

## PROJETO DE CLASSIFICADOR AUTOMÁTICO DE FASES DO SONO EM ELETROENCEFALOGRAMA<sup>1</sup>

Thiago Hideo Xavier KODAMA<sup>2</sup>  
Graduando em Engenharia Eletrônica  
IFSP/ *Campus* São Paulo

Raul de França SILVA<sup>3</sup>  
Graduando em Engenharia Eletrônica  
IFSP/ *Campus* São Paulo

Ricardo PIRES<sup>4</sup>  
Doutor em Sistemas Automáticos e  
Microeletrônicos – Université Montpellier II  
Docente no IFSP/ *Campus* São Paulo

### RESUMO

O sono é importante para o bom funcionamento do corpo humano. Alguns problemas causados pela sua falta são distúrbios psiquiátricos, déficits cognitivos, problemas de saúde e acidentes de trânsito. Ele pode ser classificado em dois tipos: sono REM (*rapid eye movement*) e sono Não-REM, sendo que este último divide-se em quatro estágios. Uma forma de se estudá-lo é realizar a polissonografia, que capta dados como eletroencefalograma (EEG), movimento dos olhos e espasmos musculares enquanto o paciente está dormindo. A classificação automática, que é o objetivo deste trabalho, pode vir a auxiliar no tratamento de distúrbios do sono, como insônia, apnéia e síndrome das pernas inquietas. Para construir um classificador, foram utilizados sinais de EEG do banco de dados do repositório Physionet, que foram processados aplicando a transformada de Fourier e divididos em sub-bandas de frequência, sendo elas: Delta, Theta, Alpha, Beta e Gamma. Os classificadores utilizados foram árvores de decisão, máquinas de vetores de suporte e redes neurais. Foram alcançados resultados de 60% de acurácia para um classificador de 6 classes. Evidenciou-se a importância de técnicas de pré-processamento para trabalhar com o EEG e uma possível aplicação em discernir um estágio específico de sono.

**Palavras-chave:** Sono; Eletroencefalograma; Inteligência Artificial; Classificação; Estágios do sono.

### *PROJECT OF AN AUTOMATIC SLEEP PHASE CLASSIFIER IN ELECTROENCEPHALOGRAM*

### ABSTRACT

Sleep is important for the proper functioning of the human body. Some problems caused by its lack are psychiatric disorders, cognitive deficits, health problems, and traffic accidents. It can be classified into two types: REM (rapid eye movement) sleep and non-REM sleep, the latter being divided into four stages. One way to study it is to perform a polysomnography, which captures data such as the electroencephalogram (EEG), eye movements, and muscle spasms while the patient is asleep. The automatic classification, which is the objective of this work, may help in the treatment of sleep disorders such as insomnia, apnea and restless legs syndrome. To build a classifier, EEG signals from the Physionet repository database were used, which were processed by applying the Fourier transform and divided into frequency sub-bands, being: Delta, Theta, Alpha, Beta and Gamma. The classifiers used were decision trees, support vector machines and neural networks. Results of 60% accuracy were achieved for a 6-class classifier. It became evident

the importance of pre-processing techniques to work with EEG and a possible application in discerning a specific sleep stage.

**Keywords:** Sleep; Electroencephalography; Artificial Intelligence; Classification; Sleep stage.

## Introdução

O sono contribui para o bom funcionamento do corpo humano. A falta de sono pode prejudicar um indivíduo de diversas maneiras, como distúrbios psiquiátricos, déficits cognitivos, problemas de saúde e riscos de acidentes de trânsito (MÜLLER; GUIMARÃES, 2007). Portanto, é importante que se tenha sono de qualidade.

Segundo Rechtschaffen e Kales (1968), o sono é classificado em dois tipos, o sono REM (*Rapid Eye Movements*), formado por um único estágio, e o Não-REM, este último formado por quatro estágios:

Estágio 1 (N1): Esse é o estágio de transição entre vigília e sono. Possui as características de um sinal EEG (eletroencefalograma) de amplitude relativamente baixa.

Estágio 2 (N2): Esse é o estágio de sono mais leve, onde ocorrem o complexo-k e episódios de fusos de sono. O complexo-k tem como característica uma componente negativa de EEG de alta amplitude seguida de uma componente positiva. O fuso de sono são ondas de duração curta que aumentam e diminuem sua amplitude.

Estágio 3 (N3): Esse é o começo do sono profundo. Ele possui a característica de ter ondas de alta amplitude, com frequência mais baixa.

Estágio 4 (N4): Esse é o sono profundo efetivamente. As características dele são muito semelhantes às do estágio anterior.

Também há o termo vigília, que é empregado para se referir ao estado em que um indivíduo está acordado.

O estudo de sono pode ser feito por meio de uma polissonografia, em que são captados alguns dados do paciente, como sinais cerebrais, movimento dos olhos e espasmos dos músculos durante o sono (MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH, 2020).

A atividade elétrica do cérebro é medida na forma do chamado eletroencefalograma (EEG), usando-se, normalmente, mais de 20 eletrodos metálicos (SEECK et al., 2017). Usualmente, é empregado o chamado sistema 10-20 de posicionamento de eletrodos na superfície do crânio, sistema este padronizado internacionalmente e descrito por Malmivuo e Plonsey (1995). Aquela atividade elétrica

deve-se à comunicação entre células cerebrais por meio de impulsos elétricos. O EEG, portanto, manifesta-se como ondas ao longo do tempo, uma para cada eletrodo. Cada um desses eletrodos e sua onda correspondente é chamado de canal. O EEG é usado no diagnóstico de distúrbios cerebrais (MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH, 2022).

A classificação automática do estágio de sono para fins de monitoramento pode auxiliar na agilidade do estudo da qualidade de sono de um paciente, visto que o estudo de informações obtidas de uma polissonografia receberia uma forma de automação, e facilitaria determinar quando algum indivíduo não está em plenas faculdades mentais devido a falta de sono, sem a necessidade da coleta e análise dos dados de um paciente com métodos apenas manuais por um especialista.

Como um exemplo da utilidade do monitoramento do sono, poder-se-ia determinar que um motorista de caminhão não está apto para dirigir em um determinado momento, visto que o sono é uma das principais condições de saúde que mais motivam acidentes de trânsito nas rodovias brasileiras, segundo levantamento da Abramet. Este fator causou cerca de 22,6 mil acidentes registrados nas rodovias no Brasil, causando 2.092 mortes e deixando 22.645 feridos, entre 2014 e julho de 2020 (ABRAMET, 2021).

