da literatura .....	69
<b>5. CONCLUSÃO E TRABALHOS FUTUROS .....</b>	<b>74</b>
<b>REFERÊNCIAS BIBLIOGRÁFICA .....</b>	<b>78</b>
<b>APÊNDICE A.....</b>	<b>83</b>
A.1 Código .....	83

# 1. INTRODUÇÃO

## 1.1. Contextualização

As deficiências neurofisiológicas podem ser precedidas do aparecimento de sintomas clínicos. Isso ocorre em diversas doenças neurológicas, conforme reportado por Khanna, Pascual-Leone e Farzan (2014). O monitoramento de biomarcadores neurofisiológicos pode permitir a detecção precoce de distúrbios e contribuir para o entendimento das deficiências subjacentes.

Uma das principais técnicas para estudo dos padrões neurofisiológicos é o Eletroencefalograma (EEG) (HU; ZHANG, 2019). A técnica consiste do monitoramento da atividade cerebral através da aquisição dos potenciais em uma série de eletrodos distribuídos pelo couro cabeludo de um paciente. O EEG é uma ferramenta poderosa e muito difundida, utilizada no diagnóstico clínico de distúrbios do sono, epilepsia, doença de Alzheimer, coma, morte cerebral e diversas outras condições patológicas (OBEID; PICONE, 2018).

A técnica do EEG possui algumas vantagens clínicas no estudo da função cerebral que potencializam sua utilização: não é invasiva, possui um custo relativo baixo, e possui uma alta resolução temporal.

Devido a suas vantagens, em particular a alta resolução temporal, é possível levantar informações importantes de coordenação neural local e de longo alcance (KHANNA *et al.*, 2015).

Apesar de todos os seus benefícios, a interpretação dos sinais de EEG é um processo dependente da análise subjetiva do examinador e demanda profissionais cujo treinamento leva anos (LOPEZ *et al.*, 2016). Isso limita sua utilização no acompanhamento de pacientes e de biomarcadores associados à detecção precoce de funcionamento anormal da rede neural.

---

Nesse sentido, existem trabalhos que buscam técnicas capazes de automatizar esse processo, visando auxiliar o diagnóstico médico de neuropatologias e o acompanhamento clínico de pacientes.

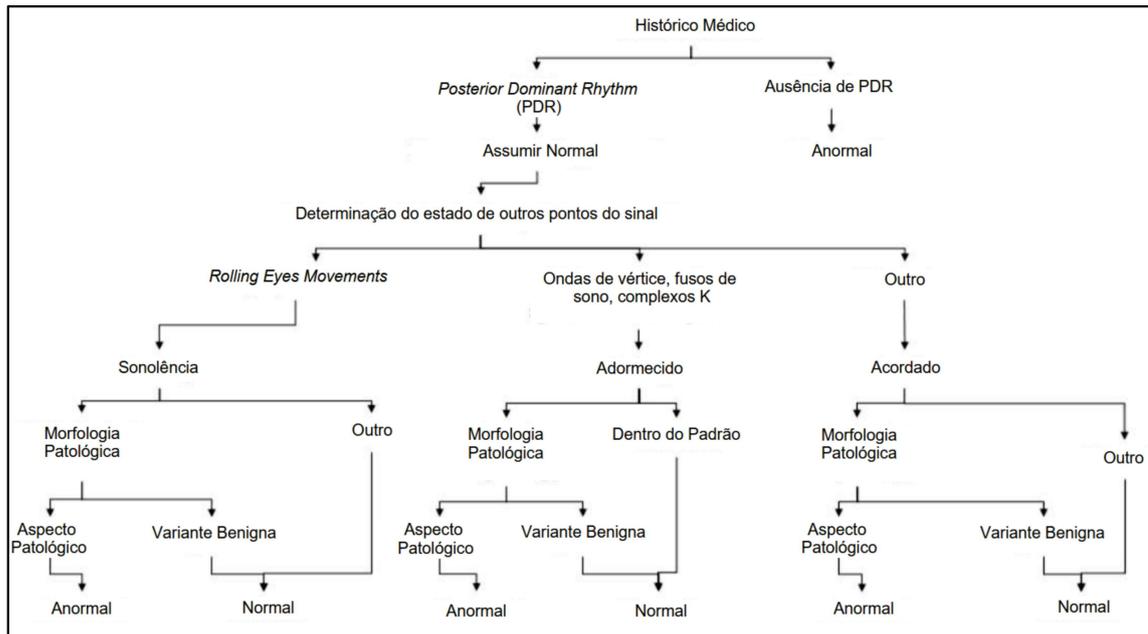
Programas de atenção à saúde à distância (telemedicina) podem ser beneficiados por sistemas capazes de detectar precocemente alterações importantes nos padrões neurofisiológicos de pacientes, permitindo intervenções mais assertivas e a tempo de impedir ou tratar a evolução dos quadros clínicos de neuropatologias. Outro fator muito importante, do ponto de vista clínico hospitalar, consiste do fato de que é possível montar uma estrutura de monitoramento contínuo do paciente em ambiente de internação hospitalar utilizando EEG, o que seria impossível com outras técnicas como fMRI (Imagem por ressonância magnética funcional, do inglês *Functional Magnetic Resonance Imaging*) ou CT (Tomografia Computadorizada, do inglês *Computerized Tomography*) (FINNIGAN; PUTTEN, 2013).

Para fins de ilustração da importância desse tipo de auxílio diagnóstico, pode ser usado como exemplo o AVC (Acidente Vascular Cerebral), que é a 2<sup>a</sup> principal causa de morte no mundo (MINISTÉRIO DA SAÚDE, 2018). Uma aplicação de monitoramento hospitalar capaz de identificar um quadro de AVC Isquêmico agudo possibilitaria um tratamento trombolítico imediato (FINNIGAN; PUTTEN, 2013), que se aplicado em até 4 horas após o início do quadro, pode impedir a evolução para óbito, além de reduzir drasticamente ou até mesmo eliminar a chance de sequelas.

## 1.2. Análise visual do sinal de EEG

Em um ambiente clínico, um neurologista treinado segue algumas etapas para classificar um sinal de EEG, que podem ser ilustrados num esquema de árvore de decisão. Para fins de ilustração do processo, na Figura 1 é mostrada uma árvore de decisão típica para classificação do sinal de EEG, através de inspeção visual por um neurologista especialista.

Figura 1. Árvore de decisão – inspeção visual por neurologista



Adaptado de Lopez (2017).

Pode-se observar pela Figura 1 que essa classificação leva em consideração uma série de fatores importantes, como histórico médico, condições clínicas e fisiológicas do paciente no momento da captura do sinal e verificações de padrões contidos no sinal. Esse tipo de análise visual é altamente sensível à subjetividade do examinador e, mais importante, demanda treinamento altamente especializado do profissional, um processo que leva anos (LOPEZ, 2017).

Isto realça a importância do uso de métodos computacionais eficazes para auxiliar os profissionais na decisão, com a expectativa de tornar o processo de avaliação dos sinais de EEG mais rápido e menos subjetivo.

### 1.3. Principal Trabalho Relacionado

O trabalho de Lopez (2017) foi o primeiro estudo identificado que trata do tema de detecção de sinais anormais de EEG usando uma base de dados ampla constituída por diversas neuropatologias. Nele, são utilizadas técnicas de aprendizado de máquina e redes profundas, entretanto, elas são aplicadas utilizando a análise clássica do sinal de EEG,

---

baseada nas características temporais e espectrais, que são detalhadas na Seção 2.1.2. A autora estabelece uma linha de base (do inglês *baseline*) da classificação automática de sinais anormais de EEG, alcançando uma taxa de erro na classificação de 21,2% (melhor resultado) quando utilizando uma rede profunda.

No trabalho de Lopez (2017), foi inicialmente estabelecido um resultado base com classificadores do tipo kNN (*k Nearest Neighbors*) e RF (*Random Forest*) em um subconjunto da base de dados denominado *short dataset*, utilizando apenas 162 sessões de EEG.

Todos os classificadores foram treinados usando as características extraídas de diferentes regiões do crânio, por pares de eletrodos, utilizando vários períodos com duração entre 3 e 9 segundos, dependendo do teste e do classificador, sendo processados vários trechos sobrepostos para uma mesma amostra.

As características utilizadas pela autora foram: *Mel Frequency Cepstral Coefficients (MFCC)*, Energia, Energia diferencial e 1ª e 2ª derivadas dessas características. A dimensão do vetor de características é reduzida por meio de PCA (*Principal Component Analysis*). O trabalho avalia uma série de classificadores intermediários no *short dataset*, devido à complexidade computacional e, por fim, é desenvolvido o sistema CNN-MLP (*Convolutional Neural Networks - Multilayer Perceptron*), que obteve o melhor resultado para a base de dados completa, denominada no trabalho como *full dataset*.

O CNN-MLP classifica os sinais através de uma rede com 3 camadas convolucionais que é treinada com o vetor de características de 4 diferentes regiões do escalpo e os dados de saída são agregados através da camada MLP, que é responsável por realizar a classificação das amostras.

O trabalho de Lopez (2017) será usado como referência para comparação de resultados do método proposto neste trabalho.

---

## 1.4. Objetivos

### 1.4.1. Objetivo Geral

O presente trabalho tem por objetivo a detecção de patologias em sinais de EEG por meio do uso de técnicas de aprendizado de máquina e de biomarcadores associados a neuropatologias, utilizando para isto análise de microestados de EEG.

Para validação da metodologia proposta, será utilizada uma base de dados pública formada por sinais de EEG de pacientes saudáveis e acometidos de patologias neurológicas.

O objetivo geral do presente trabalho é verificar a aplicabilidade da análise de microestados (LEHMANN; OZAKI; PAL, 1987), que usam a configuração topográfica do sinal e suas propriedades dinâmicas, para determinar se um sinal de EEG capturado em estado de repouso pode ou não ser classificado como normal ou anormal.

Há uma série de métodos tradicionalmente utilizados para avaliação de sinais de EEG, em particular características temporais e espectrais. Tais técnicas, entretanto, estão altamente ligadas a aspectos subjetivos de avaliação e características muito dependentes de cada indivíduo avaliado (GEMEIN *et al.*, 2020).

Dessa forma, a investigação da técnica de análise de microestados (LEHMANN; OZAKI; PAL, 1987), que tem demonstrado capacidade de discriminar aspectos intrínsecos do sinal de EEG em uma série de estudos (KHANNA *et al.*, 2015), explicando sua variância de maneira menos subjetiva, é promissora na investigação de sistemas de auxílio diagnóstico na análise e classificação do sinal de EEG. Esta investigação também é justificada por não terem sido encontrados na literatura trabalhos que utilizam microestados para uma classificação do sinal de EEG entre normal e anormal, em particular em uma base de dados ampla com diversas patologias envolvidas.

O presente trabalho busca, ainda, uma metodologia supervisionada para obtenção dos protótipos de microestados, por meio do uso de Redes Neurais Artificiais LVQ (*Learning*

---

*Vector Quantization*) (KOHONEN, 1995). Isto difere dos métodos tradicionalmente utilizados na literatura para análise de microestados, que costumam utilizar técnicas não-supervisionadas (Poulsen *et al.*, 2018).

#### 1.4.2. Objetivos Específicos

- Avaliar a eficácia do uso de microestados para detectar um sinal anormal de EEG;
- Identificar quais são as características e estatísticas de microestados que melhor discriminam sinais normais de anormais;
- Verificar o efeito de usar a rede do tipo LVQ para a obtenção dos protótipos de microestados;
- Comparação dos resultados com os obtidos por Lopez (2017).

#### 1.5. Hipóteses

Com base na revisão da literatura, os objetivos neste trabalho foram elaborados para avaliar as seguintes hipóteses:

- i) Microestados podem ser utilizados para discriminar entre sinais de EEG normais e anormais;
- ii) É possível aprimorar os protótipos de microestados por meio do uso de um método de agrupamento supervisionado combinado com o resultado de um método não-supervisionado; e
- iii) É possível obter os protótipos de microestados por meio de agrupamento supervisionado, com discriminação dos sinais e taxas de erros de classificação análogas às obtidas com o agrupamento não-supervisionado.

#### 1.6. Estrutura do Trabalho

O restante da estrutura deste trabalho é a seguinte: no Capítulo 2 são apresentados os métodos utilizados e a base de dados. No Capítulo 3 é descrita a metodologia proposta e implementada neste trabalho. No Capítulo 4, os resultados que foram obtidos são

---

apresentados. Por último, no Capítulo 5 são apresentadas as conclusões e as propostas de trabalhos futuros.

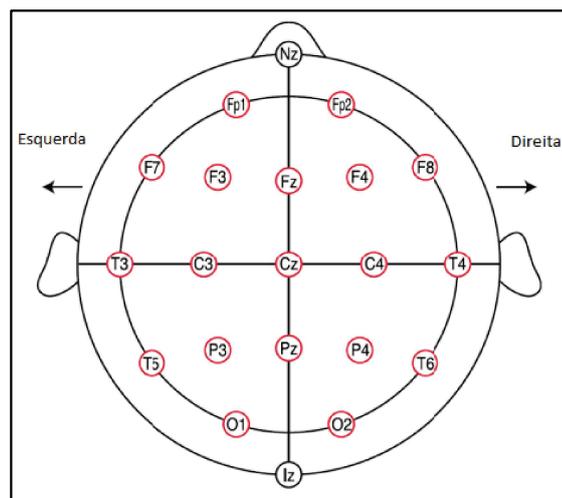
## 2. MATERIAL E MÉTODOS

### 2.1. Análise da Informação contida no Sinal de EEG

#### 2.1.1. Captura do Sinal

A captura do sinal de EEG é feita através do posicionamento de eletrodos no escalpo, utilizando-se, em geral, o sistema de montagem padrão do sistema internacional 10-20, que posiciona os eletrodos no crânio a uma distância entre si de 10% e 20% da medida total entre pontos de referência no escalpo, a saber as distâncias *nasion-inion* (Nz-Iz) e direita-esquerda (FERRELL *et al.*, 2020). A Figura 2 exemplifica o sistema de montagem.

Figura 2. Captura do EEG - Sistema Internacional 10-20



Adaptado de Okamoto *et al.* (2004)

No ambiente clínico, o fluxo típico de captura e análise do EEG envolve (LOPEZ, 2017): 1) o posicionamento dos eletrodos no crânio do paciente, em geral por técnicos; 2) a aquisição do sinal cerebral através de *software* específico; 3) interpretação do sinal por um neurologista certificado; e 4) laudo neurológico contendo histórico, medicações, descrição da gravação de EEG e alterações encontradas, bem como uma impressão geral do EEG, classificando-o como normal ou anormal.

---

### 2.1.2. Características Temporais e Espectrais do Sinal de EEG

Há uma série de técnicas para análise da informação contida no sinal de EEG, sendo as mais difundidas na literatura aquelas relacionadas à análise das características **temporais** e **espectrais** do sinal (NIAZY *et al.*, 2011; CUSTO *et al.*, 2017; MONDINI; MANGIA; CAPPELLO, 2018; SHAH *et al.*, 2018; MCMACKIN *et al.*, 2019). Em geral, elas caracterizam a morfologia temporal das formas de onda e distribuições de frequência em certos eletrodos pré-selecionados (POULSEN *et al.*, 2018).

Nessa análise clássica, um sinal de EEG é considerado normal quando as frequências, tensões, formas de onda e outras características estão contidas em certos intervalos pré-determinados e em certas regiões do escalpo, em geral observadas a partir de pares ou grupos de eletrodos, definindo-se assim a atividade compatível com o funcionamento cerebral normal.

Considerando-se a análise do ponto de vista espectral, a atividade cerebral é comumente dividida em bandas de frequência (EBERSOLE, 2014; HU; ZHANG, 2019), sendo elas: delta (0 – 4 Hz), teta (4 – 8 Hz), alfa (8 – 13 Hz), beta (13 – 25 Hz), gama (> 25 Hz).

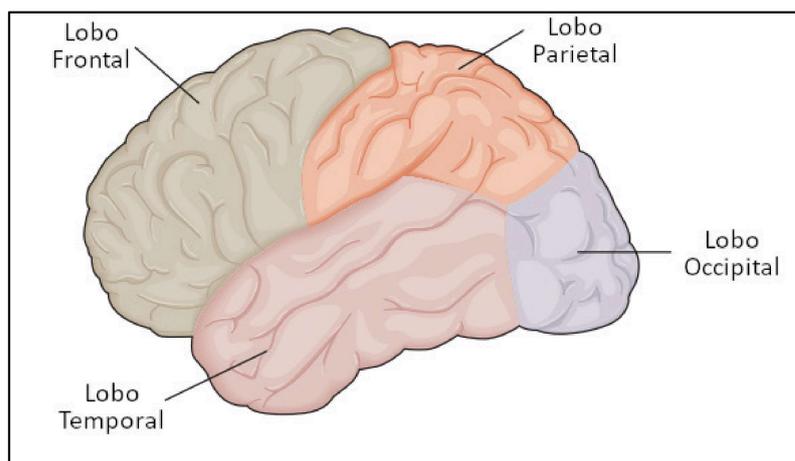
Sob essa perspectiva clássica, pode-se caracterizar o EEG normal de um adulto, entre outras, de acordo com as seguintes características (EBERSOLE, 2014; LOPEZ, 2017):

1. **Reatividade:** se refere à resposta a certas variações fisiológicas (abertura de olhos, estímulo sensorial etc.);
2. **Banda alfa:** ponto de partida para a análise visual do EEG, e posterior classificação como normal/anormal. As ondas na banda alfa se originam no lobo occipital (Figura 3) e são caracterizadas por frequências entre 8-13 Hz e amplitude entre 15-45  $\mu$ V.
3. **Banda Mi ( $\mu$ ):** essa sub-banda da banda alfa é denominada “frequência motora” e ocorre na região central com frequências entre 8-10 Hz e amplitudes análogas à banda alfa. Essa banda é suprimida unilateralmente pelo movimento de um

membro do hemisfério oposto. Também é suprimida por condições como fadiga, estímulos somatossensoriais e motores. Não é sempre detectável.

