



Detectando estágios do sono utilizando classificação binária em sinais de EEG

Detecting sleep stages using binary classification in EEG signals

RESUMO

O sono é um processo fisiológico que pode ser dividido em quatro estágios: acordado, S1, S2, S3, S4 e REM. Sendo nos últimos quatro estágios, o período em que o corpo exerce funções restauradoras. Cada estágio está ligado a características fisiológicas, cujas anomalias podem indicar doenças, tais como, insônia e epilepsia. A análise dessas etapas tem papel indispensável no diagnóstico de doenças, porém, a classificação manual é um processo demorado. Neste contexto, este trabalho tem como objetivo comparar o desempenho de classificadores na rotulação de estágios do sono em segmentos de eletroencefalograma (EEG). Para isso, foram utilizadas as leituras de EEG de 19 voluntários. Esses sinais foram divididos em intervalos de 5 segundos, com cada um sendo usado na extração de características dos domínios: tempo, frequência e tempo-frequência. Em seguida, foram criados modelos de classificação usando os métodos Árvore de Decisão, K-vizinhos mais próximo (KNN), Naive Bayes e Floresta Aleatória. Subsequentemente, os modelos foram avaliados utilizando validação cruzada e testes estatísticos de hipótese. Como resultado, obtiveram-se valores de acurácia entre 72,83% e 95,21%. Adicionalmente, foi registrada diferença estatisticamente significativa nos seguintes pares de modelos: Árvore de Decisão vs. Naive Bayes; KNN vs. Floresta Aleatória; e Naive Bayes vs. Floresta Aleatória.

PALAVRAS-CHAVE: Aprendizado de Máquina. Detecção de Estágios do Sono. Eletroencefalograma.

ABSTRACT

Sleep is a physiological process that can be divided into four stages: awake, S1, S2, S3, S4, and REM. The last four stages are the period in which the body exercises restorative functions. Each stage is linked to physiological characteristics, whose abnormalities may indicate diseases, such as insomnia and epilepsy. The analysis of these stages has an indispensable role in the diagnosis of diseases, but manual classification is a time-consuming process. In this context, this paper aims to compare the performance of classifiers in labeling sleep stages in electroencephalogram (EEG) segments. To this end, the EEG readings of 19 volunteers were used. These signals were divided into 5-second intervals, with each being used to extract features from the time, frequency, and time-frequency domains. Next, classification models were created using Decision Tree, K-nearest neighbor (KNN), Naive Bayes, and Random Forest methods. Subsequently, the models were evaluated using cross-validation and statistical hypothesis tests. As a result, accuracy values between 72.83% and 95.21% were obtained. Additionally, a statistically significant difference was recorded in the following pairs of models: Decision Tree vs. Naive Bayes; KNN vs. Random Forest; and Naive Bayes vs. Random Forest.

KEYWORDS: Machine Learning. Sleep Stage Detection. Electroencephalogram.

1 INTRODUÇÃO

Conforme apresentado no dicionário Oxford (OXFORD, 2022), sono é definido como um estado de repouso físico e mental em que ocorre a desaceleração do metabolismo, relaxamento muscular e diminuição da atividade sensorial, que pode ser dividido em dois tipos: NREM (*Non Rapid Eye Movement*), que pode ser dividido em 5 estágios (acordado, S1, S2, S3 e S4), e REM (*Rapid Eye Movement*). Cada estágio do sono está associado a uma série de características fisiológicas e neurais que podem vir a ser usadas no diagnóstico de doenças, tais como insônia, hipersonia e epilepsia (LAJNEF, T. et al., 2015).

Tendo em mente a importância médica em se desenvolver métodos de classificar estágios do sono, este trabalho tem o objetivo de utilizar diferentes métodos de aprendizado de máquina para a construção e a comparação de classificadores para a diferenciação de estágios do sono a partir de eletroencefalogramas (EEG), de maneira em



que se possa compará-los e observar quais mais se adequam a identificação dos estágios do sono.

2 TRABALHOS RELACIONADOS

Tarek Lajnef et al. (2015) ressalta que mesmo com os resultados promissores de métodos automatizados de classificação dos estágios do sono, ainda há espaço para melhorias. Foram usados neste trabalho para classificação multiclasse estágios do sono os métodos máquina de vetores de suporte (SVM, do inglês *Support Vector Machine*) com a estratégia um-contra-todos (OAA, do inglês *one-against-all*), análise discriminante linear (LDA, do inglês *Linear Discriminant Analysis*) e SVM baseado em dendograma, obtendo acurácias médias de 69,20%, 62,60% e 74,74%, respectivamente.