A classificação automática também pode auxiliar no tratamento de distúrbio do sono. Os distúrbios mais comuns são insônia, apneia e a síndrome das pernas inquietas. Pacientes com esses distúrbios podem desenvolver doenças nas artérias, infarto do miocárdio, acidente vascular cerebral, obesidade, disfunção erétil entre outros problemas (MINISTÉRIO DA SAÚDE, 2012).

### **Estudo Bibliográfico**

Em um estudo por Malafeev et al. (2018), os autores realizaram a automatização da classificação de estágios de sono por aprendizado de máquina, seguindo três possibilidades de algoritmos: *Random Forest*, *Long Short-Term Memory* (LSTM), baseada em características e, por fim, uma *Convolutional Neural Network Long Short-Term Memory*, ou CNN LSTM, baseada em dados brutos. Por esse estudo, os autores obtiveram resultados similares positivos entre os diferentes algoritmos, embora as redes neurais tenham apresentado algumas vantagens em relação ao algoritmo *Random Forest* enquanto apresentavam uma convergência satisfatória. Para validar os resultados,

utilizaram uma métrica denominada Kappa de Cohen, observando resultados bons entre seus modelos na classificação dos estados de sono. Obtiveram valores de, em média, 0.8 na escala do Kappa de Cohen (que possui um máximo de 1), com exceção do estágio N1, o primeiro estágio do sono Não-REM, no qual observaram-se valores em torno de 0.4, fato que os autores interpretaram como bom resultado, dada a falta de consenso e de consistência na classificação deste estágio por seres humanos.

Em uma pesquisa realizada por Aboalayon et al. (2016), é apresentado um estudo de classificação do sono utilizando diversos classificadores, como Máquina de Vetores de Suporte (*Support-Vector Machine*, SVM), Árvore de Decisão (*Decision Tree*, DT) e Análise de Discriminante Linear (*Linear Discriminant Analysis*, LDA). Neste estudo, os autores propõem a utilização de apenas um canal EEG, que é decomposto em sub-bandas (delta, theta, alpha, beta e gamma). Para a extração de características, ele utiliza a MMD (*Maximum-Minimum Distance*) junto com a *EnergySis* (Esis). Após todos esses tratamentos, foram aplicados os classificadores. O classificador DT apresentou melhores resultados, com média de 93% de acurácia.

Em um trabalho realizado por Ferreira (2018), é apresentado um estudo que utiliza a técnica *multitaper* como pré-processamento e redes neurais como classificador, utilizando o mesmo banco de dados que o deste estudo, que será descrito adiante. Naquela pesquisa, foi obtida uma classificação com acurácia de 86%.

Naqueles artigos, observou-se o emprego de apenas cinco classes para a predição, ignorando-se o estágio N4 do Não-REM.

### **Fundamentação teórica**

No estudo bibliográfico, constatou-se que a identificação das fases do sono está relacionada ao espectro do EEG. Foi decidido, então, empregar-se a Transformada Rápida de Fourier (FFT) para extração de características dos dados brutos. Segundo Bracewell (2000), a Transformada de Fourier é uma ferramenta matemática para análise de sinais e sistemas, que permite representar um sinal do domínio do tempo como uma soma de senos e cossenos no domínio da frequência. Ela é amplamente utilizada em diversas áreas, incluindo processamento de sinais, comunicações, acústica, óptica e física.

A Transformada de Fourier Discreta (DFT) é uma versão discretizada da Transformada de Fourier, que pode ser implementada de forma eficiente utilizando um

algoritmo conhecido como FFT, uma técnica computacional importante que permite calcular a DFT em um tempo muito mais curto do que usando-se uma implementação computacional mais direta da definição da transformada (BRACEWELL, 2000).

Para classificar as características extraídas com a DFT, é necessário algum modelo de classificador. Para este trabalho, foram feitos experimentos com os classificadores: SVM, DT e MLP.

A Máquina de Vetores de Suporte (em inglês, *Support-Vector Machine*, SVM) (CORTES; VAPNIK, 1995) é um classificador que, em seu uso, passa por uma fase de aprendizado, em que lhe são apresentados exemplos de vetores de dados e lhe são informadas suas classes correspondentes (neste projeto, REM, estágio 1, estágio 2, estágio 3 e estágio 4). O aprendizado consiste na identificação de um hiperplano separador ótimo, num espaço multidimensional, entre os pontos que representam elementos de uma classe e os pontos de outras classes. Por separador ótimo, entende-se aquele que mantém a máxima margem possível entre pontos de uma classe, de um lado, e pontos das outras classes, do outro. Para casos em que não exista esse separador no espaço multidimensional original, todos os pontos são mapeados para um espaço com mais dimensões, usando-se uma função não linear, chamada *kernel*. (Há várias opções de funções *kernel*, como a polinomial e a de base radial.) Esse mapeamento tende a aumentar a probabilidade de se encontrar um hiperplano separador, desta vez no espaço com mais dimensões. Após o aprendizado, a SVM passa a inferir as classes às quais devem pertencer novos vetores que lhe forem apresentados. A inferência é feita verificando-se a posição dos novos vetores em relação ao hiperplano separador ótimo.

A Árvore de Decisão (em inglês, *Decision Tree*, DT) é um classificador simples de ser interpretado, por se assemelhar a uma estrutura de árvore para um fluxograma. Partindo-se da raiz, essa árvore tem seus nós internos representando características dos elementos a serem classificados e os testes realizados sobre eles. Em sequência a esses testes, há ramificações em seus nós, formando os seus galhos, que representam as diferentes decisões parciais que a DT utiliza para a classificação final. Ao fim da ramificação, há as folhas, que representam as classes ou valores preditos (HAN et al., 2011).

O MLP, ou Perceptron de Múltiplas Camadas (do inglês, *Multiple Layer Perceptron*), é um dos classificadores mais populares, sendo composto por camadas de

entrada e saída visíveis de neurônios que envolvem múltiplas camadas ocultas. Os neurônios do MLP possuem uma função de ativação não linear diferenciável. Durante uma fase de treino, pesos associados às entradas dos neurônios são ajustados por meio de uma propagação retroativa de erros, conhecida como retropropagação, com base em exemplos de dados conhecidos e rotulados (HAYKIN, 2009). Na fase de operação, para novos exemplos, espera-se que os valores às saídas do MLP possibilitem a sua classificação correta.