4. **Bandas Beta e Gama:** A atividade contida nessas bandas, para o EEG adulto normal, consiste de ondas que ocorrem nos intervalos de 14 a 16 Hz e de 18 a 25 Hz, para a banda beta e no intervalo de 35 a 40 Hz, para a banda gama, em geral com amplitudes entre 5-20 mV. É rara a presença de atividade na banda acima de 25 Hz no caso de EEG de escalpo.
5. **Banda Teta:** o EEG adulto normal tende a apresentar traços com amplitude menor que 15  $\mu\text{V}$  na banda de frequência entre 6-7 Hz, nas regiões frontal (Figura 3), frontocentral e, ocasionalmente, na região ao longo da linha central. Esse tipo de onda geralmente se torna sustentado e com amplitudes aumentadas no caso de sonolência.
6. **Banda Delta:** a presença de ondas na banda delta no EEG adulto pode indicar anormalidades (EBERSOLE, 2014). Esse tipo de onda é considerado normal em crianças e pode também ocorrer, eventualmente, em adultos com mais de 60 anos.

Figura 3. Regiões do Cérebro



Adaptado de Kandel (2013).

Conforme Obeid e Picone (2018), as patologias neurológicas causam desvios nos padrões normais do sinal de EEG e podem ocorrer a qualquer instante ou mesmo de forma

---

intermitente, sendo muito difíceis de serem capturados em exames de rotina, geralmente necessitando de monitoramento a longo prazo.

Segundo Obeid e Picone (2018), historicamente foram utilizadas três abordagens para o problema de classificação automática do sinal de EEG: (i) classificação única do sinal todo, (ii) classificação de segmentos do sinal, e (iii) decodificação sequencial. Todas essas abordagens consideram apenas as características temporais e espectrais do sinal e a presença de eventos anormais, bem como a correlação entre estes. Os autores afirmam ainda que a maioria dos estudos disponíveis tiveram foco em um número muito pequeno de pacientes.

Nesse tipo de abordagem para classificação do sinal de EEG, a primeira etapa consiste em verificar algum desvio do padrão normal a partir de características temporais e espectrais. Em seguida, os trechos de interesse são processados para detecção ou não de padrões anormais e os eventos associados são anotados no sinal. Alguns métodos consideram a correlação entre os eventos. Somente após todo esse processamento, é feita uma decisão se o sinal pode ser classificado como normal dentro do conjunto de características pré-definidas para o cérebro humano saudável, passando-se a fazer uma identificação de patologias específicas, associadas às alterações verificadas no sinal. Esse processo, em geral, tenta incorporar nos modelos de aprendizado de máquina os conhecimentos dos profissionais, como aqueles ilustrados na árvore de decisão para análise visual do sinal de EEG (LOPEZ, 2017), ilustrado na Figura 1.

### **2.1.3. Características Topológicas do sinal de EEG**

Sob outra perspectiva, muitas pesquisas têm trabalhado com outras características do sinal de EEG além daquelas clássicas. Trabalhos recentes (MICHEL; KOENIG, 2018) têm apontado para uma mudança de paradigma no entendimento do funcionamento do cérebro humano, verificando que, em repouso (na ausência da execução de alguma tarefa específica), o cérebro opera alternando entre certos mapas topológicos, que podem ser

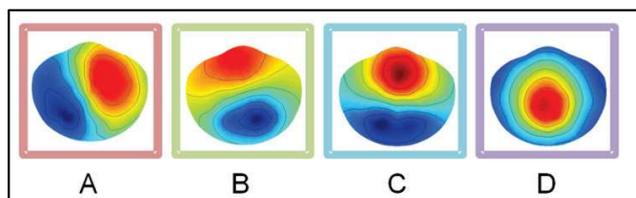
definidos como a distribuição do potencial elétrico no escalpo, e que estão envolvidos na atividade funcional. Esses mapas topológicos são gerados por redes de conectividade cerebral e são apontadas como fundamentais nos processos sensoriais e cognitivos, sendo denominadas *Rest State Networks* (RSN) (NIAZY *et al.*, 2011; CUSTO *et al.*, 2017).

Segundo estudos de conectividade cerebral utilizando fMRI (Imagem por Ressonância Magnética funcional), esses mapas surgem devido às funções do cérebro em repouso serem caracterizadas por flutuações coerentes no sinal BOLD (*blood-oxygen-level-dependant*) (MUSSO *et al.*, 2010; NIAZY *et al.*, 2011; CUSTO *et al.*, 2017). Essas flutuações são distintas das contribuições de sinal advindas da respiração e do sistema cardiovascular e operam como um estado exploratório contínuo do cérebro (HEINE *et al.*, 2012).

A ferramenta denominada análise de microestados (LEHMANN; OZAKI; PAL, 1987; KOENIG *et al.*, 2002), tem sido aplicada para o entendimento desse comportamento. Através da premissa de que as diferentes topografias cerebrais são geradas por diferentes configurações de fontes neurais, estudos demonstram que os microestados são capazes de refletir os estados funcionais globais do cérebro e estão intrinsecamente correlacionados com as RSNs de fMRI (POULSEN *et al.*, 2018).

Uma série de estudos sobre diferentes patologias ao longo dos últimos dez anos identificou um grupo de 4 microestados, que estão ilustrados na Figura 4, e são denominados na literatura de **microestados canônicos**. Os microestados canônicos são normalmente encontrados no sinal de EEG do cérebro humano saudável e são capazes de representar em média 65% da variância global do sinal de EEG (MICHEL; KOENIG, 2018).

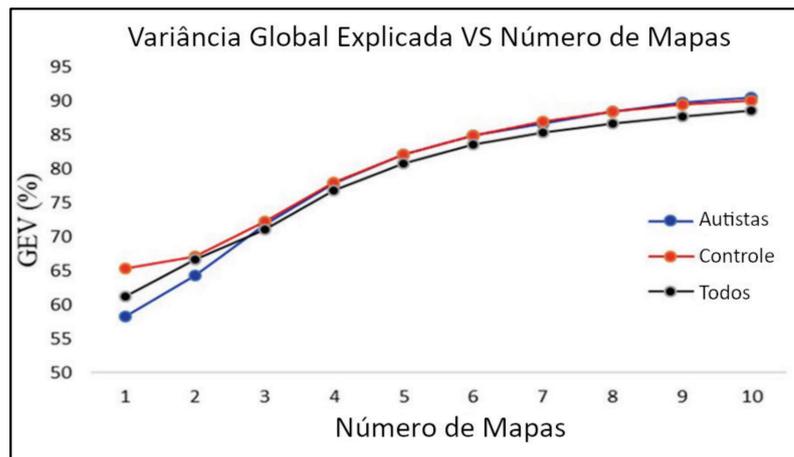
Figura 4. Microestados canônicos (A, B, C e D)



Adaptado de Croce *et al.* (2020).

Apesar dos microestados canônicos serem capazes de representar boa parte da variância global do sinal, dependendo da aplicação, o número de microestados deve ser selecionado de acordo com a base de dados (MICHEL; KOENIG, 2018). Alguns exemplos são o estudo de Custo *et al.* (2017), que obteve uma explicação da variância global do sinal de aproximadamente 60% com os 4 microestados A-D, mas conseguiu obter 84,8% com 7 protótipos de microestados. Já no trabalho de Seitzman *et al.* (2018), 15 protótipos de microestados foram necessários para representar 80% da variância global do sinal. Na Figura 5, extraída do trabalho de D’Croz-Baron *et al.* (2019), está ilustrada uma análise do comportamento da Variância Global Explicada (GEV) do sinal de EEG, para  $k \in [1,10]$  protótipos de microestados, extraídos de uma base de dados composta por pacientes controle e pacientes com autismo. No estudo é comparado o comportamento dos protótipos extraídos só do grupo controle, só do grupo com autismo e de ambos os grupos. Observa-se da figura que o GEV tende a aumentar quanto maior for o número de protótipos, contudo, a taxa de aumento do GEV tende a cair à medida que a quantidade de protótipos (número de mapas) aumenta.

Figura 5. GEV vs Número de Protótipos de Microestado



Adaptado de D’Croz-Baron *et al.* (2019).

---

#### 2.1.4. Aspectos normais e anormais dos microestados em sinais de EEG

Nos estudos levantados por Michel e Koenig (2018), os microestados do sinal normal apresentam topografias e propriedades dinâmicas consistentes, desviando desse padrão de maneiras diferentes para patologias distintas. Esses resultados sugerem que os microestados e suas propriedades dinâmicas seriam capazes de discriminar um sinal normal ou anormal de EEG de uma maneira diferente da análise clássica.

Existem alguns trabalhos que buscam fazer uma identificação de doenças específicas com uso de microestados. Por exemplo, o trabalho de M. Alves *et al.* (2020) utiliza análise de microestados e teoria de grafos para identificar pacientes com TDAH (Transtorno de Déficit de Atenção e Hiperatividade). Por outro lado, não foram identificados trabalhos que utilizem a análise de microestados para uma classificação inicial do sinal de EEG entre normal e anormal, em particular em uma base de dados ampla, como pretendido neste trabalho, a fim de possibilitar a automação da detecção de sinais anormais de EEG auxiliando o diagnóstico e laudo médico em ambiente clínico.

Os autores Michel e Koenig (2018) apresentam um sumário das alterações de microestados já detectados em pesquisas acadêmicas, que está detalhado na Tabela 1. Dessa maneira, analogamente ao conceito de EEG adulto normal, do ponto de vista das características espectrais e temporais, pode-se verificar, do ponto de vista das características topológicas do sinal de EEG, variações observadas nos quatro microestados canônicos nas diversas patologias em relação ao referencial de pacientes saudáveis. A hipótese principal levantada nesse caso é a de que isso indica que existe um grau de separabilidade entre pacientes saudáveis (EEG normal) e pacientes com algum tipo de patologia neurológica (EEG anormal).

Sob esse ponto de vista, pode-se supor que, para maiores quantidades de microestados, o comportamento das topologias e variações nas características dos microestados ocorrerá de maneira análoga, de forma que seria possível discriminar um padrão normal e caracterizar os desvios desse padrão como anormais no sinal EEG.

Tabela 1. Levantamento das alterações dos microestados canônicos

Estudo	Ano	Doença	Medicação	Pacientes	Microestados			
					A	B	C	D
König	1999	Esquizofrenia	Não	9				↑Dur, Topo
Strelets	2003	Esquizofrenia	Não	14		↓Dur, Topo		
Lehmann	2004	Esquizofrenia	Não	27	↑Occ, Cov	↓Dur	↑Occ	↓Dur
Kikuchi	2007	Esquizofrenia	Não	21		↑Occ	↑Occ	↓Dur, Cov
Nishida	2013	Esquizofrenia	Não	21	↓Dur ↑Occ	↓Dur	↑Occ	↓Dur
Andreaou	2013	Esquizofrenia	Sim	18		↑Cov		
Tomescu	2015	Esquizofrenia	Sim	27			↑Dur, Occ, Cov, GEV	↓Occ, Cov, GEV
Irisawa	2006	Esquizofrenia	Não	24		↓Dur		
Kindler	2011	Alucinação	Sim	9				↓Dur
Andreaou	2013	Alto risco de Esquizofrenia	Sim	18	↑Occ, Cov			
Tomescu	2014	22q11DS	Sim	30			↑Dur, Occ, Cov, GEV	↓Dur, Occ, Cov, GEV
Nishida	2013	Demência	Não	18			↓Dur	
Kikuchi	2011	Transtorno do Pânico	Não	18	↑Occ, Cov		↓Occ	
Gschwind	2016	Esclerose múltipla	Não	53	↑Dur	↑Dur		
Corradini	2014	Traumatismo Cranioencefálico	Não	26	↓Dur		↓Dur	↓Dur
Drissi	2016	Narcolepsia	Não	16	↓Dur, Cov, GEV	↑Dur, Cov, GEV	↑Dur, Cov, GEV	
Grieder	2016	Demência Sem.	Sim	8		Topo	Topo	
Gao	2016	Diplegia Espástica	Não	15	↑Occ	↓Dur	↑Occ	
Zappasodi	2017	Acidente Vascular Cerebral	Não	47			↓Occ,Topo (AVC esq)	↓Cov,Topo (AVC dir)

Resumo dos estudos publicados sobre alterações de microestados de EEG em doenças neuropsiquiátricas usando a abordagem convencional de agrupamento *k-means* e restringindo a análise aos quatro **mapas canônicos (A, B, C e D)** de microestados.

**Dur** = Duração, **Occ** = Ocorrência, **Cov** = Cobertura, **GEV** = Variação Global Explicada, **Topo** = Topografia.

Adaptado de Michel e Koenig (2018).

As características dos microestados referenciadas na Tabela 1, a saber Duração, Ocorrência, Cobertura, GEV e Topografia estão detalhadas na Seção 2.2.2.

---

## 2.2. Microestados de EEG

Um Microestado de EEG pode ser definido como um mapa topográfico de potencial cerebral, consistindo de configurações estacionárias por determinados períodos de tempo, em geral da ordem de centenas de milissegundos (ZAPPASODI *et al.*, 2017). Uma definição mais direta é: a orientação do dipolo elétrico do mapa topográfico permanece inalterada nestes intervalos (LEHMANN; OZAKI; PAL, 1987). Na literatura médica, os autores geralmente se referem a essas configurações como “quase estáveis”.

Algumas definições são relevantes para maior clareza de alguns termos utilizados ao longo deste texto. O termo *mapa topográfico* é aqui definido como as curvas de nível da amplitude do potencial elétrico distribuídas ao longo de todo o escalpo num dado instante de tempo, representadas por uma escala de cores. O aspecto de alguns desses mapas topográficos pode ser observado na Figura 4 e na Figura 6. Os termos *microestados* e *protótipos de microestado* (ou somente *protótipos*) podem ser definidos, respectivamente, como o grupo, contendo  $k$  mapas topográficos, escolhido para representação do sinal de EEG após a etapa de agrupamento; e qualquer grupo de mapas topográficos candidato, que poderá ser escolhido como os microestados principais.

Para a obtenção dos microestados de EEG, a técnica utilizada consiste, em suma, em localizar o grupo de mapas topográficos de potencial cerebral no sinal de EEG, filtrado na frequência do denominado canal alfa, correspondentes aos pontos de máximo na curva de *GFP* (*Global Field Power*), que melhor representam a variância global contida no sinal original.

Considere um sinal de EEG composto de um total de  $N$  pontos, capturados a uma frequência  $f_s$ , onde cada ponto do sinal se refere ao potencial elétrico de cada um dos  $C$  canais (eletrodos) posicionados sobre o couro cabeludo do paciente (por exemplo, no sistema internacional 10-20 da Figura 2), de forma que o sinal de EEG possa ser representado como uma matriz de dimensão  $C \times N$ , então o valor de  $GFP_n$ , para cada  $n = [1, 2, \dots, N]$ , é obtido através de

$$GFP_n = \sqrt{\frac{1}{C} \sum_{i=1}^C (V_i[n] - V_{méd}[n])^2} \quad , \quad (1)$$

na qual  $V_i[n]$  é o potencial do  $i$ -ésimo canal (eletrodo),  $V_{méd}[n]$  é o potencial médio de todos os  $C$  canais (eletrodos).

O  $GFP$  é uma medida singular e independente de referência, do potencial do sinal ao longo dos eletrodos. Matematicamente, o  $GFP_n$  é o desvio padrão de todos os eletrodos num dado instante de tempo  $n$  (MURRAY; BRUNET; MICHEL, 2008).

A probabilidade de encontrar um microestado bem definido é maximizada nos picos da curva  $GFP$ , porque esses picos estão associados aos pontos do sinal de EEG com a maior relação sinal ruído (SNR) (POULSEN *et al.*, 2018).