No artigo de Nan-Hung Lin et al. (2017) é argumentado a importância da detecção do estágio de sono REM no diagnóstico de doenças como o Alzheimer, depressão, insônia, epilepsia, desordem compulsiva obsessiva e esquizofrenia. Foram usadas 947 gravações de polissonografia para treino e teste de uma rede neural que classifica cada segmento como REM ou NREM, essa rede foi usada para se identificar padrões entre comportamentos irregulares no sono REM e a existência das doenças citadas anteriormente. O modelo proposto obteve uma acurácia de 93% na classificação do estágio REM, sendo este, um valor inferior à acurácia média de 97% obtido pela classificação da mesma leitura por um grupo de teste com 2500 especialistas.

Md Junayed Hasan et al. (2020) propõe a criação de um modelo de classificação para automatizar o processo de análise das etapas do sono utilizando sinais de EEG gravados de três diferentes partes do cérebro humano: córtex central, pré-frontal e lobo occipital. Para criação desse modelo, foi usado as características extraídas dos sinais de EEG no treinamento de uma rede neural residual (ResNet – do inglês *Residual Neural Network*) que alcançou uma acurácia média 87,8% e 83,7%, respectivamente, para classificação dos estágios em pessoas do gênero feminino e masculino.

Rym Nihel Sekkal et al. (2022) destaca o quão tedioso e demorado é o processo de se classificar manualmente os estágios do sono, o que leva a vários estudos sobre a automação do processo de classificação de polissonografias usando métodos de aprendizado profundo e algoritmos clássicos de aprendizado de máquina. Seu trabalho teve como objetivo comparar o desempenho entre essas duas classes de modelos de aprendizado. Os modelos escolhidos para o estudo foram perceptron multicamadas (MLP, do inglês *Multilayer Perceptron*), SVM de base radial, floresta aleatória, SVM linear, regressão logística, KNN, análise discriminante quadrática (QDA, do inglês *Quadratic Discriminant Analysis*), árvore de decisão e naive bayes, obtendo acurácias médias de 88,1%, 89,1%, 87,9%, 86,5%, 85,6%, 85,2%, 81,1%, 80,6%, e 73,9%, respectivamente.

3 MATERIAIS E MÉTODOS

3.1 FERRAMENTAS

Para o desenvolvimento deste trabalho, foi utilizado a linguagem de programação Python em conjunto com os pacotes:



- Sklearn¹: usada para aplicação de métodos de aprendizado de máquina.
- Pandas²: usado para facilitar as manipulações feitas na base de dados.
- MNE³: utilizado para o download e visualização de EEG.
- Imbalanced Learn⁴: providenciou métodos para balanceamento de classes.

3.2 BASE DE DADOS

Foram utilizadas as leituras de EEG presentes na base de dados da PhysioNet (KEMP, B. et al, 2000). As gravações desta base de dados foram obtidas no período entre 1987 a 1991 durante um estudo sobre o efeito da idade no sono. Neste estudo, foi gravado durante dois dias seguidos a polissonografia de 77 pessoas com idade entre 25 a 101 anos, com cada um desses indivíduos não utilizando quaisquer medicações que afetassem o seu sono. Os dados gravados durante a polissonografia foram: EEG Fpz-Cz, EEG Pz-Oz e eletrooculograma (EOG) horizontal, sendo gravados na frequência de 100Hz, e respiração oronasal, EMG submental e temperatura retal, gravados com frequência de 1 Hz.

Com cada leitura possuindo em torno 20 horas de duração, dos 77 indivíduos, foram selecionadas as leituras de 19, com idades entre 25 e 34 anos, para serem usadas nos treinos e testes dos classificadores de forma a diminuir a complexidade de se usar um grupo com diferentes faixas etárias (KEMP, B. et al, 2000).

3.3 EXTRAÇÃO DE CARACTERÍSTICAS

Para as leituras do EEG serem usadas para treino e teste dos classificadores foi preciso dividi-las em intervalos de 5 segundos para que se fosse possível fazer a classificação do estágio a cada janela desse intervalo, em cada segmento foram extraídas características considerando os seguintes domínios: tempo (série temporal), frequência (espectro de potência) e tempo-frequência (espectrograma). No domínio da frequência e tempo-frequência foram extraídas as características de cada uma das bandas de frequências cerebrais: Delta (1Hz z 4Hz), Theta (4Hz z 8Hz), Alpha (8Hz a 12Hz), Beta (12Hz a 30Hz), Gamma (30Hz a 60Hz) e espectro inteiro (1Hz a 60Hz) (OLIVA, J. T.; ROSA, J. L. G, 2021). Foram extraídas 16 características de cada banda no domínio tempo-frequência, 27 características de cada banda no domínio da frequência e 27 características do espectro de potência, somando-se assim um total de 285 características computadas por intervalo de 5 segundos. As características que foram extraídas neste trabalho são apresentadas no material complementar⁵.