Para se realizar a análise dos resultados de um classificador, costuma ser utilizada a ferramenta matriz de confusão, a qual mapeia, após o treino e teste de um classificador, os casos verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN) para cada uma das características que compõem as classes. Com essas quatro informações de um determinado teste, é possível, por meio de simples operações matemáticas, obter a precisão, que é a métrica que avalia a quantidade de verdadeiros positivos sobre todos os resultados positivos, e a sensibilidade, uma métrica que avalia a capacidade da rede de detectar com sucesso os resultados positivos. Essas duas métricas são obtidas conforme as seguintes relações (MARIANO, 2021):

$$precisão = \frac{VP}{(VP+FP)} \quad (1)$$

$$sensibilidade = \frac{VP}{(VP+FN)} \quad (2)$$

Utilizando a acurácia e a sensibilidade, é possível calcular o *F1 score*, a média harmônica calculada com base na precisão e sensibilidade:

$$F1\ score = 2 * \frac{precisão * sensibilidade}{precisão+sensibilidade} \quad (3)$$

Portanto, o *F1 score* reflete bem a efetividade de um algoritmo de inteligência artificial, especialmente ao se lidar com bancos de dados que estejam desbalanceados, como o utilizado neste estudo.

## Metodologia

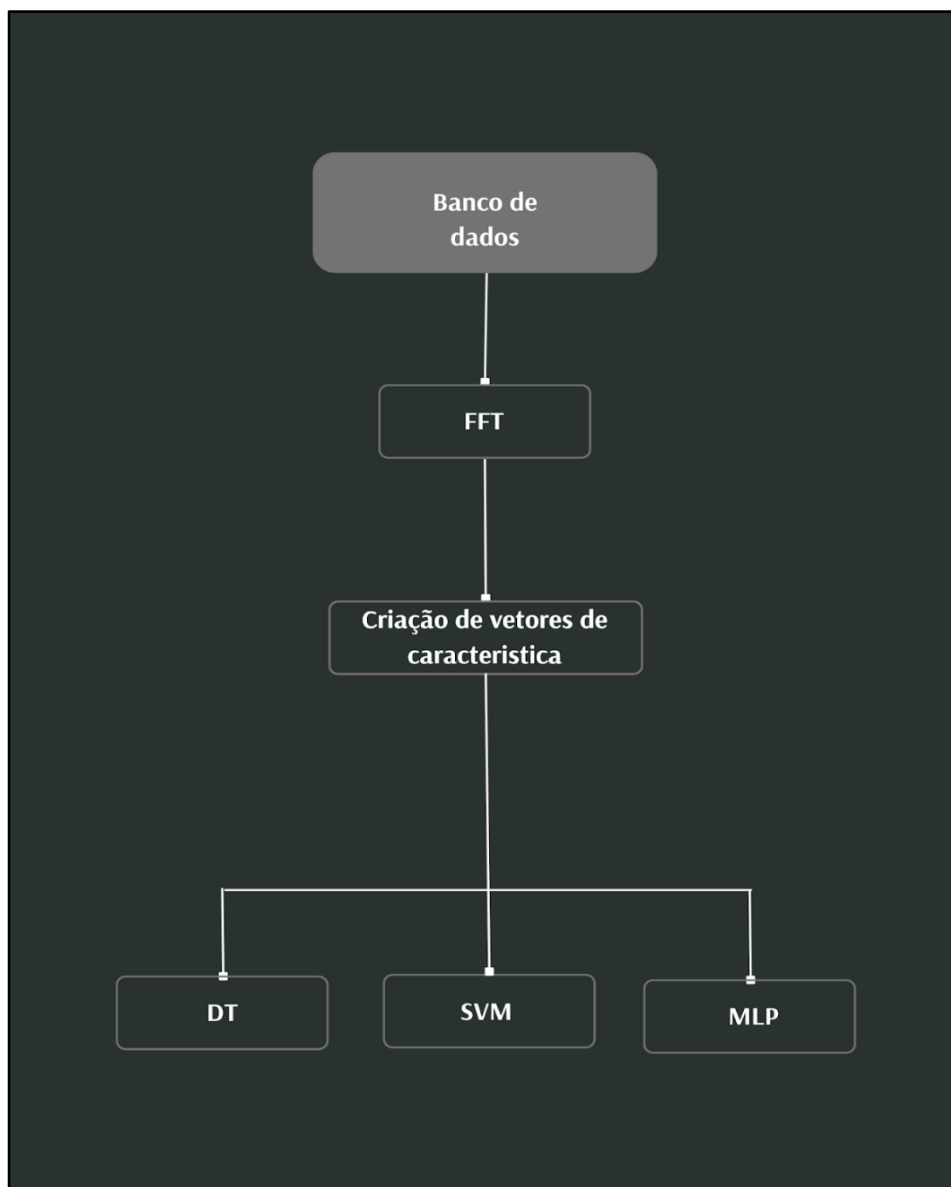
Os sinais EEG utilizados neste trabalho foram obtidos no banco de dados do repositório Physionet (2013) (KEMP, 2000). O banco é separado em arquivos para cada

paciente, cada um destes contando com dois conjuntos de dois arquivos para a primeira e para a segunda noite do estudo. O primeiro arquivo do conjunto contém os dados de EEG em arquivos nomeados no formato ST7 ssN J0-PSG.edf, em que ss é o número do paciente e N é o número da noite. O segundo arquivo tem as anotações para a respectiva polissonografia da noite e paciente do mesmo conjunto e contém os registros para aquela noite, informando qual o estágio de sono para um determinado momento do espectro do EEG, bem como a duração da ocorrência do estágio em segundos.

As polissonografias foram realizadas por técnicos, que classificaram os estágios do sono em seis: a vigília, classificada como W, de *waked* (acordado), estágio 1, estágio 2, estágio 3, e estágio 4 do sono Não-REM e o sono REM. Se essas seis classes forem submetidas a um classificador de tão baixa acurácia que o resultado da classificação se torne aleatório, esta acurácia seria de em torno de 1/6, ou cerca de 17%, sendo esse o pior cenário.