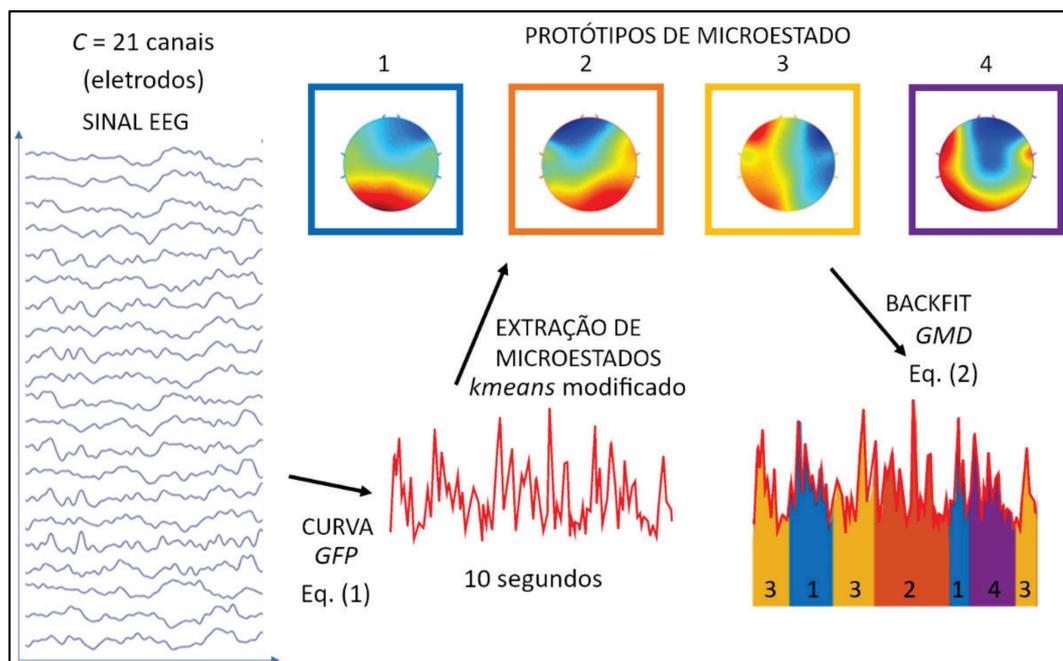
De forma a determinar os microestados, os pontos de máximo local (picos da curva  $GFP$ ), obtidos através da Equação (1), são fornecidos a um algoritmo de análise espacial para agrupamento (do inglês *clustering*). O centroide de cada grupo (ou *cluster*) representará um protótipo de microestado. Podem ser utilizadas diversas técnicas de agrupamento, sendo as principais o TAAHC (*Topographic Atomize and Agglomerate Hierarchical Clustering*) e o  $k$ -means modificado (POULSEN *et al.*, 2018).

Segundo Michel e Koenig (2018), o  $k$ -means modificado é o método mais utilizado na literatura, sendo a escolha dessa técnica adequada por permitir comparações futuras com outros estudos.

Após a obtenção dos protótipos principais de microestados por um algoritmo de agrupamento, o sinal de EEG original pode ser representado como uma sequência desses mapas topográficos através do processo de *backfit* (ver Seção 2.2.1).

Na Figura 6, pode-se verificar o aspecto de um conjunto de  $k = 4$  protótipos de microestado, calculados a partir de um trecho de 10 segundos de um sinal de EEG, que contém os 21 canais (eletrodos) do sistema internacional 10-20 (Figura 2), capturado à uma frequência de 250 Hz.

Figura 6. Microestados extraídos do sinal de EEG



Fonte: o autor.

Apesar de não existirem métodos específicos para seleção da quantidade de protótipos numa análise baseada em microestados (POULSEN *et al.*, 2018), no presente trabalho foram avaliadas as quantidades de microestados para representação do sinal de EEG entre 2 e 16, na fase preliminar do estudo (KÜSTER; CIARELLI, 2021). Os testes de classificação apresentaram resultados de particular interesse com a seleção de pelo menos 8 microestados. Os melhores resultados preliminares foram obtidos com 16 microestados. A análise da técnica se limitou a 16 microestados, pois quantidades maiores de microestados não puderam ser investigadas, devido às limitações na própria *toolbox* utilizada e ao custo computacional de quantidades maiores de protótipos (POULSEN *et al.*, 2018).

### 2.2.1. *Backfit e Temporal Smoothing*

O processo de *backfit* se refere à representação do sinal de EEG original como uma sequência de protótipos de microestados, previamente obtidos através do algoritmo de agrupamento.

Para tanto, é utilizada uma medida de distância entre cada mapa topográfico (representado na forma de um vetor contendo os potenciais instantâneos dos  $C$  eletrodos) capturado no  $n$ -ésimo intervalo de tempo ( $\mathbf{x}_n$ ) e cada um dos  $k$  protótipos de microestado obtidos no agrupamento ( $\mathbf{a}_k$ ), sendo associado a cada ponto do sinal de EEG o protótipo mais similar, através de uma etiqueta (ou *label*)  $l_n$ .

A medida de distância (ou dissimilaridade) utilizada no presente trabalho é o GMD (*Global Map Dissimilarity*), que pode ser calculada entre o mapa topográfico de  $\mathbf{x}_n$  e o protótipo de microestado  $\mathbf{a}_k$ , conforme Murray, Brunet e Michel (2008), por

$$\text{GMD}(\mathbf{x}_n, \mathbf{a}_k) = \frac{1}{\sqrt{C}} \left\| \frac{\mathbf{x}_n}{\text{GFP}_n} - \frac{\mathbf{a}_k}{\text{GFP}_k} \right\|_2, \quad (2)$$

na qual  $C$  representa o número de canais (eletrodos),  $\|\cdot\|_2$  é a norma  $L_2$ , e os valores de  $\text{GFP}_n$  e  $\text{GFP}_k$  são obtidos pela Equação (1), e normalizam cada um dos mapas topográficos, para independência de escala na medida de distância (POULSEN *et al.*, 2018).

A cada um dos  $n \in N$  pontos do sinal de EEG é assinalado o protótipo de microestado associado ao índice  $k$ , que possui a menor distância GMD com o mapa topográfico do ponto  $n$ , ou seja,

$$l_n = \text{argmin}_k \{ \text{GMD}(\mathbf{x}_n, \mathbf{a}_k) \}, \quad (3)$$

na qual  $l_n$  é a etiqueta (ou *label*) atribuída ao  $n$ -ésimo ponto do sinal de EEG. O vetor contendo o protótipo de microestado associado ao índice  $k$  ( $\mathbf{a}_k$ ) pode ser, então, referenciado também como  $\mathbf{a}_{l_n}$ .

Como o sinal de EEG possui uma série de ruídos e artefatos indesejados, que podem contribuir para o aparecimento de trechos muito curtos assinalados a um determinado protótipo de microestado, por vezes até mesmo amostras individuais, após etiquetar cada  $\mathbf{x}_n$  com o protótipo de microestado mais similar através do processo de *backfitting*, é feito um pós-processamento denominado *temporal smoothing*, que consiste em eliminar intervalos etiquetados com determinado protótipo de microestado com duração muito

---

curta, inferior à duração mínima tipicamente observada em microestados, que é da ordem de centenas de milissegundos (POULSEN *et al.*, 2018).

O processo consiste em remapear os trechos de sinal assinalados a um determinado microestado menores que um limiar consistente com a ordem de duração típica de um microestado, para o microestado adjacente que possua a menor distância GMD. O processo é repetido até que não haja trechos muito curtos no sinal.

### 2.2.2. Características dos Microestados

A partir da representação do sinal de EEG por uma série de microestados, é possível extrair algumas características e estatísticas úteis para sua discriminação. A seguir são apresentadas as características citadas e utilizadas neste trabalho (KHANNA *et al.*, 2015; POULSEN *et al.*, 2018):

- i. **Topografia:** refere-se ao aspecto visual do vetor contendo a distribuição dos potenciais elétricos dos  $C$  canais (eletrodos) num dado instante de tempo, representado como um mapa topográfico do crânio, utilizando uma escala de cores para ilustração da amplitude dos potenciais, interpolando os valores intermediários (onde não há eletrodos posicionados);
- ii. **Duração:** pode ser definida como o tempo de duração de um dado microestado no sinal, medida em [ $ms$ ];
- iii. **Ocorrência:** se refere à média da quantidade de vezes por segundo em que um determinado microestado é dominante, medida em [ $s^{-1}$ ];
- iv. **Cobertura:** é o percentual de tempo em que um determinado microestado permanece dominante no sinal todo, medida em [%];
- v. **Probabilidades de transição:** mensuram as probabilidades de um dado protótipo de microestado de transicionar para outros protótipos de microestados, medida pela matriz de transição [ $k \times k$ ], para  $k$  protótipos de microestado; e

- vi. **Variância Global Explicada (GEV)**: pode ser definida como a capacidade de um protótipo de microestado de explicar a variância contida na base de dados, medida em [%], e é calculada, para cada um dos  $k$  protótipos de microestado, por (MURRAY; BRUNET; MICHEL, 2008)

$$GEV_k = \sum_n^{N_k} GEV_n \quad , \quad (4)$$

na qual  $N_k$  é um subconjunto de  $N$  que contém os pontos assinalados ao  $k$ -ésimo protótipo de microestado ( $l_n = k$ ), e onde  $GEV_n$  é calculada para o  $n$ -ésimo ponto do sinal de EEG, e mede a similaridade desse ponto do sinal com o protótipo de microestado ao qual ele foi assinalado ( $\mathbf{a}_{l_n}$ , tal que  $l_n = k$ ), por

$$GEV_n = Corr(\mathbf{x}_n, \mathbf{a}_{l_n})^2 \cdot \frac{GFP_n^2}{\sum_{n'}^N GFP_{n'}^2} \quad , \quad (5)$$

em que  $N$  é a quantidade total de pontos contidos no sinal de EEG,  $n'$  são todos os pontos do sinal,  $\mathbf{x}_n$  é o mapa topográfico do  $n$ -ésimo ponto do sinal de EEG,  $\mathbf{a}_{l_n}$  é o vetor que contém o protótipo de microestado assinalado ao  $n$ -ésimo ponto do sinal de EEG, e  $Corr$  é a medida de correlação espacial. Ambos  $\mathbf{x}_n$  e  $\mathbf{a}_{l_n}$  são vetores de dimensão  $C$ .

A  $GEV$  representa o quadrado da correlação entre o mapa topográfico do ponto  $n$  do sinal de EEG e seu protótipo de microestado, normalizada pela fração do quadrado do  $GFP$  do ponto  $n$ , em relação ao quadrado do  $GFP$  total de todos os pontos contidos no sinal.

### 2.3. Learning Vector Quantization

De forma a melhorar os resultados da detecção de sinais anormais de EEG, é proposto utilizar neste trabalho uma rede neural LVQ (*Learning Vector Quantization*) (KOHONEN, 1995) com o objetivo de obter protótipos de microestados que discriminam melhor os sinais normais de EEG daqueles anormais, de maneira supervisionada. O LVQ

foi escolhido devido à sua capacidade de obter protótipos de maneira supervisionada, uma vez que os métodos tradicionais para obter os microestados são obtidos de forma não-supervisionada.

Uma rede LVQ é um classificador especificamente projetado para classificação estatística, capaz de definir regiões referentes a cada classe no espaço de dados de entrada (KOHONEN, 1995). O *Vector quantization* (VQ) é um método clássico de aproximar uma função de densidade de probabilidade (*probability density function* (PDF))  $p(\mathbf{a})$ , de uma variável aleatória  $\mathbf{a} \in R^n$ , usando uma quantidade finita de protótipos de *clusters* (DU, 2010).

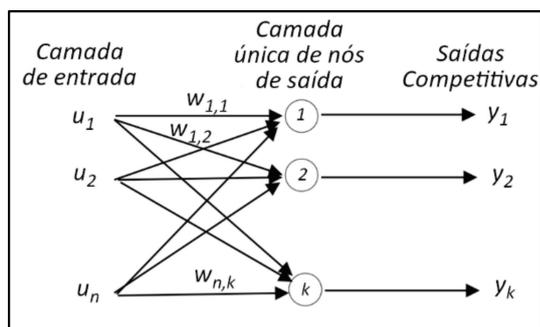
O LVQ possui neurônios em sua estrutura, na camada oculta, cujos pesos são ajustados para permitir uma melhor discriminação das diferentes classes na camada de saída da rede, cujo mecanismo de ajuste é discutido a seguir.

### 2.3.1. Implementação da Rede LVQ

Uma Rede LVQ possui uma arquitetura análoga a uma Rede Neural Competitiva, apenas com a diferença de que cada saída é associada a uma única classe (MELIN *et al.*, 2014).

Na Figura 7 tem-se a ilustração de uma arquitetura de rede neural competitiva.

Figura 7. Arquitetura de uma Rede Neural Competitiva



Adaptado de Dongoran *et al.* (2018).

Nesse tipo de arquitetura, a camada oculta é uma camada do tipo competitiva, onde cada um dos neurônios, expresso em função do vetor de pesos  $\mathbf{w}_i$ , calcula uma medida de similaridade entre um dado vetor de entrada  $\mathbf{u}$  e os pesos, em geral utilizando a

---

distância Euclidiana. Desse modo, a saída da rede será o resultado da competição, onde o neurônio vencedor terá saída 1 e todos os outros 0 (DU, 2010; MELIN *et al.*, 2014).

A Rede LVQ, implementada com a arquitetura de rede neural competitiva, possui uma camada oculta com a quantidade de neurônios igual à quantidade de *clusters*. Cada neurônio  $i$  corresponde a um dos *clusters* e os pesos  $\mathbf{w}_i$  correspondem ao centroide do *cluster*. Em particular nesse trabalho, cada *cluster* representará um protótipo de microestado que se deseja obter ou ajustar.

Existem diferentes possibilidades para o treinamento da rede LVQ, sendo o algoritmo utilizado neste trabalho um algoritmo de treinamento supervisionado, tendo sido utilizada uma rede do tipo LVQ1, detalhada a seguir.

### 2.3.2. Treinamento LVQ1

Para o treinamento de uma rede do tipo LVQ1, o primeiro passo é a inicialização dos pesos  $\mathbf{w}_i$ . Esse passo é muito importante, pois pode definir se haverá uma convergência adequada do processo de treinamento para cada um dos *clusters* ou se algum ficará subutilizado (DU, 2010). Esse problema de subutilização é causado pelo fato de que apenas o protótipo vencedor é atualizado a cada etapa do treino, como será detalhado a seguir. Para minimizar a chance de que a subutilização aconteça, devem-se inicializar os pesos com vetores adequados (DU, 2010).

A inicialização da rede LVQ é feita de duas maneiras distintas no presente trabalho, para cada uma das seguintes implementações: i) ajuste dos protótipos previamente obtidos pelo *k-means* modificado – os pesos são inicializados com os protótipos a serem ajustados, para que o algoritmo possa encontrar um mínimo local ao redor dos centroides obtidos pelo *k-means* modificado (BLACHNIK; DUCH, 2011); ii) Agrupamento utilizando somente a rede LVQ – os pesos são inicializados com valores aleatórios para que o treino possa analisar todo o espaço de busca (DU, 2010).

A cada época de treinamento, uma das amostras de treino é apresentada à rede, associada à sua respectiva classe, sendo então calculadas as distâncias entre o vetor de entrada e os centroides de cada *cluster*, representados pelos vetores de peso  $\mathbf{w}_i$ .

O centroide do *cluster* vencedor  $\mathbf{w}_c$  é, então, movido na direção do vetor de entrada  $\mathbf{u}(t)$ , caso pertençam à mesma classe, ou afastado caso pertençam a classes diferentes (KOHONEN, 1995), usando para isso uma taxa de aprendizagem definida por  $\alpha(t)$ , conforme

$$\mathbf{w}_c(t+1) = \mathbf{w}_c(t) + \alpha(t)[\mathbf{u}(t) - \mathbf{w}_c(t)], \quad (6)$$

se  $\mathbf{u}$  e  $\mathbf{w}_c$  pertencem à mesma classe,

$$\mathbf{w}_c(t+1) = \mathbf{w}_c(t) - \alpha(t)[\mathbf{u}(t) - \mathbf{w}_c(t)], \quad (7)$$

se  $\mathbf{u}$  e  $\mathbf{w}_c$  pertencem a classes diferentes,

$$\mathbf{w}_i(t+1) = \mathbf{w}_i(t) \text{ para } i \neq c. \quad (8)$$

A taxa de aprendizagem  $\alpha(t)$  é definida no intervalo  $\alpha(t) \in [0,1]$  e é, em geral, implementada como uma função monotônica decrescente sendo recomendável utilizar valores iniciais  $\alpha(t) < 0,1$ . A regra exata da taxa de aprendizagem não é crucial, desde que a quantidade de épocas de treinamento seja suficiente (KOHONEN, 1995). A distância entre um dado vetor de entrada  $\mathbf{u}$  e o centroide de cada *cluster*  $\mathbf{w}_i$  é calculado pela distância Euclidiana por (DONGORAN *et al.*, 2018)

$$\text{dist}(\mathbf{u}, \mathbf{w}_i) = \sqrt{(\mathbf{u} - \mathbf{w}_i)^T (\mathbf{u} - \mathbf{w}_i)} \quad , \quad (9)$$

em que  $\mathbf{u}$ ,  $\mathbf{w}_i$  são vetores coluna e  $T$  é a transposta.

Implementada conforme o algoritmo da LVQ1, a rede é treinada de maneira supervisionada num esquema de “recompensa” e “punição”, ajustando a posição do centroide de cada *cluster* de forma a minimizar o erro de classificação. Analogamente ao algoritmo do *k-means* modificado, utilizado na literatura para obtenção de microestados, o LVQ utiliza a estratégia *winner-take-all* (do inglês, o vencedor leva tudo) (DONGORAN *et al.*, 2018).

---

### 2.3.3. *Early stopping*

Algoritmos de aprendizado de máquina podem se especializar demais nos dados de treinamento, causando *overfitting* e reduzindo drasticamente a capacidade de generalização para novos dados. Quando ocorre *overfitting*, a taxa de erro para os dados de treino é mínima, entretanto o erro para novos dados de entrada é muito grande. Além do problema de *overfitting*, o tempo de processamento dos dados, quando utilizadas muitas amostras de treino, passa a se tornar um problema considerável (MATHWORK, 2020).

Dessa forma, o treinamento da rede LVQ foi implementado utilizando *Early Stopping*, que consiste em validar as taxas de erro de treinamento ao longo do tempo e a probabilidade de aumento ou diminuição dessa taxa, através de um subconjunto aleatório extraído dos dados de treino, após uma dada quantidade de épocas de treinamento.

Esse esquema de treino com validação consegue reduzir consideravelmente o tempo de treino e a quantidade de épocas necessárias para a convergência da rede LVQ.