Para se construírem os vetores de características, o banco de dados foi dividido em janelas de EEG de 5 segundos, pois, analisando-se as suas anotações, todos os períodos dos estágios tinham durações múltiplas de 5 segundos. Em seguida, foi aplicada a transformada de Fourier, utilizando a biblioteca *numpy* (NUMPY, 2023), em cada uma dessas janelas. Os coeficientes obtidos da transformada de Fourier foram divididos em sub-bandas, sendo elas: a Delta (0,5-4 Hz), a Theta (4-8 Hz), a Alpha (8-12 Hz), a Beta (15-30 Hz) e a Gamma (30-128 Hz), de forma análoga a feita no trabalho de Malafeev (2018). Utilizando a média dos módulos dos coeficientes em cada faixa, foram montados os vetores de características contendo, assim, 5 elementos cada um. Foi desenvolvido um programa utilizando a linguagem de programação Python (PYTHON SOFTWARE FOUNDATION, 2023), da qual foram utilizadas as bibliotecas *numpy* e *pandas* (PANDAS, 2023). A biblioteca *pandas* permite a leitura e a manipulação do banco de dados e a biblioteca *numpy* facilita as tratativas matemáticas utilizadas. Para exemplificar o fluxo do processo foi criado o fluxograma visto na Figura 1. Nele, o passo a passo é ilustrado: primeiro os dados são extraídos do banco de dados, em seguida passam pela transformada de Fourier, cujo resultados formam os vetores de característica de acordo com as sub-bandas definidas e por fim são inseridos como entrada nos classificadores.

Figura 1: Diagrama do processo.



Fonte: Autores.

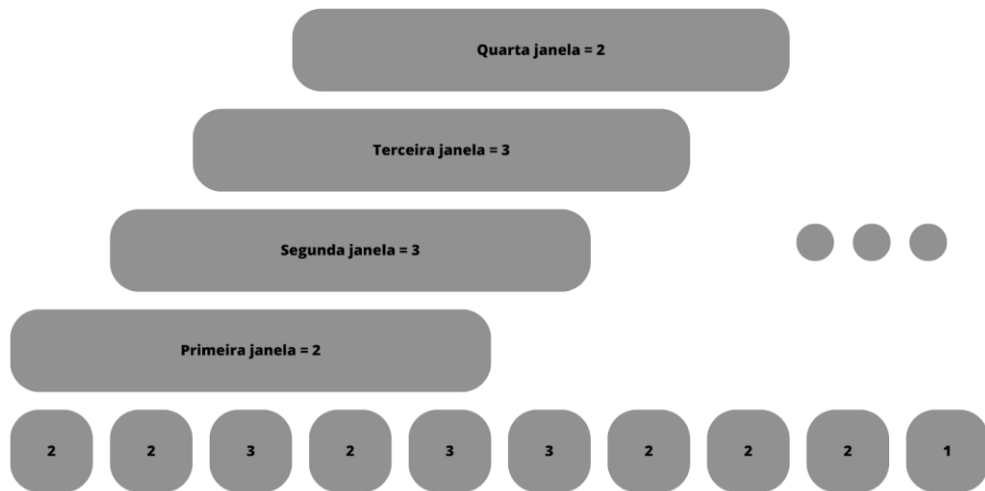
Para escolher o classificador, foram realizados experimentos, nos quais foram comparados diversos métodos, a fim de se encontrar aquele que proporcionasse a maior acurácia. Com base no estudo de Aboalayon et al. (2016), foram testados os classificadores DT e SVM. Foi, também, testado o classificador MLP, utilizado no trabalho de Ferreira (2018). Para os experimentos, foi aplicado cada um dos classificadores nos EEG dos quatro primeiros pacientes, medida a acurácia e obtida a matriz de confusão.

Foi também elaborado um teste em que se criava um ‘cursor’ que agrupava 5



vetores de classe. Esse cursor se movimenta por todos os vetores de classe, à medida que ele incrementa um próximo resultado ele se desfaz de seu valor mais anterior. Nesse teste, o cursor de agrupamento ganha o valor da classe de maior incidência, e se duas ou mais classes possuírem a mesma quantidade de elementos dentro do grupo, a primeira classe a aparecer define o valor do cursor, assim, foi possível criar um teste que determina em faixas de tempo a classe mais predominante. No experimento, esse cursor foi denominado de janela majoritária. Na Figura 2 há uma ilustração desse procedimento.

**Figura 2:** Funcionalidade do modo janela com prioridade às classes majoritárias.



Fonte: Autores.

O primeiro modelo implementado foi o de árvore de decisão (DT), neste classificador foram utilizados os parâmetros de base da biblioteca scikit-learn, com exceção dos fatores “criterion” e “max-depht”. O “criterion” se refere a qualidade das divisões da árvore, o critério utilizado neste parâmetro foi a entropia, que apesar de demorar mais por utilizar um algoritmo logarítmico ela tende a conseguir melhores resultados. Quanto ao parâmetro "max\_depth", ele estabelece a profundidade máxima da árvore, e no programa foi definido como "10".

Em relação ao modelo de perceptron de múltiplas camadas (MLP), também foram utilizados os parâmetros de base da biblioteca e ajustados os parâmetros "activation",

"solver" e "hidden\_layer\_sizes". O parâmetro "activation" corresponde à função de ativação utilizada nas camadas ocultas, sendo selecionada a função "relu" que corresponde a função de unidade linear retificada que pode ser conferida na Função 4. Quanto ao parâmetro "solver", trata-se do otimizador de pesos, e foi escolhido o método "lbfgs" que é um otimizador da família de métodos quasi-Newton. Por fim, o parâmetro "hidden\_layer\_sizes" representa o número de neurônios no modelo em cada camada oculta, e para ele foram configuradas 4 camadas que possuíam 200, 150, 100 e 50 neurônios respectivamente.

$$f(x) = \max(0, x) \quad (4)$$

No caso da SVM, foram utilizados os parâmetros na configuração base da biblioteca e ajustados os parâmetros "kernel", "gamma" e "tol". Foi adotado o kernel "RBF" (*Radial Basis Function*), que consegue combinar vários kernels polinomiais e de diferentes graus para conseguir separar dados em espaços dimensionais superiores a 2D. Já o parâmetro "gamma" define a abertura do kernel RBF. Para ele, foi selecionada a opção "auto", permitindo que a biblioteca determinasse o valor mais apropriado automaticamente. Por último, o parâmetro "tol" refere-se à tolerância para o critério de parada. A utilizada foi de 0.000001.

## Resultados

As análises dos classificadores foram feitas nos conjuntos de pacientes e noites ST7011J0-PSG, ST7012J0-PSG, ST7022J0-PSG e ST7041J0-PSG e também em um teste final, em em que foram consolidados os dados desses pacientes, a fim de validar se um banco de dados contendo dados de diversos pacientes gerava um resultado mais satisfatório. Foram testadas também algumas técnicas de balanceamento de classes, como a sobreamostragem, que gera dados de forma artificial para as classes de menor quantidade. (Essa técnica foi referenciada, adiante, no trabalho, como "Over"). Também foi utilizada a técnica de subamostragem, que mantém os dados das classes com menor quantidade e diminui a quantidade dos dados com maior quantidade. Esse método foi referenciado no trabalho como "Under". O termo "janela" se refere a técnica de agrupamento majoritário de 5 janelas de 5 segundos explicado anteriormente. O banco de dados conta com dois canais de EEG: Fpz-Cz e Pz-Oz. Inicialmente, realizaram-se testes

utilizando os dois canais. Mas, posteriormente, verificou-se que resultados melhores foram apresentados utilizando o canal Fpz-Cz individualmente. Os resultados a seguir foram obtidos utilizando apenas esse canal.