### 2.4. Seleção Sequencial de Características e Visualização

A Seleção Sequencial de Características (*Sequential Feature Selection*) é utilizada para verificação de quais características dos dados possui maior relevância na discriminação de amostras de diferentes classes na base de dados.

Esse tipo de seleção de características possui fundamentalmente dois componentes. O primeiro deles é uma função de custo, que representa o critério de busca que o algoritmo tenta minimizar. Neste trabalho foi utilizada para minimização a medida de taxa de erro, calculada como a razão entre a quantidade de amostras atribuída à classe errada e o total de amostras sendo classificadas (ver Seção 2.5.6).

---

O segundo componente é o algoritmo de busca sequencial, que adiciona ou remove as características de um *subset* predeterminado de características candidatas, para avaliação de quais conjuntos de características diminuem o critério ao longo do tempo.

O método de seleção de características possui duas variantes, adicionando ou removendo as características sequencialmente. A que foi utilizada neste trabalho é a *Sequential Forward Selection* (SFS), onde as características são adicionadas de maneira sequencial até que a adição de novas características não altere mais o critério além do limiar mínimo estabelecido. O método utiliza validação cruzada *k-fold* para avaliar o comportamento do critério em diferentes partições dos dados (GUYON; ELISSEEFF, 2003). Os métodos de particionamento dos dados estão descritos na Seção 2.5.5.

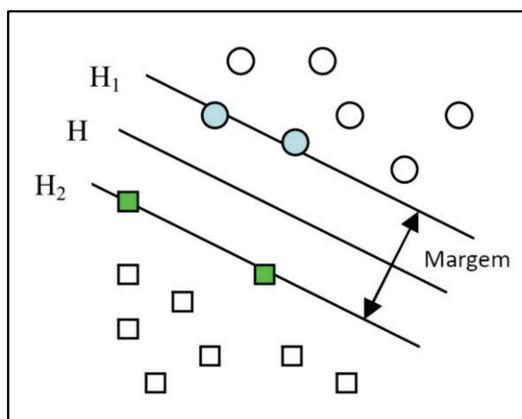
Complementarmente, foi gerada uma visualização, sobre as amostras de teste representadas pelo vetor de características, após definição de quais características serão utilizadas para treino e classificação, projetadas no plano bidimensional através do t-SNE (*t-distributed Stochastic Neighbor Embedding*) com distância Euclidiana para análise visual da separação entre classes (MONTALVÃO; MIRANDA; DORIZZI, 2021).

## 2.5. Classificadores

### 2.5.1. *Support Vector Machine*

O SVM (*Support Vector Machine*) é um tipo de classificador que tenta definir um hiperplano capaz de separar as amostras de diferentes classes (JIANG; HE; GUO, 2010; HU; ZHANG, 2019). Um classificador do tipo SVM constrói o modelo de forma que o hiperplano de separação entre classes aproxima a fronteira ótima de decisão, maximizando a margem de separação entre classes no espaço de características (HU; ZHANG, 2019). Na Figura 8 pode-se verificar o princípio de construção do SVM, onde  $H_1$  e  $H_2$  são os vetores de suporte (*Support Vectors*), construídos a partir das amostras localizadas na borda de cada *cluster*, e  $H$  é a fronteira de separação entre classes que maximiza a margem (*margin*).

Figura 8. Exemplo de um classificador SVM



Adaptado de Jiang, He e Guo (2010).

A fronteira de decisão é uma reta ou hiperplano, quando o espaço de características possui mais que duas dimensões, de forma a maximizar a margem de separação entre classes. No caso do SVM não-linear, é aplicado um *kernel*  $K(\mathbf{x}, \mathbf{y})$ , que faz um mapeamento dos vetores de características para um novo espaço, geralmente de dimensão maior do que o espaço original, de forma que uma separação linear entre os dados possa ocorrer (JIANG; HE; GUO, 2010).

Algumas das principais funções utilizadas para o *kernel* do SVM são a polinomial, a radial e a gaussiana (JIANG; HE; GUO, 2010), respectivamente, por

$$K(\mathbf{x}, \mathbf{y}) = (\mathbf{x} \cdot \mathbf{y} + 1)^m, \quad (10)$$

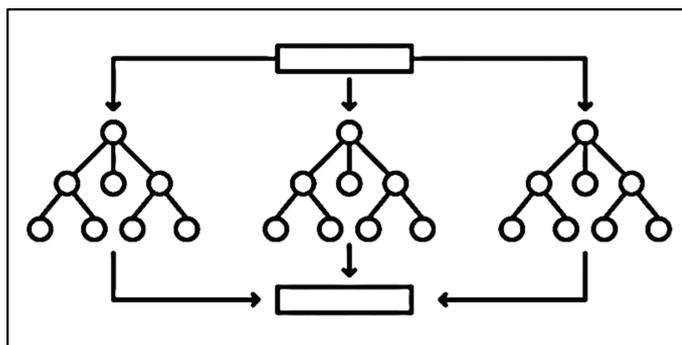
$$K(\mathbf{x}, \mathbf{y}) = \exp(-\gamma \|\mathbf{x} - \mathbf{y}\|^2), \text{ e} \quad (11)$$

$$K(\|\mathbf{x} - \mathbf{y}\|) = \exp\left(\frac{-\|\mathbf{x} - \mathbf{y}\|^2}{(2\sigma)^2}\right). \quad (12)$$

### 2.5.2. *Random Forest*

Um Classificador do tipo RF (*Random Forest*) consiste de um agrupamento (em inglês denominado *ensemble*) de classificadores do tipo *tree predictor*. Cada classificador individual é treinado com diferentes parâmetros de inicialização de forma que, ao agrupar vários destes classificadores, a capacidade de discriminar diferentes classes é maximizada.

Figura 9. Exemplo de um classificador RF



Fonte: *shutterstock.com*.

A Figura 9 ilustra um classificador do tipo *Random Forest*. Para classificar uma nova amostra apresentada ao classificador do tipo RF, cada classificador individual que compõem o *ensemble* “vota” na classe à qual aquela amostra deveria ser atribuída, e a classe mais popular é a vencedora (BREIMAN, 2001).

Esse tipo de classificador possui uma melhoria considerável na acurácia de classificação em relação aos classificadores individuais e o erro de generalização converge a um limite mínimo, conforme o número de árvores  $N_t$  fica grande o suficiente (BREIMAN, 2001).

### 2.5.3. *k Nearest Neighbors*

O algoritmo de classificação kNN (*k Nearest Neighbors*) é um dos mais clássicos da literatura. O kNN é um classificador não-paramétrico, que classifica cada amostra de acordo com a maioria de seus vizinhos, considerada uma medida de distância entre as amostras (BABLANI; EDLA; DODIA, 2018).

Tradicionalmente, na ausência de algum conhecimento *a priori* sobre os dados e sua distribuição, a distância Euclidiana é utilizada, entretanto, há uma série de medidas de distância que podem ser utilizadas com o classificador do tipo kNN, por exemplo a distância de Mahalanobis, a distância de Chebychev, entre outras.

---

#### 2.5.4. Otimização Bayesiana

A Otimização Bayesiana consiste da busca pelo mínimo de uma função objetivo  $f(\mathbf{x})$ , em que  $\mathbf{x} \in \mathcal{X}$ , que é assumido pertencer a  $\mathbb{R}^D$ . Diferente de outros tipos de otimização, o algoritmo de otimização Bayesiana considera toda a informação passada a respeito de  $f(\mathbf{x})$ , para decidir para qual lugar de  $\mathcal{X}$  deve-se fazer a próxima avaliação da função que está sendo minimizada, não apenas no gradiente local ou em aproximações Hessianas (SNOEK; LAROCHELLE; ADAMS, 2012).

Isso resulta em um procedimento capaz de localizar o mínimo de funções que não são convexas com uma quantidade relativamente pequena de avaliações (SNOEK; LAROCHELLE; ADAMS, 2012).

Para isso, é necessário definir uma distribuição de probabilidades prévias (*priors*), em geral com distribuição gaussiana, e escolher uma função de aquisição utilizada para construção do modelo *a posteriori* para cálculo do próximo ponto a ser avaliado.

Com isto, é construída e refinada uma superfície (*surrogate function*) que é uma estimativa da função objetivo e que é utilizada para definir as próximas regiões a serem exploradas.

#### 2.5.5. *Holdout e cross validation*

Conforme Faceli (2011), ao utilizar um mesmo conjunto de amostras para treino e teste de preditores, num processo denominado ressubstituição, em geral, obtém-se taxas de erro ou acerto otimistas. Esse tipo de avaliação do desempenho de um preditor é denominado aparente.

De forma a melhorar a estimativa do desempenho dos preditores, devem-se utilizar técnicas de amostragem como *holdout*, validação cruzada (do inglês *cross validation*) e outros, para os casos em que se tem um único conjunto de dados. Esse tipo de

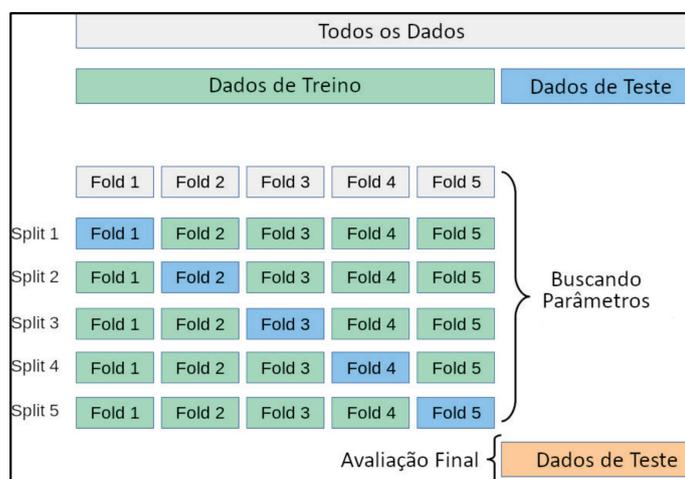
amostragem favorece a aproximação de estatísticas mais realistas para as métricas de classificação (FACELI, 2011).

No esquema de amostragem *holdout*, o conjunto de dados é dividido de forma que um percentual das amostras é deixado de fora do conjunto de treino, sendo utilizado para teste.

Já na validação cruzada *k-fold* os dados são particionados em  $k$  subconjuntos, que podem ser estratificados ou não, sempre com tamanhos aproximadamente iguais. Cada um dos *folds* possui, então, subconjuntos que são utilizados de maneira separada para treino, validação e teste. Um exemplo de amostragem com 5 *folds* está ilustrado na Figura 10.

Dessa forma, pode-se utilizar a média da taxa de erro ou acerto, bem como o desvio padrão para uma estimativa mais realista do desempenho dos preditores utilizados para classificar os dados.

Figura 10. Exemplo de partição *k-fold* para validação cruzada



Adaptado de Scikit-Learn Developers (2011).

Um desvio padrão muito grande pode indicar, por exemplo, que existe muita variabilidade nos resultados, o que significa que o modelo de preditor utilizado apresenta sensibilidade às amostras utilizadas para treino, mesmo sendo todas oriundas de uma mesma distribuição (FACELI, 2011).

A análise do desvio padrão pode auxiliar na escolha entre dois algoritmos com desempenho médio próximo, sendo o de menor desvio padrão preferível por ser mais

---

estável a variações nos dados de treino. O mais adequado, entretanto, é a aplicação de análises estatísticas para comparação do desempenho dos classificadores, utilizando-se testes de hipóteses estatísticas (FACELI, 2011).

### 2.5.6. Métrica dos classificadores

Para definição do desempenho dos classificadores implementados no presente trabalho e, em particular para possibilitar a comparação com outros trabalhos da literatura, é utilizada a métrica de taxa de erro, calculada por (LOPEZ *et al.*, 2016)

$$\text{Erro} = \left(1 - \frac{N_A}{N_T}\right) \times 100\% \quad , \quad (13)$$

onde  $N_A$  é o número de sessões de EEG atribuídas à classe correta e  $N_T$  é o número total de sessões de EEG sendo classificadas entre normal/anormal.

### 2.5.7. Teste de Hipóteses

Uma vez definidas as taxas de erro de cada tipo de classificador, treinado e testado utilizando-se das técnicas de otimização com validação cruzada *k-fold* e *holdout*, é necessário testar estatisticamente a relevância dos resultados obtidos.

Nesse sentido, foram utilizados dois testes estatísticos para comparação do desempenho dos classificadores, comparando o desempenho entre si bem como comparando o desempenho para os diferentes métodos de agrupamento implementados.

O primeiro teste estatístico é o teste de Friedman (HOLLANDER; WOLFE, 2013), que avalia a hipótese nula  $H_0$  de que todos os classificadores possuem o mesmo desempenho contra a hipótese alternativa  $H_1$  de que os classificadores avaliados não têm o mesmo desempenho. Esse teste é o mais recomendado na literatura para esta tarefa (DEMŠAR, 2006). Adicionalmente, outro teste estatístico *post-hoc* é feito para comparar quais pares de classificadores são estatisticamente diferentes dos outros, e quais não são. O teste mais

---

recomendado na literatura para esse fim é o teste de Wilcoxon-Mann-Whitney (DEMŠAR, 2006; BENAOLI; CORANI; MANGILI, 2016).

O teste de hipótese de Friedman assume as duas condições a seguir (HOLLANDER; WOLFE, 2013): Todos os valores testados são obtidos de populações com a mesma distribuição; e todas as observações são mutuamente independentes.

O teste de Wilcoxon-Mann-Whitney (HOLLANDER; WOLFE, 2013) consiste de um teste não paramétrico pareado, aplicado a duas amostras independentes. Esse teste verifica a hipótese nula  $H_0$  de que as duas amostras pertencem a uma mesma distribuição contínua com iguais medianas. O teste também assume a independência entre as amostras. Com a combinação dos dois testes é possível verificar se há diferença estatisticamente relevante nos resultados dos classificadores.

O nível de significância utilizado em ambos os testes estatísticos foi de  $\alpha = 0,05$ .

## 2.6. Base de Dados

A pesquisa utilizou a base denominada TUAB (*The TUH Abnormal EEG Corpus*) (LOPEZ, 2017) que consiste de um subconjunto da base de dados denominada TUEG (*The Temple University Hospital EEG Data Corpus*) (OBEID; PICONE, 2016). As bases são disponibilizadas<sup>1</sup> pelo *Neural Engineering Data Consortium* (OBEID; PICONE, 2016). A base de dados TUEG é o maior conjunto de dados clínicos de EEG disponibilizada publicamente (OBEID; PICONE, 2016). Na Tabela 2 está descrita a quantidade de pacientes que constam da base de dados TUAB, versão 2.0.0, consistindo de um total de 2.383 pacientes. Importante ressaltar que nas versões anteriores da base constavam alguns dados que estavam duplicados e que foram mesclados, dessa forma a versão 2.0.0 tem algumas sessões a menos que a anterior que, inclusive, foi a base utilizada pelo trabalho de Lopez (2017).

---

<sup>1</sup> [www.nedcdata.org](http://www.nedcdata.org)

---

**2.6.1. Base TUAB versão 2.0.0**

A base de dados TUAB, na versão 2.0.0, é dividida em 2.993 sessões de EEG, que totalizam 1.142 horas, subdivididas em *Train* e *Evaluation* sem sobreposição de pacientes. O conjunto de pacientes está detalhado na Tabela 2. Todos os arquivos da base possuem pelo menos 15 minutos de duração da gravação do sinal de EEG (LOPEZ, 2017).

Tabela 2. Conjunto de Pacientes TUAB

Descrição	Normal	Anormal	Total
<i>Evaluation</i>	148	105	253
<i>Train</i>	1.237	893	2.130
Total	1.385	998	2.383

Na Tabela 3 há um detalhamento das sessões da base de dados TUAB. A quantidade de sessões é maior que a quantidade de pacientes porque alguns dos pacientes acompanhados pelo *Temple University Hospital* possui mais de uma sessão de EEG registrada em seu histórico médico.

Tabela 3. Conjunto de Sessões TUAB

Descrição	Normal	Anormal	Total
<i>Evaluation</i>	150	126	276
<i>Train</i>	1.371	1.346	2.717
Total	1.521	1.472	2.993

A base de dados TUAB, em todas as versões, é composta somente por sinais de EEG capturados na montagem denominada TCP/AR. A nomenclatura TCP é referente à montagem bipolar *Temporal Central Parassagital* dos eletrodos (FERRELL *et al.*, 2020) no sistema 10-20 (ilustrada na Figura 2, Seção 2.1.1). A terminologia AR (*Average Referenced*) indica que o sinal é referenciado à média, o que é importante para a análise de microestados, uma vez que na literatura e nas ferramentas computacionais é usual assumir essa referência (POULSEN *et al.*, 2018).