Os resultados de acurácia para cada paciente, bem como os resultados gerais do teste final consolidado, podem ser visualizados nas Tabelas 1, 2, 3, 4 e 5. Cada tabela apresenta informações detalhadas sobre as métricas obtidas. A primeira coluna de cada tabela corresponde à técnica utilizada, indicando o método empregado na análise dos dados. A segunda coluna fornece os valores de acurácia alcançados por cada técnica. As colunas subsequentes apresentam os valores do F1 score específico para cada uma das seis classes em relação às demais, sendo importante destacar que a classe 0 representa o estado de vigília, enquanto a classe 5 representa o sono REM. Desta forma, as tabelas fornecem uma visão abrangente dos resultados alcançados, permitindo uma análise mais detalhada e comparativa entre as diferentes técnicas empregadas.

**Tabela 1:** Resultados de acurácia para o conjunto ST7011J0-PSG.

PINI	Acurácia	F1-0	F1-1	F1-2	F1-3	F1-4	F1-5
DT	0.47	0.47	0.22	0.62	0.17	0.35	0.18
DT Over	0.46	0.46	0.21	0.59	0.21	0.40	0.30
DT Under	0.39	0.47	0.21	0.51	0.23	0.28	0.25
MLP	0.65	0.71	0.24	0.78	0.24	0.49	0.60
MLP Over	0.53	0.66	0.39	0.60	0.28	0.54	0.57
MLP Under	0.55	0.68	0.38	0.63	0.27	0.53	0.00
SVM	0.67	0.72	0.06	0.79	0.00	0.50	0.66
SVM Over	0.60	0.70	0.31	0.70	0.30	0.55	0.61
SVM Under	0.61	0.71	0.30	0.72	0.21	0.54	0.65
DT Janela	0.78	0.34	0.07	0.88	0.00	0.11	0.04
DT Over Janela	0.69	0.28	0.00	0.83	0.00	0.11	0.00
DT Under Janela	0.61	0.27	0.09	0.77	0.11	0.39	0.04
MLP Janela	0.80	0.19	0.00	0.89	0.00	0.00	0.00
MLP Over Janela	0.28	0.34	0.00	0.37	0.05	0.12	0.10
MLP Under Janela	0.41	0.35	0.05	0.54	0.02	0.16	0.06
SVM Janela	0.82	0.26	0.00	0.90	0.00	0.00	0.00
SVM Over Janela	0.17	0.40	0.16	0.17	0.01	0.20	0.09
SVM Under Janela	0.18	0.35	0.17	0.21	0.01	0.14	0.09

**Tabela 2:** Resultados de acurácia para o conjunto ST7012J0-PSG.

P1N2	Acurácia	F1-0	F1-1	F1-2	F1-3	F1-4	F1-5
DT	0.66	0.58	0.32	0.78	0.28	0.00	0.59
DT Over	0.60	0.57	0.26	0.73	0.29	0.00	0.50
DT Under	0.56	0.62	0.29	0.69	0.34	0.00	0.46
MLP	0.73	0.71	0.34	0.84	0.39	0.00	0.75
MLP Over	0.62	0.67	0.37	0.70	0.46	0.00	0.68
MLP Under	0.64	0.70	0.39	0.72	0.45	0.00	0.70
SVM	0.66	0.71	0.39	0.73	0.47	0.00	0.74
SVM Over	0.70	0.73	0.35	0.77	0.47	0.00	0.78
SVM Under	0.70	0.72	0.33	0.78	0.45	0.00	0.76
DT Janela	0.88	0.20	0.00	0.94	0.00	0.00	0.42
DT Over Janela	0.84	0.29	0.00	0.91	0.00	0.00	0.16
DT Under Janela	0.74	0.23	0.00	0.85	0.05	0.00	0.21
MLP Janela	0.89	0.00	0.00	0.94	0.00	0.00	0.11
MLP Over Janela	0.52	0.13	0.00	0.69	0.07	0.00	0.17
MLP Under Janela	0.43	0.07	0.00	0.61	0.09	0.00	0.26
SVM Janela	0.90	0.17	0.00	0.95	0.00	0.00	0.46
SVM Over Janela	0.54	0.20	0.08	0.69	0.05	0.00	0.28
SVM Under Janela	0.54	0.21	0.00	0.68	0.07	0.00	0.30

**Tabela 3:** Resultados de acurácia para o conjunto ST7022J0-PSG.

P2N2	Acurácia	F1-0	F1-1	F1-2	F1-3	F1-4	F1-5
DT	0.59	0.62	0.24	0.67	0.38	0.74	0.52
DT Over	0.53	0.54	0.25	0.61	0.34	0.73	0.42
DT Under	0.51	0.54	0.26	0.58	0.38	0.69	0.41
MLP	0.66	0.60	0.29	0.74	0.42	0.83	0.62
MLP Over	0.60	0.55	0.25	0.68	0.47	0.82	0.57
MLP Under	0.62	0.58	0.35	0.69	0.47	0.82	0.58
SVM	0.65	0.58	0.36	0.72	0.51	0.86	0.60
SVM Over	0.67	0.59	0.37	0.76	0.50	0.85	0.64
SVM Under	0.66	0.60	0.31	0.76	0.49	0.86	0.62
DT Janela	0.56	0.42	0.00	0.70	0.02	0.47	0.27
DT Over Janela	0.60	0.05	0.00	0.74	0.02	0.47	0.27
DT Under Janela	0.54	0.25	0.04	0.70	0.09	0.49	0.28
MLP Janela	0.69	0.17	0.00	0.82	0.00	0.32	0.08
MLP Over Janela	0.34	0.07	0.00	0.39	0.08	0.37	0.31
MLP Under Janela	0.35	0.15	0.00	0.41	0.09	0.37	0.37
SVM Janela	0.66	0.22	0.00	0.79	0.03	0.49	0.33
SVM Over Janela	0.40	0.42	0.00	0.45	0.16	0.48	0.34
SVM Under Janela	0.40	0.41	0.00	0.44	0.16	0.49	0.38

**Tabela 4:** Resultados de acurácia para o conjunto ST7041J0-PSG.

P4N1	Acurácia	F1-0	F1-1	F1-2	F1-3	F1-4	F1-5
DT	0.48	0.04	0.02	0.63	0.20	0.33	0.39
DT Over	0.40	0.07	0.07	0.56	0.19	0.21	0.37
DT Under	0.34	0.10	0.13	0.49	0.15	0.30	0.30
MLP	0.54	0.02	0.09	0.69	0.04	0.30	0.48
MLP Over	0.33	0.19	0.16	0.37	0.14	0.39	0.49
MLP Under	0.35	0.21	0.12	0.43	0.19	0.39	0.47
SVM	0.57	0.00	0.00	0.70	0.00	0.22	0.51
SVM Over	0.50	0.17	0.05	0.64	0.16	0.43	0.50
SVM Under	0.51	0.20	0.06	0.65	0.18	0.42	0.51
DT Janela	0.81	0.00	0.00	0.89	0.00	0.00	0.10
DT Over Janela	0.72	0.00	0.00	0.83	0.00	0.00	0.16
DT Under Janela	0.64	0.00	0.00	0.78	0.00	0.12	0.20
MLP Janela	0.85	0.00	0.00	0.92	0.00	0.00	0.00
MLP Over Janela	0.13	0.00	0.00	0.00	0.00	0.60	0.23
MLP Under Janela	0.09	0.00	0.00	0.00	0.00	0.08	0.23
SVM Janela	0.85	0.00	0.00	0.92	0.00	0.00	0.00
SVM Over Janela	0.13	0.02	0.00	0.08	0.00	0.06	0.28
SVM Under Janela	0.14	0.01	0.00	0.11	0.02	0.02	0.30

**Tabela 5:** Resultados de acurácia para o teste geral consolidado.

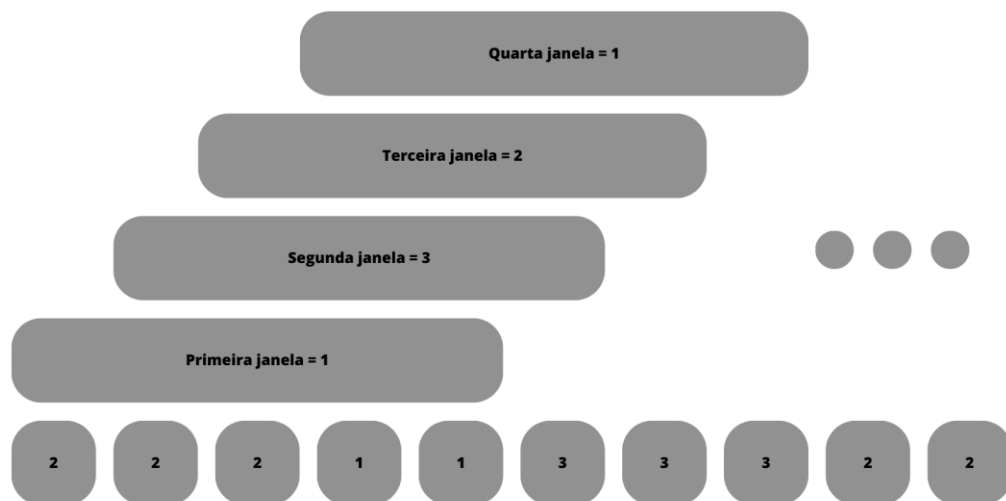
Total	Acurácia	F1-0	F1-1	F1-2	F1-3	F1-4	F1-5
DT	0.35	0.28	0.27	0.53	0.18	0.02	0.04
DT Over	0.21	0.22	0.16	0.28	0.22	0.19	0.16
DT Under	0.20	0.21	0.16	0.32	0.14	0.03	0.06
MLP	0.37	0.32	0.24	0.53	0.00	0.00	0.00
MLP Over	0.33	0.40	0.15	0.49	0.10	0.00	0.14
MLP Under	0.32	0.40	0.14	0.50	0.08	0.06	0.10
SVM	0.38	0.24	0.35	0.53	0.00	0.00	0.00
SVM Over	0.28	0.39	0.17	0.39	0.20	0.10	0.25
SVM Under	0.28	0.39	0.19	0.38	0.18	0.10	0.24
DT Janela	0.49	0.05	0.16	0.66	0.04	0.00	0.00
DT Over Janela	0.23	0.22	0.23	0.34	0.09	0.04	0.08
DT Under Janela	0.31	0.20	0.27	0.44	0.06	0.00	0.02
MLP Janela	0.49	0.03	0.03	0.66	0.00	0.00	0.00
MLP Over Janela	0.47	0.28	0.03	0.65	0.00	0.00	0.03
MLP Under Janela	0.47	0.27	0.02	0.66	0.00	0.00	0.02
SVM Janela	0.50	0.00	0.07	0.66	0.00	0.00	0.00
SVM Over Janela	0.26	0.27	0.07	0.45	0.10	0.01	0.12
SVM Under Janela	0.25	0.26	0.08	0.43	0.05	0.00	0.11

Foi notado que, apesar de alguns métodos apresentarem uma alta acurácia, eles erravam muito algumas classes, classes estas que aparecem por breves períodos contínuos apenas, no EEG. Contudo, utilizando as técnicas de sobreamostragem e de subamostragem, que nivelam as ocorrências de cada classe, constatou-se que a acurácia caiu.

Para tentar mitigar esse problema, foi testada mais uma técnica que consistia basicamente em uma varredura para tentar evidenciar essas classes minoritárias. Se existisse alguma ocorrência dessas classes durante uma análise pelo método janela, ela era exposta como sendo a classe daquele bloco, como ilustrado pela Figura 3.

Nela, a primeira sequência de cinco janelas de EEG, à esquerda, tem duas janelas classificadas como sendo da classe 1 e três janelas da classe 2. Como a classe 1 é minoritária na sequência, a sequência como um todo é classificada como sendo da classe 1. Na próxima sequência de cinco janelas, parcialmente sobreposta à primeira sequência, a classe minoritária é a 3 e assim por diante. Quando há empate nos números de janelas de classes minoritárias, é adotada, para a sequência, a classe da primeira minoritária a aparecer na sequência.

**Figura 3:** Funcionalidade do modo janela com prioridade às classes minoritárias.



Fonte: Autores.

Durante a análise dos resultados, observou-se que as classes de curta duração

apresentam recorrência. Essa característica acabou comprometendo a precisão das classes com maior incidência. Essa observação pode ser evidenciada em um exemplo específico de resultados de um dos testes, conforme demonstrado na Tabela 6. Nessa tabela, é possível constatar a diminuição da assertividade nas classes mais frequentes devido à presença recorrente das classes de curta duração.