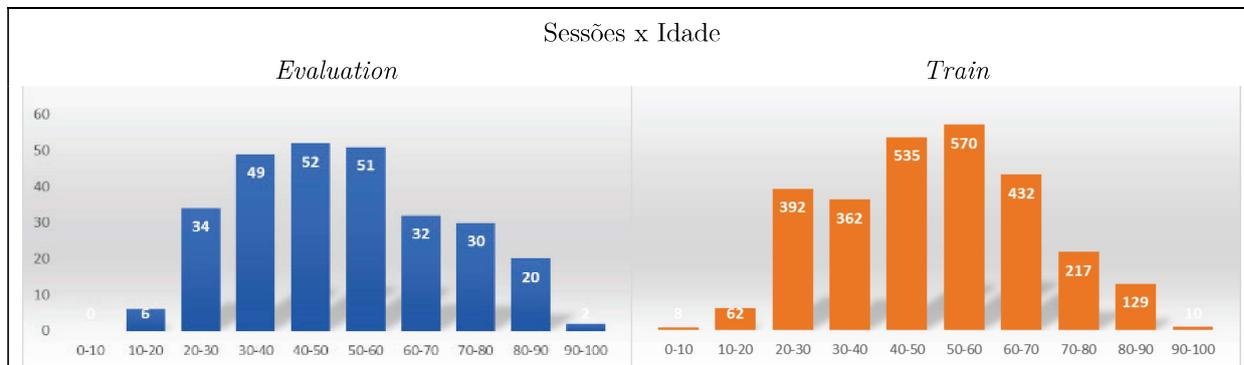
Na Tabela 4 e na Figura 11, podem-se verificar as distribuições de gênero e de idade entre as sessões da base.

Tabela 4. Distribuição de gênero por sessão (TUAB)

Gênero	<i>Evaluation</i>		<i>Train</i>	
	M	F	M	F
Total	128	148	1.270	1.447

Todas as sessões possuem um arquivo de texto associado contendo o laudo clínico do neurologista, que possuem um breve histórico médico do paciente, remédios utilizados, impressões clínicas e uma classificação do EEG entre normal e anormal pelo especialista. Os sinais classificados pelo neurologista como anormais consistem de uma diversidade de desordens neurológicas, sendo algumas das doenças citadas nos laudos: AVC, concussão e epilepsia (OBEID; PICONE, 2016; LOPEZ, 2017).

Figura 11. Distribuição de idade por sessão



Fonte: (LOPEZ, 2017).

As taxas de amostragem variam na base de dados, sendo sempre de, no mínimo, 250 Hz (HARATI *et al.*, 2014). Os sinais foram capturados em diversas configurações de ambiente clínico do hospital universitário *Temple University Hospital (TUH)*, em estado de repouso (LOPEZ, 2017). A base de dados original foi processada automaticamente para a remoção dos seguintes padrões indesejados no sinal: movimento dos olhos, artefatos e ruído de origem não cerebral (HARATI *et al.*, 2014).

---

### 2.6.2. Pré-processamento

As seguintes etapas de pré-processamento foram utilizadas no método proposto neste trabalho:

- seleção dos 21 canais de EEG do sinal, pois os sinais possuíam outros canais que não eram de EEG como, por exemplo, canal de eletrocardiograma (ECG) (FERRELL *et al.*, 2020);
- mapeamento dos eletrodos de cada canal em cada ponto do crânio conforme montagem 10-20 TCP/AR no EEGLab;
- reamostragem dos sinais para 250 Hz, caso necessário (para compatibilidade no EEGLab);
- aplicação de filtro do tipo passa banda com frequências de corte de 1 e 30 Hz – utilizando a função ‘*pop\_eegfiltnew()*’ do EEGLab, que aplica um filtro do tipo janela de *Hamming* (ZAPPASODI *et al.*, 2017; POULSEN *et al.*, 2018); e
- seleção de intervalo de 10 segundos para cada amostra com *offset* de 3 minutos do início da sessão.

Conforme Lopez (2017), 10 segundos representa o tempo médio de duração do EEG, necessário para que um especialista treinado seja capaz de classificá-lo como normal ou anormal com base na inspeção visual. No trabalho de Lopez (2017), alguns dos algoritmos utilizados analisam diversos trechos do sinal para cada sessão de EEG, em particular por utilizar análise temporal e espectral do sinal.

Devido à forma como é feita a implementação da análise de microestados, em particular à quantidade de espaço de memória e armazenamento, principalmente ao custo computacional, que é da ordem de  $O(n^2)$  para o *k-means* (DU, 2010), por exemplo, a análise de trechos muito grandes de sinal ou mais de um trecho por sessão torna-se inviável.

---

Assim, a metodologia implementada foi a análise de um trecho único para cada sessão de EEG, cujo intervalo de tempo fosse o mais próximo daquele que um profissional qualificado utiliza. Nas etapas preliminares da pesquisa, resultados empíricos, considerado o *tradeoff* entre tamanho do trecho analisado, complexidade computacional e nível de erro dos classificadores, foi definido o intervalo de 10 segundos para cada amostra, ou seja, para cada sessão de EEG da base de dados.

Por último, foram descartados os 3 minutos iniciais de cada sessão de EEG, como forma de desconsiderar alterações no sinal em repouso oriundas do ambiente clínico (como ruídos e interferências externas, desconforto do paciente com os eletrodos e com o ambiente, etc.). Esse intervalo foi selecionado empiricamente nas etapas preliminares da pesquisa.

### 3. METODOLOGIA E IMPLEMENTAÇÃO

#### 3.1. Seleção de dados e pré-processamento

Utilizando a base de dados descrita na Seção 2.6, inicialmente foram selecionados conjuntos de treino balanceados entre as classes normal e anormal. Os testes sobre a base de dados foram executados para 162 e 500 sessões de treino e, por fim, para todas as 2717 sessões de treino.

Todas as amostras são pré-processadas conforme discutido na Seção 2.6.2, e cada amostra alimentada aos algoritmos de análise de microestados e classificação se refere a um trecho de 10 segundos de EEG extraído de cada uma das sessões selecionadas, desconsiderados os 3 minutos iniciais do sinal.

#### 3.2. Obtenção dos Microestados e pós-processamento

Para obtenção dos microestados, todos os trechos devem ser alimentados para o algoritmo de obtenção dos protótipos de microestados a ser utilizado. Para isso, todas as amostras de treino, independente de rótulo, são selecionadas para o cálculo de GFP conforme Equação (1). São selecionados até 100 pontos de máximo da curva GFP de cada uma das amostras de treino, escolhidos aleatoriamente (POULSEN *et al.*, 2018).

O resultado desse processo é salvo em um novo *set* do EEGLab. Esse novo *set* contendo todos os pontos de máximo da curva GFP das amostras de treino é utilizado para obtenção dos protótipos de microestados.

A obtenção dos protótipos dos microestados para a primeira hipótese é feita com o algoritmo do *k-means* modificado; de forma a obter os microestados aprimorados para a segunda hipótese, os pesos da rede LVQ são inicializados com os protótipos obtidos pelo *k-means* modificado. A rede LVQ é, então, treinada com os picos GFP das amostras de treino do sinal EEG, cada pico associado à respectiva classe da amostra, utilizando como

---

hiperparâmetros uma taxa de aprendizagem de  $\alpha = 0,005$  e máximo de 50 épocas, com o *Early Stopping* conforme descrito na Seção 2.3.3. Esses hiperparâmetros foram obtidos empiricamente, de forma que, nos testes executados, a convergência se deu sempre antes de chegar ao limite de épocas. Após o treino da rede LVQ, seus pesos são normalizados pela norma  $L_2$ , e são chamados aqui de microestados ajustados. Para a terceira hipótese, foram utilizados os mesmos hiperparâmetros e a normalização utilizados para a segunda hipótese, entretanto, a rede LVQ é inicializada com pesos aleatórios, de forma que o treinamento supervisionado obtém os protótipos de microestados de maneira independente do *k-means* modificado.

Após a definição dos protótipos dos microestados para cada uma das hipóteses, eles são importados para execução do *backfit* de cada amostra de treino e de teste individualmente. Após o *backfit* é feito o pós-processamento discutido na Seção 2.2.1, denominado *temporal smoothing*, cujo limiar é determinado empiricamente, para rejeição de segmentos muito curtos. Neste trabalho, foi utilizado o limiar de 30 ms, levando-se em consideração, também, que a duração típica de um microestado é da ordem de centenas de milissegundos (LEHMANN; OZAKI; PAL, 1987; ZAPPASODI *et al.*, 2017; FU; CHEN; XIONG, 2018).

### 3.3. Extração e Seleção de Características, Treino e Classificação

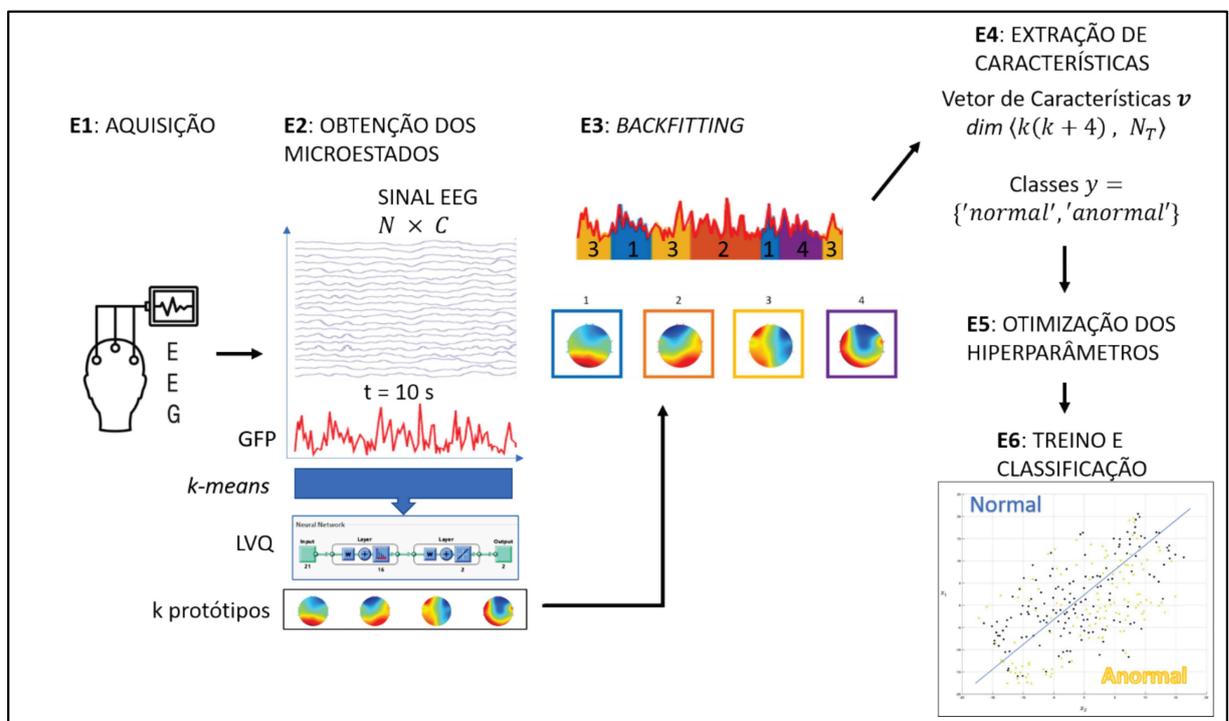
A última etapa antes do treino dos classificadores é a extração de características. Para tanto, é feito o cálculo das estatísticas dos protótipos de microestados obtidos na etapa de segmentação do sinal de EEG. As características e estatísticas extraídas da configuração topográfica do sinal de EEG através da representação pelos microestados, descritas na Seção 2.2.2, são utilizadas para gerar o vetor  $\mathbf{v}$  de características (ou *features*, em inglês), utilizado para o treino e teste dos classificadores.

Cada amostra selecionada da base de dados, que se refere a um trecho de 10 segundos de sinal EEG de cada uma das sessões utilizadas, passa a ser representada por seu vetor de

características  $\mathbf{v}$ , cuja dimensão será de um total de  $k(k + 4)$  atributos, sendo Duração, Ocorrência, Cobertura e GEV para cada microestado, totalizando  $k \times 4$ , mais  $k^2$  Probabilidades de Transição entre cada um dos  $k$  protótipos de microestados (ver Seção 2.2.2). Dessa forma, o vetor de características  $\mathbf{v}$  representando cada amostra de sinal de 10 segundos, extraídos de cada uma das sessões de EEG da base de dados, irá possuir o tamanho 96 para  $k = 8$  microestados e 320 para  $k = 16$  protótipos de microestados.

O número do total de sessões da base a serem classificadas, com base nos seus respectivos vetores de características ( $\mathbf{v}$ ), extraídos das amostras de 10 segundos de sinal de EEG de cada sessão, é o número  $N_T$ , e as classes atribuídas a cada uma das  $N_T$  sessões é  $\mathbf{y} = \{\text{'normal'}, \text{'anormal'}\}$ , com base na atribuição realizada pelo classificador treinado com os vetores de características. O número de sessões atribuídas à classe correta no processo de classificação é  $N_A$ . A Figura 12 apresenta um diagrama com o fluxo dos dados durante o processamento.

Figura 12. Diagrama de fluxo de dados



Fonte: o autor.

---

Com a suposição de que nem todas as características possuem a mesma relevância para discriminar as classes no espaço de características das amostras, representadas por seus vetores de características, foi executada uma etapa de Seleção de Características (*Sequential Feature Selection* – SFS). A seleção de características foi executada com um subconjunto extraído dos dados de treino composto de 500 sessões, utilizando um particionamento *10-fold* para validação cruzada do algoritmo de SFS. Nesta etapa, foi utilizado apenas o *k-means* modificado, utilizando  $k = 8$  protótipos de microestado. Um classificador SVM foi utilizado para verificar se há melhoria estatisticamente relevante nas taxas de erro de classificação considerando o subconjunto de características selecionado pelo algoritmo em relação ao uso de todas as características para representação dos dados e classificação.

Para fins de inspeção visual da distribuição das sessões da base entre as duas classes, é gerada a visualização das amostras (trechos de 10 segundos de cada sessão com *offset* de 3 minutos) representadas por seus vetores de características ( $\mathbf{v}$ ), projetadas no plano bidimensional pelo t-SNE (*t-distributed Stochastic Neighbor Embedding*).

Após a definição das características a serem consideradas no treino e teste dos algoritmos, foram utilizados no presente trabalho os seguintes classificadores e as faixas de busca dos hiperparâmetros, realizada pela Otimização Bayesiana, com validação cruzada *10-fold* sobre as sessões de treino e utilizando a função de aquisição *expected-improvement-plus*, com um número máximo de 30 iterações:

- a) Para o SVM são otimizados a função do *kernel* (polinomial, gaussiana e linear), escala do *kernel* (0,001 a 1000), e *box constraint* (0,001 a 1000);
- b) Para a RF é otimizado o número de árvores  $N_t$  (1 a 500); e
- c) Para o kNN é otimizado o valor de  $k$  (1 a 25) e a distância (chebychev, euclidiana e minkowski).

Na Figura 13 estão descritas e detalhadas todas as etapas da Metodologia do presente trabalho.

Figura 13. Etapas da Metodologia e Implementação



Fonte: o autor.

## 4. RESULTADOS E DISCUSSÃO

De forma a implementar e testar o uso da análise de microestados para classificação da base de dados TUAB, foi utilizado o EEGLab, versão 2020.0, para a plataforma computacional Matlab® (versão 2017a). O EEGLab (DELORME; MAKEIG, 2004) é uma ferramenta aberta para tratamento e análise de sinais de EEG contando com uma série de recursos pré-definidos, inclusive uma série de *toolboxes*, dentre as quais a *Microstate toolbox* (POULSEN *et al.*, 2018), que foi utilizada para as etapas da análise dos microestados (os *scripts* implementados para o presente trabalho podem ser acessados conforme APÊNDICE A).

### 4.1. Resultados da Seleção de Características e Visualização das Classes

Para o teste de seleção de características foram utilizadas 500 sessões de treino,  $k = 8$  microestados e agrupamento *k-means* modificado, utilizando uma partição *10-fold*. A função objetivo a ser minimizada foi a taxa de erro de classificação, utilizando um classificador do tipo SVM, com os critérios de parada de taxa de erro menor que  $1 \times 10^{-6}$ , ou obtenção de um valor de mínimo local, com execução sem limite na quantidade de iterações.

É possível verificar na Tabela 5 que o resultado de classificação com todos os atributos teve um desempenho médio estatisticamente superior ao da classificação com os subconjuntos de características selecionados pelo SFS, de acordo com o teste de Wilcoxon. Uma hipótese sobre o motivo desse resultado é a possibilidade do algoritmo SFS convergir a algum ponto de mínimo local da função a ser minimizada, assim, terminando a sua execução de forma prematura. Também foi observado, ao executar sucessivos testes de seleção de características, que o algoritmo alternava entre as características que eram selecionadas, de forma aparentemente aleatória, o que pode

indicar que não há uma grande diferença na capacidade de discriminação dos diferentes conjuntos de características testados pelo SFS para representação do sinal de EEG.