**Tabela 6:** Exemplo de resultados de acurácia usando varredura em busca de classes minoritárias.

Teste	Acurácia	F1-0	F1-1	F1-2	F1-3	F1-4	F1-5
MLP Under Janela	0.22	0.26	0.21	0.00	0.15	0.24	0.25

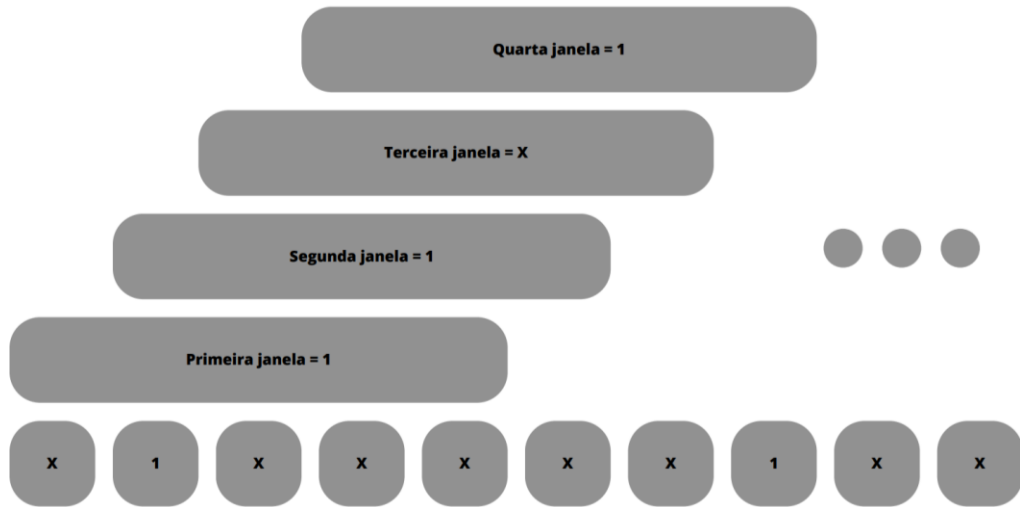
A classe 2 é a classe com maior recorrência e duração. Nesse método, ela acabou desaparecendo. Esse sintoma se repetiu para os outros conjuntos de pacientes e noites testados.

Verificou-se também que no teste em que foi criado um banco de dados com diversos pacientes não houve um ganho muito significativo.

Portanto, com base nos resultados, o método mais eficiente foi o SVM utilizando a técnica de sobreamostragem, que conseguiu uma acurácia média de 62%, e um *F1 score* médio de 0,50.

Seguindo uma outra linha de raciocínio, decidimos analisar a eficácia para cada classe individualmente. O experimento foi inicialmente realizado dando prioridade a classe 1, o primeiro estágio do sono Não-REM, que é uma das classes com menor ocorrência no sono dos pacientes. A nova janela prioritária verificava se havia alguma ocorrência da classe 1 durante a varredura, e caso encontrasse alguma, a tratava como a classe do bloco. Se a janela não detectasse nenhuma ocorrência do estágio 1, a classe do bloco era tratada como uma classe qualquer, que foi denominada X. A Figura 4 ilustra a operação da nova janela prioritária.

**Figura 4:** Funcionalidade do modo janela com prioridade à uma única classe.



Fonte: Autores.

Após a aplicação da nova janela prioritária, o classificador foi treinado e testado utilizando o conjunto ST7042J0-PSG. Para verificar a efetividade do experimento, foi coletado o novo *F1 score* da classe 1 para julgar se houve aumento na performance do classificador SVM Over em discernir ocorrências minoritárias após aplicado o critério de prioridade. Em sequência, o mesmo experimento foi refeito para todas as classes e todos os novos valores de *F1 score* foram comparados aos valores de um treinamento pré-teste. O resultado pode ser verificado na Tabela 7.

**Tabela 7:** Resultados do teste de janela prioritária para uma única classe.

Teste	Acurácia	F1-Estágio pré-teste	F1-Estágio priorizado	F1-X
Estágio 0	0.35	0.28	0.27	0.53
Estágio 1	0.21	0.22	0.16	0.28
Estágio 2	0.20	0.21	0.16	0.32
Estágio 3	0.37	0.32	0.24	0.53
Estágio 4	0.33	0.40	0.15	0.49
Estágio 5	0.32	0.40	0.14	0.50

O experimento evidenciou que houve um aumento significativo para todas as classes, com exceção da classe 4, que é a menos recorrente nas anotações presentes no espectro do EEG de cada paciente no banco de dados estudado, havendo, também, uma completa inexistência de anotações sobre a classe 4 para alguns conjuntos de pacientes e



noites.

Como forma de tentar transportar os resultados desses testes para o algoritmo de classificador, estabeleceu-se uma hierarquia de prioridades. A janela prioritária agora verificava inicialmente se havia ocorrência da classe 4, a menos recorrente. Caso não encontrasse a classe 4, ele passaria a buscar a 1, e assim seguindo a lista das menos presentes, buscando posteriormente se era classe 3, em seguida 0, 5 e por último a classe 2, a mais abundante de todas. O resultado para esse último experimento está na Tabela 8.

**Tabela 8:** Resultados do experimento de janela prioritária hierárquica generalizada.

Teste prioritário	Acurácia	F1-0	F1-1	F1-2	F1-3	F1-4	F1-5
SVM Over	0.35	0.26	0.23	0.20	0.33	0.53	0.23

Evidentemente, este experimento não teve bons resultados. O principal motivo para isso é que, embora as classes minoritárias possuam poucas ocorrências, elas são suficientemente segmentadas durante todo o espectro de sono EEG, segundo as anotações dos técnicos de sono. Portanto, a classe 2, que sempre retornava valores satisfatórios em seu *F1 score*, recebeu uma queda significativa em sua acurácia, enquanto que não foi observada uma melhora nos resultados obtidos pelas classes minoritárias.

## Conclusão

Dos experimentos realizados, conclui-se que, com as técnicas aqui usadas, a acurácia na classificação tende a ser maior quando se busca uma fase específica do sono no EEG do que quando se busca classificar as janelas de EEG, simultaneamente, em seis classes. Isso poderia ser ampliado em um projeto que busque encontrar, por exemplo, o estágio 1 do sono Não-REM para identificar quando alguém ainda acordado apresenta sinais de sonolência, se tornando uma forma de alertar um motorista que ele não está apto a continuar dirigindo, necessitando de descanso.

Apesar de o resultado obtido ser inferior aos dos artigos estudados, existem fatores que possam ter contribuído para isso, como a diferença no número de classes trabalhadas, pois em todos aqueles artigos, os autores utilizaram cinco classes e, neste artigo, experimentou-se classificar todas as seis classes possíveis. Outro fator importante pode ser a forma como os dados foram pré-processados, pois apenas foi aplicada a FFT sobre o EEG, sem mais nenhuma técnica de pré-processamento que os outros estudos

utilizaram.

Contudo, os resultados foram satisfatórios. Foi possível alcançar uma acurácia média de 62%, consideravelmente maior que aquela que seria obtida aleatoriamente, de 17%. Foi observado que as faixas de sono variam muito e que, apesar de algumas classes aparecerem com uma certa recorrência, a duração delas é muito baixa.

Levando-se em conta as publicações comentadas no estudo bibliográfico que usaram algum tipo de pré-processamento e os resultados aqui obtidos sem pré-processamento, conclui-se que o pré-processamento, provavelmente, é importante para se obter alta acurácia na classificação. Para trabalhos futuros sugere-se, então, usar alguma forma de pré-processamento dos dados e, também, tentar tratar o problema de outro ângulo, onde poder-se-iam olhar essas fases recorrentes e curtas como fases de transição e assim, talvez, até predizer em qual próximo estágio o usuário irá entrar.

## Referências

ABOALAYON, K. et al. **Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation**. 2016. Disponível em <https://www.mdpi.com/1099-4300/18/9/272>. Acesso em: 07 jun. 2023.

ABRAMET. **Problemas na saúde de motoristas causaram mais de 280 mil acidentes nas rodovias desde 2014, aponta Abramet**. 2020. Disponível em <https://abramet.com.br/noticias/problemas-na-saude-de-motoristas-causaram-mais-de-280-mil-acidentes-nas-rodovias-desde-2014-aponta-abramet/> Acesso em: 08 nov. 2022.

BRACEWELL, R. **The Fourier transform and its applications**. McGraw-Hill Education. Third Edition. p. 5-8, 258-264. 2000.

CORTES, C.; VAPNIK, V. **Support-vector networks**. Machine learning, v. 20, n. 3, p. 273-297, 1995.

FERREIRA, J. **Classificação dos estágios do sono utilizando análise espectral e redes neurais convolutivas**. 2018. Disponível em <https://monografias.ufma.br/jspui/bitstream/123456789/3542/1/JULIA-FERREIRA.pdf>. Acesso em: 07 jun. 2023.

HAN, J. et al. **Data Mining: Concepts and Techniques**. 3. ed. [S. l.]: Morgan Kaufmann, 2011. cap. 1, p. 18-19. Disponível em: <http://myweb.sabanciuniv.edu/rdehkharghani/files/2016/02/The-Morgan-Kaufmann-Series-in-Data-Management-Systems-Jiawei-Han-Micheline-Kamber-Jian-Pei-Data-Mining.-Concepts-and-Techniques-3rd-Edition-Morgan-Kaufmann-2011.pdf>. Acesso em: 24 maio 2023.

HAYKIN, S. **Neural networks and learning machines**. 3. ed. [S. l.]: Pearson, 2009. cap. 4, p. 123-125. Disponível em: <https://lps.ufrj.br/~caloba/Livros/Haykin2009.pdf>. Acesso em: 24 maio 2023.

KEMP, B. et al. **Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG**. IEEE Transactions on Biomedical Engineering, v. 47, n. 9, p. 1185-1194, 2000. DOI 10.1109/10.867928. Disponível em: <https://ieeexplore.ieee.org/document/867928>. Acesso em: 7 jun. 2023.

MALAFEEV, Alexander et al. **Automatic Human Sleep Stage Scoring Using Deep Neural Networks**. Frontiers: in Neuroscience, [s. l.], v. 12, 6 nov. 2018. DOI 10.3389/fnins.2018.00781. Disponível em: <https://www.frontiersin.org/articles/10.3389/fnins.2018.00781>. Acesso em: 23 fev. 2023.

MARIANO, D. **Métricas de avaliação em machine learning**. 2021. Disponível em <https://diegomariano.com/metricas-de-avaliacao-em-machine-learning/>. Acesso em: 07 jun. 2023.

MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. **EEG (electroencephalogram)**. 2022. Disponível em <https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>. Acesso em: 26 jan. 2023.

MAYO FOUNDATION FOR MEDICAL EDUCATION AND RESEARCH. **Polysomnography (sleep study)**. 2020. Disponível em <https://www.mayoclinic.org/tests-procedures/eeg/about/pac-20393875>. Acesso em: 26 jan. 2023.

MINISTÉRIO DA SAÚDE. **Distúrbios do sono**. 2012. Disponível em <https://bvsmis.saude.gov.br/disturbios-do-sono/>. Acesso em: 07 jun. 2023.

MÜLLER, M. R.; GUIMARÃES, S. S. **Impacto dos transtornos do sono sobre o funcionamento diário e a qualidade de vida**. Estudos de Psicologia (Campinas), 24 (4). p. 519-528, 2007. Disponível em: <https://www.scielo.br/j/estpsi/a/gTGLpgtmtMnTrcMyhGFvNpG/?format=pdf&lang=pt>. Acesso em: 7 jun. 2023.

NUMPY. **Numpy**. 2023. Disponível em <https://numpy.org/>. Acesso em 13 jun. 2023.

PANDAS. **Pandas**. 2023. Disponível em <https://pandas.pydata.org/>. Acesso em 13 jun. 2023.

PHYSIONET. **Sleep-EDF Database Expanded**. 2013. Disponível em <https://www.physionet.org/content/sleep-edfx/1.0.0/>. Acesso em: 07 jun. 2023.

PYTHON SOFTWARE FOUNDATION. **Python**. 2023. Disponível em <https://www.python.org/>. Acesso em 13 jun. 2023..

RECHTSCHAFFEN, A.; KALES, A. **Techniques, and scoring system for sleep stages of human subjects: a manual of standardised terminology**. Los Angeles:

UCLA Brain Information Service, 1968.

SCIKIT-LEARN. **Scikit-learn - Machine Learning in Python**. Disponível em <https://scikit-learn.org/stable/>. Acesso em 13 jun. 2023.

SEECK, M. et al. **The standardized EEG electrode array of the IFCN**. *Clinical Neurophysiology*, v. 128, p. 2070-2077, 2017.

Enviado: 19/06/2023