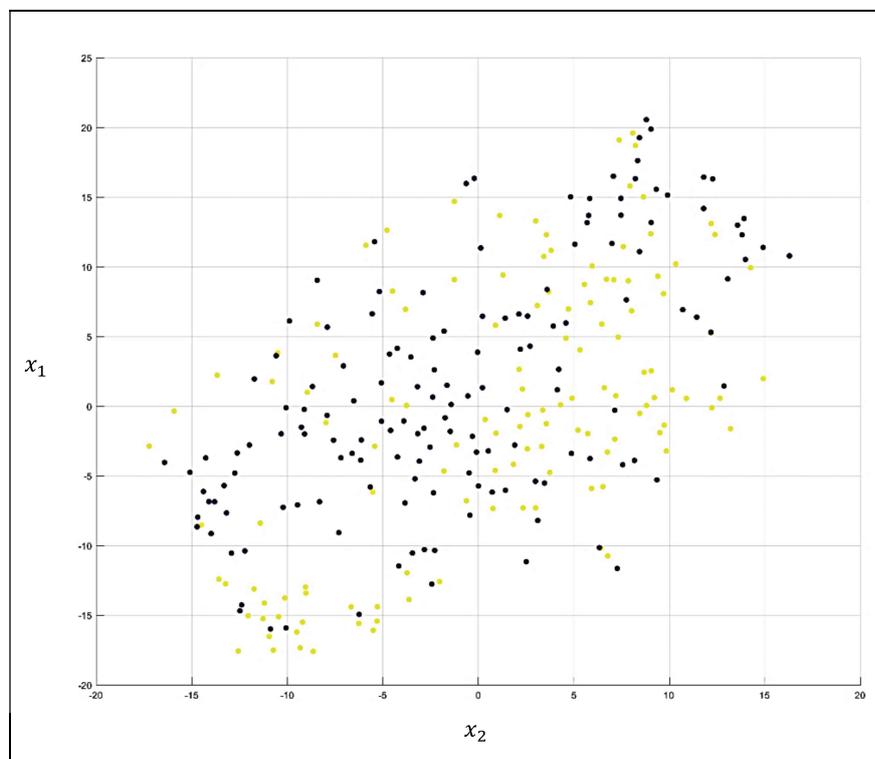
Tabela 5. Comparativo do classificador SVM com e sem SFS

	Todos os Atributos			SFS			Wilcoxon Test
	Média	Mediana	Desvio	Média	Mediana	Desvio	<i>p-value</i>
SVM	24,28%	24,60%	$\pm 0,70\%$	32,97%	33,00%	$\pm 1,45\%$	0,0002

Como o custo computacional dessa seleção de características é bastante alto e, uma vez que a SFS não demonstrou uma melhora estatística no desempenho médio do classificador testado, utilizando subconjuntos dos atributos para discriminação dos dados, nas etapas seguintes foram utilizados todos os atributos para treinamento dos classificadores.

Foi gerada a projeção com o t-SNE bidimensional e distância Euclidiana (ver Seção 2.4) para visualização das sessões de teste da base, utilizando a representação com todos os atributos, conforme pode-se verificar na Figura 14.

Figura 14. t-SNE bidimensional, usando 276 sessões de teste, aplicado sobre o vetor de características de cada amostra



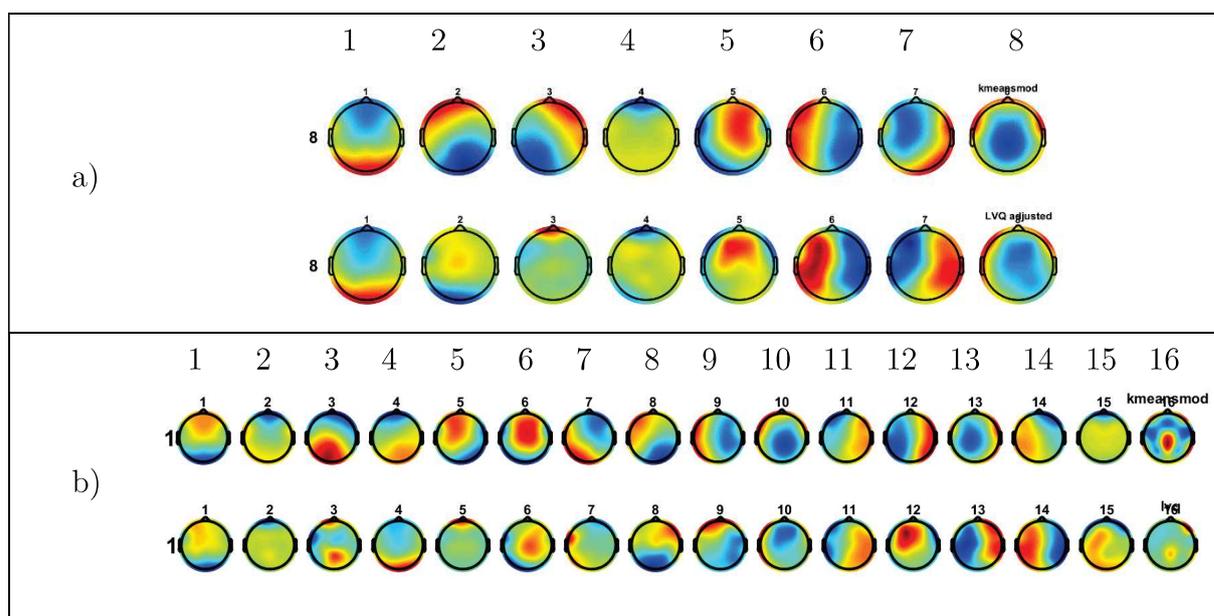
Fonte: o autor.

É possível observar na Figura 14 que não existe uma fronteira muito clara de separação entre as classes anormal (pontos em amarelo) e normal (pontos em azul escuro). A visualização pelo t-SNE ilustra, assim, que há uma sobreposição muito grande entre diferentes classes, ressaltando a complexidade do problema de detecção de sinais anormais de EEG.

#### 4.2. Análise dos Algoritmos de Agrupamento

O teste com diferentes tipos de configurações para a obtenção dos microestados, executados sobre um subconjunto de 500 sessões selecionadas aleatoriamente do conjunto de treino, foram: i) Agrupamento clássico não supervisionado com o *k-means* modificado; ii) Agrupamento não supervisionado com o *k-means* modificado e ajuste supervisionado dos protótipos com a rede LVQ; e iii) Agrupamento supervisionado somente com a rede LVQ.

Figura 15. Protótipos de microestado gerados somente com *k-means* modificado (acima) e gerados por *k-means* modificado e ajustados com LVQ (abaixo)

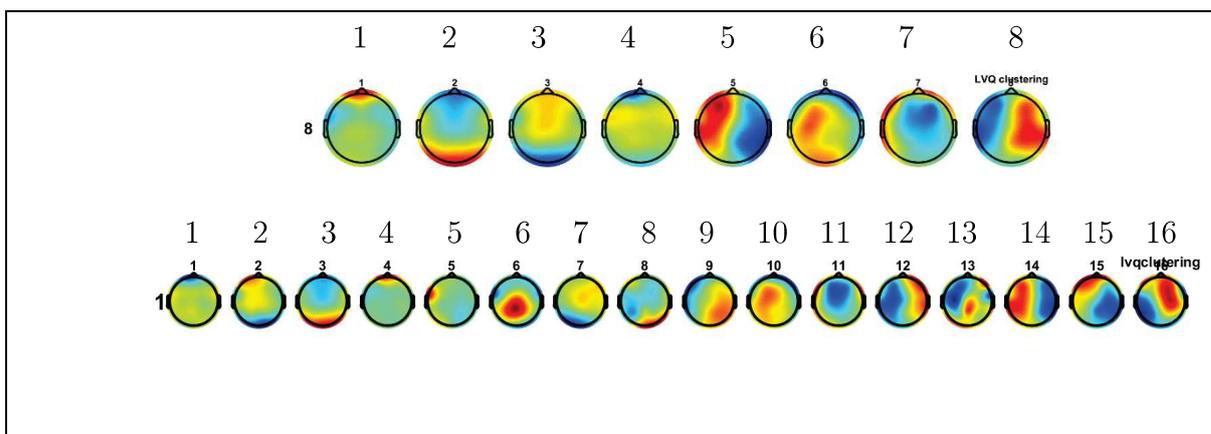


a) 8 protótipos e b) 16 protótipos de microestados

Fonte: o autor.

O aspecto dos microestados para  $k = 8$  e  $k = 16$  protótipos está ilustrado na Figura 15 sendo os protótipos de microestados obtidos pelo *k-means* modificado na primeira linha e os protótipos obtidos pelo *k-means* modificado e ajustados pela rede LVQ na segunda linha. Os protótipos de microestado utilizando somente a rede LVQ como agrupamento estão ilustrados na Figura 16.

Figura 16. Protótipos de microestado gerados apenas com a rede LVQ para  $k = 8$  e 16



Fonte: o autor.

Nas Figuras 15 e 16, os números acima das topografias se referem a cada um dos protótipos de microestado obtidos. Na Figura 15, os mapas obtidos apenas pelo *k-means* modificado (acima) e os mapas obtidos pelo ajuste com a rede LVQ (abaixo) são apresentados na sequência. Dessa forma, é possível identificar quais protótipos a rede LVQ ajustou.

É possível observar na Figura 15, por inspeção visual, que o LVQ faz uma série de ajustes nos protótipos previamente obtidos pelo *k-means* modificado. Por ser um algoritmo supervisionado, o LVQ busca ajustar os protótipos de microestado com a meta de maximizar a discriminação entre os dados, com base nas classes de cada sessão de treino (normal e anormal). É interessante observar que o LVQ parece selecionar alguns dos protótipos de microestado previamente obtidos pelo *k-means* modificado, mantendo-os quase inalterados, enquanto parece alterar completamente outros protótipos, como é o caso dos protótipos de número 2, 3 e 5, para o caso de  $k = 8$ , por exemplo. Outra

---

consideração a ser apontada é, para o caso  $k = 16$ , o fato de que o LVQ parece reagrupar protótipos muito semelhantes obtidos pelo *k-means* modificado, como é o caso dos protótipos de número 2 e 15 do *k-means* modificado e do protótipo de número 2 do LVQ. É possível conjecturar que isso ocorra devido à implementação feita dos algoritmos, uma vez que os pesos da rede LVQ buscam maximizar a discriminação entre classes, de maneira supervisionada, já o *k-means* modificado apenas agrupa os protótipos de forma não supervisionada, o que pode gerar protótipos duplicados (*i.e.* muito semelhantes).

Quando um protótipo discrimina bem a dinâmica associada à alguma patologia, por exemplo, o algoritmo de treinamento supervisionado favorece a sua manutenção e, caso este comece a ir em direção à dinâmica associada ao comportamento normal, o treino supervisionado penaliza, trazendo-o de volta à condição anterior, de forma a manter uma maior discriminação entre classes.

Como os protótipos muito parecidos contribuem de maneira mais ou menos equivalente na discriminação entre classes, os protótipos mais diferentes após o ajuste irão determinar um possível nível de variação nos resultados de classificação dos sinais de EEG de diferentes classes.

É esperado da rede LVQ um nível de melhoria na classificação em relação ao *k-means* modificado, uma vez que sua superfície discriminante, determinada por mapas auto organizáveis de Kohonen (1995), é muito mais complexa que a superfície determinada por *sets* de Voronoi (DU, 2010) inerente ao *k-means* modificado.

### 4.3. Comparativo dos Algoritmos de Agrupamento e dos Classificadores

Utilizando cada um dos algoritmos de agrupamento, os classificadores foram treinados, com o subconjunto de 500 sessões de treino, com uma partição *10-fold* para validação cruzada na Otimização Bayesiana para definição dos hiperparâmetros.

Os classificadores foram testados sobre 10 execuções de *holdout*, com 90% das sessões utilizadas para treino dos classificadores com os hiperparâmetros obtidos pela Otimização

Bayesiana, e 10% das sessões utilizadas para teste, sempre com diferentes partições, estratificadas e de igual proporção.

Os classificadores alcançaram as médias de taxa de erro de classificação  $\mu_E$  e desvio padrão  $\sigma_E$ , conforme detalhado na Tabela 6, para 8 e 16 microestados. Os números entre colchetes identificam as diferentes configurações testadas.

As configurações com as menores taxas de erro de classificação sobre o conjunto de teste, para cada um dos testes, ou seja, utilizando 8 e 16 microestados para discriminação das amostras, foram a [5] totalizando  $23,19\% \pm 0,74\%$  e a [14] totalizando  $25,36\% \pm 0,76\%$ , respectivamente, ambas obtidas pelo classificador *Random Forest* com o uso do *k-means* modificado associado ao ajuste dos protótipos pela rede LVQ.

Tabela 6. Comparação da taxa de erro para diferentes tipos de agrupamento com 500 sessões de treino, 8 e 16 microestados

8 microestados						
	<i>k-means</i> modificado		<i>k-means</i> com ajuste LVQ		LVQ	
	$\mu_E$	$\sigma_E$	$\mu_E$	$\sigma_E$	$\mu_E$	$\sigma_E$
SVM	[1] 24,28%	$\pm 0,70\%$	[4] 24,27%	$\pm 0,69\%$	[7] 30,43%	$\pm 0,69\%$
RF	[2] 26,45%	$\pm 0,93\%$	[5] 23,19%	$\pm 0,74\%$	[8] 29,71%	$\pm 0,92\%$
kNN	[3] 40,94%	$\pm 1,06\%$	[6] 29,35%	$\pm 0,92\%$	[9] 40,58%	$\pm 1,09\%$
16 microestados						
	<i>k-means</i> modificado		<i>k-means</i> com ajuste LVQ		LVQ	
	$\mu_E$	$\sigma_E$	$\mu_E$	$\sigma_E$	$\mu_E$	$\sigma_E$
SVM	[10] 36,96%	$\pm 1,31\%$	[13] 25,73%	$\pm 0,93\%$	[16] 28,62%	$\pm 0,94\%$
RF	[11] 26,09%	$\pm 0,92\%$	[14] 25,36%	$\pm 0,76\%$	[17] 26,45%	$\pm 0,70\%$
kNN	[12] 36,59%	$\pm 0,82\%$	[15] 43,12%	$\pm 0,91\%$	[18] 42,39%	$\pm 0,89\%$

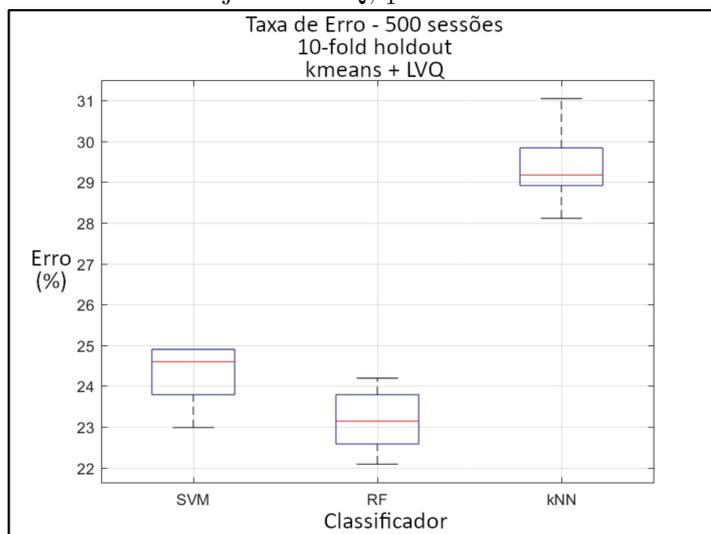
Em geral, o SVM obteve resultados que não foram muito distantes dos resultados do classificador RF, em algumas configurações conseguindo superá-lo, como por exemplo na configuração [1].

O classificador kNN foi superado pelos classificadores RF e SVM em todos os diferentes tipos de agrupamento. Além disso, o classificador RF obteve taxas de erro de classificação

mais consistentes para os diferentes tipos de agrupamento utilizados, sugerindo uma menor sensibilidade a variações nas amostras utilizadas para treino.

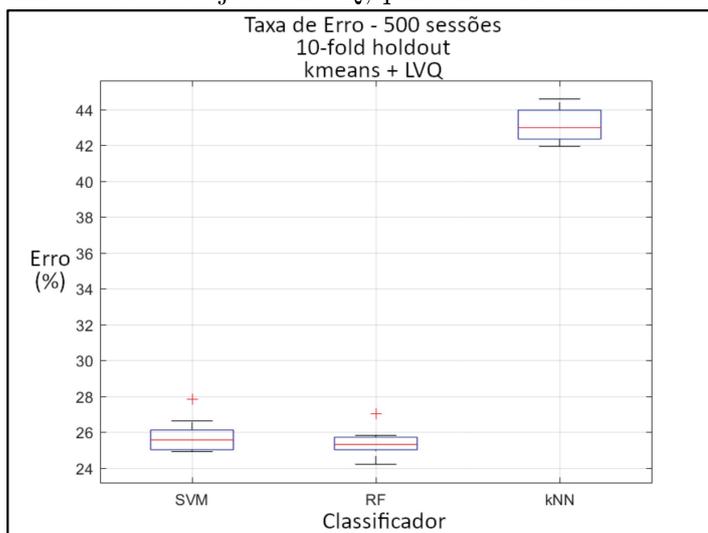
Nas Figuras 17 e 18, podem-se verificar as distribuições e a comparação das taxas de erro das configurações [4], [5] e [6]; e [13], [14] e [15]; através dos *boxplots*, para a agrupamento utilizando o *k-means* modificado com ajuste pela rede LVQ.

Figura 17. *Boxplot* do erro de classificação para 500 sessões de treino e agrupamento com *k-means* e ajuste LVQ, para  $k = 8$  microestados



Fonte: o autor.

Figura 18. *Boxplot* do erro de classificação para 500 sessões de treino e agrupamento com *k-means* e ajuste LVQ, para  $k = 16$  microestados



Fonte: o autor.

---

Para verificar a relevância dos resultados obtidos, foi executado o teste estatístico de Friedman, conforme discutido na Seção 2.5.7. O teste de Friedman é aplicado individualmente para cada execução, ou seja, com 8 e 16 microestados. Na comparação entre as configurações [1] a [9] ( $k = 8$ ), o teste confirma a Hipótese de que existe diferença estatisticamente relevante no *rank* médio.

Novamente, na comparação entre as configurações [10] a [18] ( $k = 16$ ), o teste confirma a Hipótese de que existe diferença estatisticamente relevante no *rank* médio. Testes adicionais de Wilcoxon pareado (ver Seção 2.5.7) revelam que as configurações [5] e [14] são estatisticamente superiores aos demais resultados para 8 e 16 microestados, respectivamente.

Dessa forma, pode-se afirmar que o melhor classificador e o agrupamento que melhor conseguiu discriminar os dados foram, respectivamente, o classificador *Random Forest* e a agrupamento *k-means* com ajuste LVQ, para  $k = 8$  microestados, alcançando uma taxa média de erro de classificação de  $23,19\% \pm 0,74\%$ , utilizando todos os atributos, a saber Ocorrência, Duração, Cobertura, GEV e Probabilidades de Transição.

#### 4.4. Comparativo de classificação para 8 e 16 microestados

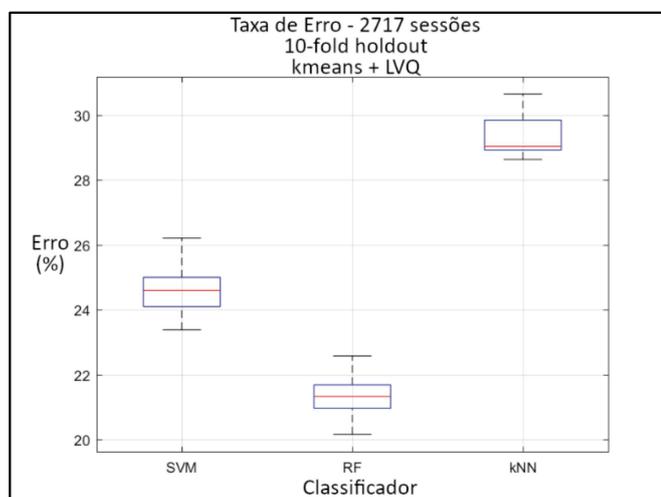
Com a informação do melhor método de agrupamento, passa-se à análise do comportamento dos classificadores para todos os dados da base. Assim, são obtidos os resultados dos testes executados para os três classificadores, utilizando todas as amostras, utilizando apenas o melhor algoritmo de agrupamento determinado anteriormente, ou seja, o *k-means* modificado com o ajuste da rede LVQ, para  $k = 8$  microestados e  $k = 16$  microestados.

A Tabela 7 apresenta os resultados obtidos com a implementação apresentada, com todas as amostras para os 3 classificadores utilizados. Os *boxplots* estão nas Figuras 19 e 20.

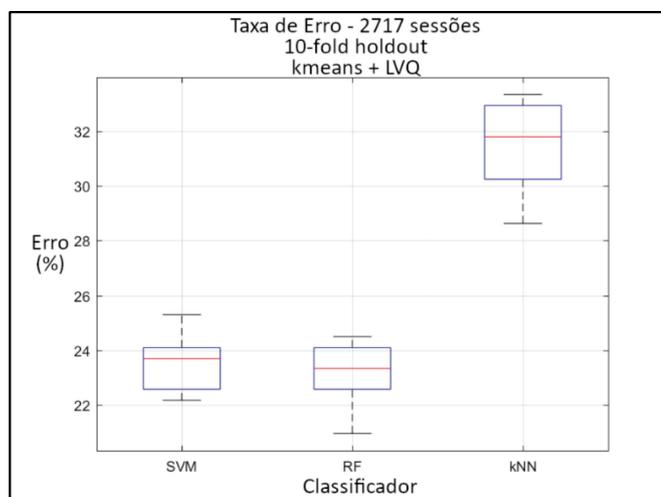
Tabela 7. Comparação da taxa de erro para 2717 sessões de treino, 8 e 16 microestados,

agrupamento com *k-means* e ajuste LVQ

	<i>k-means</i> com ajuste LVQ			
	8 microestados		16 microestados	
	$\mu_E$	$\sigma_E$	$\mu_E$	$\sigma_E$
SVM	[4] 24,64%	$\pm 0,80\%$	[13] 23,55%	$\pm 0,92\%$
RF	[5] 21,38%	$\pm 0,72\%$	[14] 23,19%	$\pm 1,09\%$
kNN	[6] 29,35%	$\pm 0,67\%$	[15] 31,52%	$\pm 1,57\%$

Figura 19. *Boxplot* do erro de classificação para 2717 sessões de treino e agrupamentocom *k-means* e LVQ, para  $k = 8$  microestados

Fonte: o autor.

Figura 20. *Boxplot* do erro de classificação para 2717 sessões de treino e agrupamentocom *k-means* e LVQ, para  $k = 16$  microestados

Fonte: o autor.

Novamente, para verificar a relevância dos resultados obtidos com os testes utilizando todas as amostras, foi executado o teste estatístico de Friedman (ver Seção 2.5.7). Os testes de Friedman confirmam a Hipótese de que para  $k = 8$  e para  $k = 16$ , há configurações com um *rank* médio estatisticamente superior aos demais. Utilizando, por fim o melhor resultado para 8 microestados e o melhor resultado para 16 microestados, foi executado o teste de Wilcoxon (ver Seção 2.5.7). O teste aponta que a configuração [5] é estatisticamente superior à configuração [14], com  $p\text{-value} = 0,0024$ .

#### 4.5. Comparativo com os resultados da literatura

Para a comparação dos classificadores kNN e RF de Lopez (2017) com o método proposto neste trabalho, foi executado inicialmente um teste para os classificadores SVM, RF e kNN com Otimização Bayesiana usando validação cruzada *10-fold* para definição dos hiperparâmetros e 10 execuções de *holdout* sobre um subconjunto estratificado de dados extraído aleatoriamente das sessões de treino, do mesmo tamanho daquele utilizado no trabalho de Lopez (2017) para os classificadores citados, no denominado *short set*, a saber 162 sessões de treino, sendo 80 anormais e 82 normais, bem como 106 sessões de teste, sendo 55 anormais e 51 normais.

Muito importante destacar que Lopez (2017) não informa quais sessões foram utilizadas para treino e teste, por isto que em cada execução do *holdout* foram utilizadas sessões diferentes. Os resultados obtidos por Lopez (2017) e por este trabalho no referido teste são apresentados na Tabela 8.

Tabela 8. Comparativo de resultados para 162 sessões de treino (*short set*)

	Este trabalho			(LOPEZ, 2017)
	k-means com ajuste LVQ			<i>short set</i>
	Média	Mediana	Desvio	Erro
SVM	37,74%	37,89%	$\pm 1,68\%$	-
RF	36,79%	37,50%	$\pm 1,65\%$	31,70%
kNN	46,22%	46,32%	$\pm 1,99\%$	41,80%

---

É importante ressaltar que a autora apresentou somente uma taxa de erro de classificação de uma única execução de sua metodologia para cada classificador, de forma que não foi possível executar algum teste estatístico para comparar os resultados.

Avaliando apenas as médias deste trabalho em comparação com a taxa de erro do trabalho de Lopez (2017), verifica-se que os classificadores RF e kNN obtiveram resultados um pouco piores, quando considerada essa mesma quantidade de sessões daquela utilizada pela autora no *short set* (162 de treino e 106 de teste). Esse teste foi executado no trabalho da autora para determinar um valor de base para comparação com os demais testes efetuados posteriormente com outros classificadores em seu trabalho, sobre a base de dados toda.

Na etapa preliminar desta pesquisa (KÜSTER; CIARELLI, 2021), foi verificado que a taxa de erro diminuiu, consideravelmente, utilizando a metodologia apresentada com um número maior de sessões utilizadas no treino dos classificadores. Dessa forma, se for retomado o resultado de Küster e Ciarelli (2021) para 500 sessões de treino, por exemplo, as taxas de erro dos classificadores RF e kNN já são reduzidas para 27,78% e 26,98%, respectivamente. Essa mesma constatação foi feita por Lopez (2017), que afirma que os modelos mais complexos, que serão brevemente descritos a seguir neste texto, também melhoravam as taxas de erro com o aumento do número de sessões de treino.

Nesse mesmo sentido, passa-se a investigar os resultados obtidos pela autora Lopez (2017) quando consideradas todas as sessões da base. Todos os resultados de Lopez (2017) e deste trabalho estão elencados na Tabela 9, sendo que os diferentes classificadores estão ordenados de forma decrescente em relação às taxas de erro obtidas. Analisando os resultados para a base toda, é possível verificar que os resultados obtidos neste trabalho superam alguns dos classificadores da literatura.

Tabela 9. Sumário de taxa de erro de classificação para todos os dados da base TUAB

Classificador	Trabalho	Erro (%)	$\sigma_E$ (%)
PCA-HMM	Lopez, (2017)	32,60	–
kNN ( <i>k-means</i> + LVQ, $k = 8$ )	Este trabalho	29,35	$\pm 0,67$
GMM-HMM	Lopez, (2017)	26,10	–
SVM ( <i>k-means</i> + LVQ, $k = 8$ )	Este trabalho	24,64	$\pm 0,80$
<i>Epoch-Based</i> HMM- <i>Majority</i>	Lopez, (2017)	24,60	–
<i>Vote Epoch-Based</i> HMM-SdA	Lopez, (2017)	22,10	–
RF ( <i>k-means</i> + LVQ, $k = 8$ )	Este trabalho	21,38	$\pm 0,72$
CNN-MLP	Lopez, (2017)	21,20	–

No trabalho de Lopez (2017), os classificadores executados sobre todas as sessões da base utilizaram os atributos: *Mel Frequency Cepstral Coefficients (MFCC)*, Energia, Energia diferencial e 1<sup>a</sup> e 2<sup>a</sup> derivadas dessas características. A dimensão do vetor de características foi reduzida por meio de PCA (*Principal Component Analysis*), sendo a taxa de erro de cada configuração referente a uma única execução feita em um *holdout* aplicado sobre a base toda para definição de sessões de treino e teste.

A autora Lopez (2017) informa que, para os sistemas baseado em HMM (*Hidden Markov Models*), ou seja, PCA-HMM, GMM-HMM e *Epoch-Based HMM-Majority* e *Vote Epoch-Based HMM-SdA*, foram utilizados 10 minutos de cada sessões de treino da base, para otimização dos hiperparâmetros. Uma vez que estes foram definidos, os testes foram executados sobre trechos entre 5 e 25 minutos de duração para cada sessão de EEG. Enfim, os melhores resultados foram obtidos para uma quantidade de processamento sequencial de 10 minutos para cada sessão de EEG, sendo os resultados de taxa de erro de classificação executados com todas as sessões de teste da base de dados.

Nesses sistemas, são processadas diversas janelas (ou épocas) de cada sessão da base, e a sessão é assinalada à classe mais votada. Na nomenclatura utilizada pela autora, *Majority* se refere à maioria de votos em determinada classe. O SdA (*Stacked denoising Autoencoder*) é utilizado numa etapa de pós-processamento, antes da contabilização final de votos, integralizando a maioria de votos em subperíodos da sessão maiores que as

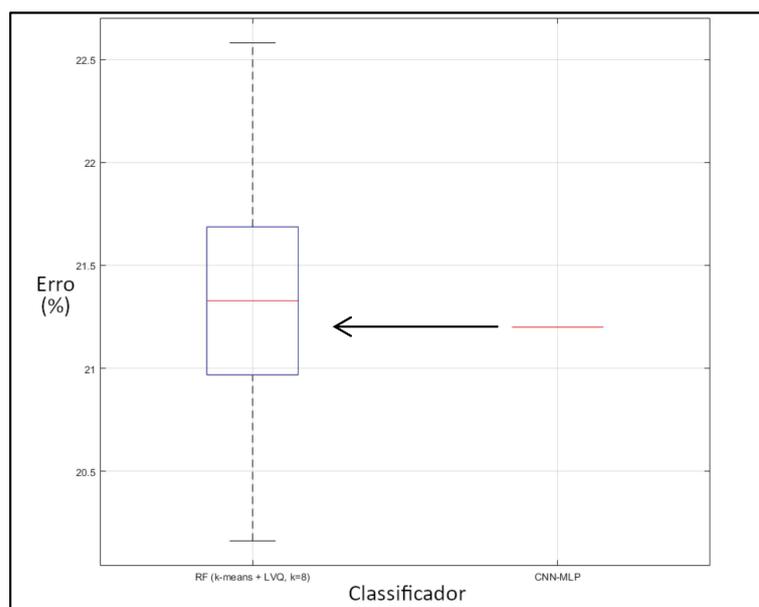
janelas processadas inicialmente pelos classificadores baseados em HMM. O SdA é utilizado devido sua capacidade de tratar sinais ruidosos e não-ideais (LOPEZ, 2017).

O classificador RF do presente trabalho foi capaz de superar todos os classificadores baseados em HMM do trabalho de Lopez (2017), como pode-se verificar na Tabela 9. Esse é o primeiro resultado favorável ao uso de análise de microestados para discriminação de anormalidades nos sinais de EEG.

O melhor sistema desenvolvido no trabalho de Lopez (2017), a saber o CNN-MLP, utilizou os mesmos atributos citados pela autora, já informados acima, extraídos individualmente de cada janela do sinal de EEG. Segundo a autora, o sistema é capaz de explorar melhor as características temporais do sinal de EEG, mas o uso de canais múltiplos também permite a inferência de correlações espaciais. O sistema treinado com gradiente estocástico descendente e Adam (*Adaptive Moment Estimation*) alcançou os melhores resultados ao processar as sessões usando todos os canais da montagem TCP/AR, com janelas de 7 segundos e com 3 camadas convolucionais conectadas à rede MLP que opera como um classificador da saída da última camada convolucional.

Figura 21. *Boxplot* do erro de classificação do melhor resultado deste trabalho e erro do

melhor resultado de Lopez (2017)



Fonte: o autor.

---

Os classificadores RF deste trabalho e CNN-MLP de Lopez (2017), para todas as amostras da base TUAB, obtiveram taxas de erro bastante parecidas, como pode ser observado no *boxplot* da Figura 21. Novamente, como a autora informa apenas uma taxa de erro de classificação, não é possível realizar um teste estatístico para comparar os resultados.

O erro do melhor classificador deste trabalho, a saber, o classificador *Random Forest*, com  $k = 8$  microestados e agrupamento *k-means* modificado com ajuste LVQ, utilizando 10 segundos de cada sessão da base, conforme elencado na Tabela 7, foi de  $21,38\% \pm 0,72\%$  que, quando comparado com o melhor resultado informado por Lopez (2017), de 21,20% para o classificador CNN-MLP, levando-se em consideração o desvio padrão, podem ser considerados muito próximos.

Uma vez que a metodologia utilizada no presente trabalho para discriminação dos sinais de EEG é muito diferente daquela implementada no trabalho de Lopez (2017), os resultados aqui obtidos corroboram também com a conclusão da autora, de que técnicas modernas de aprendizado de máquinas são capazes de modelar e discriminar anormalidades no sinal de EEG com uma taxa de erro de cerca de 21,20%.

Os resultados obtidos no presente trabalho corroboram também com a hipótese de que a análise de microestados pode ser útil para discriminação dos sinais anormais de EEG, superando alguns dos classificadores encontrados na literatura, conforme pode-se verificar na Tabela 9.

## 5. CONCLUSÃO E TRABALHOS FUTUROS

Neste trabalho foi proposto o uso da análise de microestados de EEG para detectar sinais anormais, em contraponto aos métodos tradicionais. Também foi proposto o uso da rede LVQ para obter de forma supervisionada os protótipos de microestados, diferente dos métodos comumente utilizados que são não-supervisionados.

Com base nos resultados obtidos é possível inferir que:

- i) A análise de microestados, que utiliza as características topológicas de um sinal de EEG, pode ser utilizada para discriminar sinais normais de sinais anormais de EEG, sem definição de alguma neuropatologia *a priori*;
- ii) É possível ajustar os protótipos de microestados obtidos pelo método clássico *k-means* modificado, por meio de um algoritmo supervisionado (neste trabalho sendo usada a rede LVQ), possibilitando uma ligeira melhora na discriminação das diferentes classes atribuídas a cada sinal de EEG;
- iii) É viável obter os protótipos de microestados por meio da rede LVQ, entretanto, a agrupamento utilizando somente o LVQ alcançou resultados um pouco piores que o *k-means* modificado sozinho.

Comparando as taxas de erro de classificação dos melhores resultados da literatura, de 21,20% (LOPEZ, 2017) com os melhores resultados do presente trabalho, a saber,  $21,38 \pm 0,72\%$ , é possível afirmar que são da mesma ordem, apesar de utilizarem metodologias completamente diferentes para discriminação e classificação dos sinais de EEG.

Isso sugere que a análise de microestados é um método promissor na tarefa de detecção de sinais anormais de EEG como etapa inicial do processo de identificação automática de distúrbios neurológicos, entretanto, é necessário alcançar taxas de erro ainda bastante inferiores a essas para uso clínico, uma vez que os neurologistas sugerem que, para

---

monitoramento em tempo real num ambiente clínico, é necessário uma sensibilidade de pelo menos 95% e uma especificidade menor que 5% (GOLMOHAMMADI *et al.*, 2019). Apesar de ainda não ter alcançado os parâmetros necessários para uso clínico, a detecção de sinais de EEG anormais por meio do uso de microestados se mostra vantajoso na comparação com as técnicas apresentadas na literatura, uma vez que foi possível alcançar taxas de erro da mesma ordem da literatura, em que se usam classificadores mais complexos e que demandam custo computacional muito maior, necessitando da escolha de canais específicos para extração das características temporais e espectrais, classicamente utilizadas, bem como da redução de dimensionalidade do vetor de características para viabilizar a análise.

Além de obter resultados muito semelhantes utilizando classificadores mais simples, foi possível obter tais resultados com trechos consideravelmente pequenos, de apenas 10 segundos de sinal, sendo que a maioria dos classificadores da literatura utilizam vários minutos de sinal para realizar a classificação. Isso demonstra a grande capacidade dos microestados de discriminar os sinais de EEG, capturando muito bem sua variância global.

Apesar da rede LVQ utilizada ter apresentado uma melhoria estatística na discriminação dos sinais ao ser utilizada para ajuste dos protótipos de microestado obtidos inicialmente pelo *k-means* modificado, quando aplicada sozinha para obtenção dos protótipos, ao contrário do que seria esperado, ela obteve um resultado um pouco inferior ao resultado do *k-means* modificado sozinho.

Uma possível explicação para este fato é a subutilização de alguns dos *clusters* no processo de treino da rede. Isso ocorre devido à atualização apenas do centroide do *cluster* vencedor à cada nova iteração durante o treino, de forma que pesos inicializados com valores aleatórios podem cair muito longe da região de interesse, não sendo atualizados nunca e ficando subutilizados.

---

Nesse sentido, podem-se propor algumas sugestões para trabalhos futuros, como o teste com classificadores que utilizem outras abordagens de aprendizado de máquina, baseadas em redes convolucionais e redes de aprendizado profundo, que são técnicas promissoras a serem investigadas associadas à análise de microestados para melhoria dos resultados de classificação.

Uma outra área de investigação é a extração de mais características dos microestados, que podem ser úteis para auxiliar na melhor discriminação de sinais anormais de EEG. Também podem ser avaliadas técnicas de detecção de anomalias, na qual os algoritmos são treinados somente com dados de normalidade e identificam o desvio da condição normal. Algoritmos como *Autoencoder* e *One Class SVM* são bons candidatos para serem avaliados.

Como os microestados são capazes de capturar a dinâmica espacial da rede neuronal e sua evolução temporal, eles possuem informações relevantes de contexto quando correlatados no tempo. Nesse sentido, um campo a ser investigado no processo de identificação de sinais anormais de EEG é o encadeamento de amostras por meio de técnicas de regressão de dados.

É necessária, também, a investigação da eventual subutilização no processo de treino da rede LVQ, quando esta é utilizada na etapa de obtenção dos microestados. Nesse sentido, pode-se testar técnicas de inicialização dos pesos da rede, como as elencadas por Du (2010) e por Blachnik e Duch (2011), por exemplo.

Uma outra sugestão, é utilizar uma quantidade maior de neurônios na rede LVQ, modificando a função de treinamento para atualizar não só o neurônio vencedor, mas também sua vizinhança, por meio de uma função-peso, como uma função gaussiana, por exemplo. A seguir, avaliam-se os neurônios ligados ao mesmo rótulo, para agrupá-los. Essa metodologia possibilita a utilização de um número maior de protótipos de microestados, sendo necessário um procedimento posterior para selecionar os protótipos mais relevantes.

---

Por fim, o teste de outros algoritmos para a obtenção dos protótipos de microestados também se impõe, uma vez que a rede do tipo LVQ depende do conhecimento das classes dos sinais de treino para obtenção dos *clusters*. Isso pode ser inviável em algumas aplicações que pretendam utilizar a análise de microestados de EEG. Um candidato promissor nesta tarefa são os mapas auto organizáveis de Kohonen (1995).

## REFERÊNCIAS BIBLIOGRÁFICA

- BABLANI, A.; EDLA, D. R.; DODIA, S. Classification of EEG data using k-nearest neighbor approach for concealed information test. *Procedia Computer Science*, v. 143, p. 242–249, 2018. Disponível em: <<https://doi.org/10.1016/j.procs.2018.10.392>>.
- BENAVOLI, A.; CORANI, G.; MANGILI, F. Should we really use post-hoc tests based on mean-ranks? *Journal of Machine Learning Research*, v. 17, p. 1–10, 2016.
- BLACHNIK, M.; DUCH, W. LVQ algorithm with instance weighting for generation of prototype-based rules. *Neural Networks*, v. 24, n. 8, p. 824–830, 2011. Disponível em: <<http://dx.doi.org/10.1016/j.neunet.2011.05.013>>.
- BREIMAN, L. Random Forests. *Machine Learning*, v. 45, p. 5–32, 2001. Disponível em: <<https://doi.org/10.1023/A:1010933404324>>.
- CROCE, P.; QUERCIA, A.; COSTA, S.; ZAPPASODI, F. EEG microstates associated with intra- and inter-subject alpha variability. p. 1–11, 2020.
- CUSTO, A.; VAN DE VILLE, D.; WELLS, W. M.; TOMESCU, M. I.; BRUNET, D.; MICHEL, C. M. Electroencephalographic Resting-State Networks: Source Localization of Microstates. *Brain Connectivity*, v. 7, n. 10, p. 671–682, 2017.
- D’CROZ-BARON, D. F.; BAKER, M.; MICHEL, C. M.; KARP, T. EEG microstates analysis in young adults with autism spectrum disorder during resting-state. *Frontiers in Human Neuroscience*, v. 13, n. June, p. 1–9, 1 fev. 2019.
- DELORME, A.; MAKEIG, S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. v. 134, p. 9–21, 2004.
- DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. *Journal of Machine Learning Research*, v. 7, p. 1–30, 2006.
- DONGORAN, A.; RAHMADANI, S.; ZARLIS, M.; ZAKARIAS. Feature weighting using particle swarm optimization for learning vector quantization classifier. *Journal of Physics: Conference Series*, v. 978, n. 1, 2018.
- DU, K. L. Clustering: A neural network approach. *Neural Networks*, v. 23, n. 1, p. 89–107, 2010.
- EBERSOLE, J. S. *Current Practice of Clinical Electroencephalography*. 4. ed. Philadelphia: Wolters KluWer Health, 2014.

---

FACELI, K. *Inteligência artificial: uma abordagem de aprendizado de máquina*. Rio de Janeiro, RJ: LTC, 2011.

FERRELL, S.; MATHEW, V.; REFFORD, M.; TCHIONG, V.; AHSAN, T.; OBEID, I.; PICONE, J. *The Temple University Hospital EEG Corpus: Electrode Location and Channel Labels*. [s.l.: s.n.]. Disponível em: <par.nsf.gov/servlets/purl/10199699>.

FINNIGAN, S.; PUTTEN, M. J. A. M. van. EEG in ischaemic stroke: Quantitative EEG can uniquely inform (sub-)acute prognoses and clinical management. *Clinical Neurophysiology*, v. 124, n. 1, p. 10–19, 2013. Disponível em: <<http://dx.doi.org/10.1016/j.clinph.2012.07.003>>.

FU, Y.; CHEN, J.; XIONG, X. Calculation and Analysis of Microstate Related to Variation in Executed and Imagined Movement of Force of Hand Clenching. *Computational Intelligence and Neuroscience*, v. 2018, 2018.

GEMEIN, L. A. W.; SCHIRRMEISTER, R. T.; CHRABAŚCZCZ, P.; WILSON, D.; BOEDECKER, J.; SCHULZE-BONHAGE, A.; HUTTER, F.; BALL, T. Machine-learning-based diagnostics of EEG pathology. *NeuroImage*, v. 220, n. May, 2020.

GOLMOHAMMADI, M.; HARATI NEJAD TORBATI, A. H.; LOPEZ DE DIEGO, S.; OBEID, I.; PICONE, J. Automatic analysis of EEGs using big data and hybrid deep learning architectures. *Frontiers in Human Neuroscience*, v. 13, n. March, p. 1–14, 2019.

GUYON, I.; ELISSEEFF, A. An Introduction to Variable and Feature Selection Isabelle. *Journal of Machine Learning Research*, v. 1, n. 3, p. 1157–1182, 2003. Disponível em: <[http://www.crossref.org/deleted\\_DOI.html](http://www.crossref.org/deleted_DOI.html)>.

HARATI, A.; LOPEZ, S.; OBEID, I.; PICONE, J.; JACOBSON, M. P.; TOBOCHNIK, S. The TUH EEG CORPUS: A big data resource for automated EEG interpretation. In: 2014 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), *Anais...IEEE*, dez. 2014. Disponível em: <<http://ieeexplore.ieee.org/document/7002953/>>.

HEINE, L.; SODDU, A.; GÓMEZ, F.; VANHAUDENHUYSE, A.; TSHIBANDA, L.; THONNARD, M.; CHARLAND-VERVILLE, V.; KIRSCH, M.; LAUREYS, S.; DEMERTZI, A. Resting state networks and consciousness Alterations of multiple resting state network connectivity in physiological, pharmacological, and pathological consciousness states. *Frontiers in Psychology*, v. 3, n. AUG, p. 1–12, 2012.

HOLLANDER, M.; WOLFE, D. A. *Nonparametric Statistical Methods*. 3. ed. [s.l.] John Wiley & Sons, 2013.

- 
- HU, L.; ZHANG, Z. *EEG signal processing and feature extraction*. [s.l: s.n.]
- JIANG, Z.; HE, J.; GUO, P. Feature data optimization with LVQ technique in semantic image annotation. *Proceedings of the 2010 10th International Conference on Intelligent Systems Design and Applications, ISDA'10*, p. 906–911, 2010.
- KANDEL, E. R. *PRINCIPLES OF NEURAL SCIENCE*. 5. ed. New York: Mc Graw Hill, 2013.
- KHANNA, A.; PASCUAL-LEONE, A.; FARZAN, F. Reliability of resting-state microstate features in electroencephalography. *PLoS ONE*, v. 9, n. 12, p. 1–21, 2014.
- KHANNA, A.; PASCUAL-LEONE, A.; MICHEL, C. M.; FARZAN, F. Microstates in resting-state EEG: Current status and future directions. *Neuroscience and Biobehavioral Reviews*, v. 49, p. 105–113, 2015.
- KOENIG, T.; PRICHEP, L.; LEHMANN, D.; SOSA, P. V.; BRAEKER, E.; KLEINLOGEL, H.; ISENHART, R.; JOHN, E. R. Millisecond by millisecond, year by year: Normative EEG microstates and developmental stages. *NeuroImage*, v. 16, n. 1, p. 41–48, 2002.
- KOHONEN, T. *Learning Vector Quantization*. Berlin: Springer, 1995.
- KÜSTER, D. W.; CIARELLI, P. M. Classificação Automática de Sinais Anormais de EEG por meio de Microestados e Aprendizado de Máquina. In: *Anais...2021*. Disponível em:  
<[https://www.sba.org.br/open\\_journal\\_systems/index.php/sbai/article/view/2797](https://www.sba.org.br/open_journal_systems/index.php/sbai/article/view/2797)>.
- LEHMANN, D.; OZAKI, H.; PAL, I. EEG alpha map series: brain micro-states by space-oriented adaptive segmentation. *Electroencephalography and Clinical Neurophysiology*, v. 67, n. 3, p. 271–288, 1987.
- LOPEZ, S. *Automated Interpretation of Abnormal Adult Electroencephalograms*. 2017. Temple University, 2017. Disponível em:  
<[www.isip.piconepress.com/publications/ms\\_theses/2017/abnormal/thesis/](http://www.isip.piconepress.com/publications/ms_theses/2017/abnormal/thesis/)>.
- LOPEZ, S.; SUAREZ, G.; JUNGREIS, D.; OBEID, I.; PICONE, J. Automated Identification of Abnormal Adult EEGs. *Physiology & behavior*, v. 176, n. 12, p. 139–148, 2016.
- M. ALVES, L.; F. CÔCO, K.; L. DE SOUZA, M.; M. CIARELLI, P. Graph Theory Analysis of Microstates in Attention-Deficit Hyperactivity Disorder. In: Anais do Congresso Brasileiro de Automática 2020, *Anais...sbabra*, 7 dez. 2020. Disponível em:  
<[https://www.sba.org.br/open\\_journal\\_systems/index.php/sba/article/view/1481](https://www.sba.org.br/open_journal_systems/index.php/sba/article/view/1481)>.

---

MATHWORK. *Improve Shallow Neural Network Generalization and Avoid Overfitting: Multiple Neural Networks*. Disponível em: <<https://www.mathworks.com/Help/DeepLearning/Ug/Improve-Neural-Network-Generalization-and-Avoid-Overfitting.Html#Bt0Cnqo-1>>. Acesso em: 1 fev. 2022.

MCMACKIN, R.; MUTHURAMAN, M.; GROPPA, S.; BABILONI, C.; TAYLOR, J. P.; KIERNAN, M. C.; NASSEROLESLAMI, B.; HARDIMAN, O. Measuring network disruption in neurodegenerative diseases: New approaches using signal analysis. *Journal of Neurology, Neurosurgery and Psychiatry*, v. 90, n. 9, p. 1011–1020, 2019.

MELIN, P.; AMEZCUA, J.; VALDEZ, F.; CASTILLO, O. A new neural network model based on the LVQ algorithm for multi-class classification of arrhythmias. *Information Sciences*, v. 279, p. 483–497, 2014. Disponível em: <<http://dx.doi.org/10.1016/j.ins.2014.04.003>>.

MICHEL, C. M.; KOENIG, T. EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review. *NeuroImage*, v. 180, n. December 2017, p. 577–593, 2018. Disponível em: <<https://doi.org/10.1016/j.neuroimage.2017.11.062>>.

MINISTÉRIO DA SAÚDE. Principais causas de morte no Brasil. v. 2016, p. 3–6, 2018. Disponível em: <<http://svs.aids.gov.br/dantps/centrais-de-conteudos/paineis-de-monitoramento/mortalidade/gbd-brasil/principais-causas/>>.

MONDINI, V.; MANGIA, A. L.; CAPPELLO, A. Single-session tDCS over the dominant hemisphere affects contralateral spectral EEG power, but does not enhance neurofeedback-guided event-related desynchronization of the non-dominant hemisphere's sensorimotor rhythm. *PLoS ONE*, v. 13, n. 3, p. 1–21, 2018.

MONTALVÃO, J.; MIRANDA, L.; DORIZZI, B. Straightforward Working Principles Behind Modern Data Visualization Approaches. v. 9, 2021.

MURRAY, M. M.; BRUNET, D.; MICHEL, C. M. Topographic ERP analyses: A step-by-step tutorial review. *Brain Topography*, v. 20, n. 4, p. 249–264, 2008.

MUSSO, F.; BRINKMEYER, J.; MOBASCHER, A.; WARBRICK, T.; WINTERER, G. Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks. *NeuroImage*, v. 52, n. 4, p. 1149–1161, 2010. Disponível em: <<http://dx.doi.org/10.1016/j.neuroimage.2010.01.093>>.

NIAZY, R. K.; XIE, J.; MILLER, K.; BECKMANN, C. F.; SMITH, S. M. *Spectral characteristics of resting state networks*. 1. ed. [s.l.] Elsevier B.V., 2011. v. 193

---

OBEID, I.; PICONE, J. The temple university hospital EEG data corpus. *Frontiers in Neuroscience*, v. 10, n. MAY, 2016.

OBEID, I.; PICONE, J. Machine Learning Approaches to Automatic Interpretation of EEGs. *Signal Processing and Machine Learning for Biomedical Big Data*, p. 271–300, 16 nov. 2018.

OKAMOTO, M.; DAN, H.; SAKAMOTO, K.; TAKEO, K.; SHIMIZU, K.; KOHNO, S.; ODA, I.; ISOBE, S.; SUZUKI, T.; KOHYAMA, K.; DAN, I. Three-dimensional probabilistic anatomical cranio-cerebral correlation via the international 10 – 20 system oriented for transcranial functional brain mapping. v. 21, p. 99–111, 2004.

POULSEN, A. T.; PEDRONI, A.; LANGER, N.; HANSEN, L. K. Microstate EEGlab toolbox: An introductory guide. *bioRxiv*, p. 289850, 2018.

SCIKIT-LEARN DEVELOPERS. *Cross-validation: evaluating estimator performance*. Disponível em: <[https://scikit-learn.org/stable/modules/cross\\_validation.html](https://scikit-learn.org/stable/modules/cross_validation.html)>.

SEITZMAN, B. A.; ABELL, M.; BARTLEY, S. C.; MOLLY, A.; BOLBECKER, A. R.; HETRICK, W. P.; SCIENCES, B.; HOSPITAL, D. C. M.; CARE, B. H.; TOWNSHIP, P.; ERICKSON, M. A.; BOLBECKER, A. R.; HETRICK, W. P. HHS Public Access. *NeuroImage*, v. 146, p. 533–543, 2018.

SHAH, V.; VON WELTIN, E.; LOPEZ, S.; MCHUGH, J. R.; VELOSO, L.; GOLMOHAMMADI, M.; OBEID, I.; PICONE, J. The temple university hospital seizure detection corpus. *Frontiers in Neuroinformatics*, v. 12, n. November, p. 1–6, 2018.

SNOEK, J.; LAROCHELLE, H.; ADAMS, R. P. Practical Bayesian Optimization of Machine Learning Algorithms. 13 jun. 2012. Disponível em: <<http://arxiv.org/abs/1206.2944>>.

ZAPPASODI, F.; CROCE, P.; GIORDANI, A.; ASSENZA, G.; GIANNANTONI, N. M.; PROFICE, P.; GRANATA, G.; ROSSINI, P. M.; TECCHIO, F. Prognostic Value of EEG Microstates in Acute Stroke. *Brain Topography*, v. 30, n. 5, p. 698–710, 2017.

## APÊNDICE A

### A.1 Código

Os *scripts* implementados em plataforma Matlab®, versão 2017a, EEGLab versão 2020 (DELORME; MAKEIG, 2004) e Microstate Toolbox versão 1.0 (POULSEN *et al.*, 2018), estão disponíveis no GitHub. O repositório pode ser acessado através do QR CODE abaixo e do link <<https://github.com/davidwkuster/AbClass>>.