3.4 CONSTRUÇÃO DE MODELOS PREDITIVOS

Com as leituras obtidas foram geradas duas listas, ambas com as características extraídas. A lista 1 foi balanceada utilizando o método de undersampling em relação aos estágios do sono, de maneira a diminuir a quantidade de falsos positivos. A lista 1 foi

¹ Sklearn: <https://scikit-learn.org/stable/>

² Pandas: <https://pandas.pydata.org/>

³ MNE: <https://mne.tools/stable/index.html>

⁴ Imbalanced Learn: <https://imbalanced-learn.org/stable/>

⁵ Material Suplementar: ocultado



usada para treinar os modelos de classificação usando a estratégia OAA e a lista 2 para o teste dos classificadores. Os classificadores usados neste trabalho foram:

- Árvores de Decisão (AD): método de aprendizado de máquina supervisionado não paramétrico usado para classificação e regressão. Funciona dividindo o problema estudado em nós de uma árvore e percorrendo até que seja atingido um nó folha que representa uma decisão (ALPAYDIN, E, 2020).
- KNN: método de aprendizado de máquina supervisionado não paramétrico usado para classificação e regressão. Se utiliza da proximidade entre os pontos de dados para fazer previsões e classificações (ALTMAN, N. S., 1992).
- Naive Bayes (NB): método de aprendizado supervisionado de máquina baseado na aplicação do teorema de Bayes assumindo a independência condicional entre as características (Sklearn - Naive Bayes, 2022).
- Floresta Aleatória (RF): método de aprendizado de máquina supervisionado não paramétrico usado para classificação e regressão. Funciona criando várias Árvores de Decisão por meio da seleção aleatória de atributos (HO, T. K, 2002).

3.5 AVALIAÇÃO DOS MODELOS

Na avaliação, foi utilizada a validação cruzada *k-fold*, que consiste em dividir os dados em *k* partições de tamanhos iguais, onde a *i*-ésima partição é o conjunto de teste e as *k-1* partições restantes compõem o conjunto de treinamento (KOHAVI, 1995). Foi usado neste trabalho a validação cruzada para 10 *folds*, dos quais foram obtidas as medidas estatísticas média e desvio padrão calculadas para cada fold. Além disso, foram aplicados os testes estatísticos de hipótese de *Friedman* e *Nemenyi* para se verificar se existe diferença estatística significativa entre os modelos de classificação.

4 RESULTADOS

A tabela abaixo apresenta os valores de Acurácia Média (AM) e Desvio Padrão (DP) referente aos modelos preditivos utilizados para classificar cada estágio do sono.

Tabela 1 -

Estágios	Árvore de Decisão		KNN		Naive Bayes		Floresta Aleatória	
	AM (%)	DP (%)	AM (%)	DP (%)	AM (%)	DP (%)	AM (%)	DP (%)
Acordado	94,36	0,22	76,93	0,53	74,38	0,26	95,21	0,26
S1	80,15	2,14	79,26	1,01	74,88	1,59	82,07	1,44
S2	87,9	0,37	75,8	0,85	72,83	0,61	89,88	0,37
S3	89,87	1,36	75,02	2,26	77,93	2,47	91,59	1
S4	90,17	2,16	76,88	3,13	80,54	5,58	92,99	2,52
REM	83,19	1,24	77,45	0,98	75,48	0,96	85,48	0,83

Fonte: Autoria própria (2022).

O modelo de FA alcançou os maiores valores de AM em metade dos estágios do sono (Acordado, S1, S4). NB apresentou os menores valores de AM na maioria dos estágios (Acordado, S1, S2, REM). Realizando o teste estatístico de *Friedman*, foi verificado que em todos os estágios do sono houve diferença estatística significativa entre



os classificadores utilizados. Para verificar em quais pares de modelos houve diferença estatisticamente significativa, foi aplicado o pós-teste de *Nemenyi* (HOLLANDER, M.; WOLFE, D. A, 1999). Após a aplicação do pós-teste para cada estágio do sono, em todos os casos foi registrada diferença estatisticamente significativa entre os seguintes pares de modelos: AD vs. NB; KNN vs. RF; e NB vs. RF. Adicionalmente, para o estágio S3 ainda foi registrada diferença estatisticamente significativa na comparação entre os modelos AD e KNN. Com esses resultados pode-se afirmar que os modelos de AD e RF foram os que mais se destacaram positivamente por não possuírem diferença estatisticamente significativa entre si e terem alcançado acurácia média maior que 80% em todos os estágios.

5 CONCLUSÃO

Este trabalho teve como objetivo comparar métodos de aprendizado de máquina na classificação de estágios de sono através de leituras de EEG. Na fase experimental deste trabalho foram alcançados resultados promissores, com nenhum dos classificadores obtendo acurácia média menor que 70%, porém, com a aplicação dos testes de *Friedman* e *Nemenyi*, foi constatado que houve diferença estatística significativa entre os classificadores nos diferentes estágios do sono, com os modelos de AD e RF se destacando positivamente na classificação de todos os estágios do sono por não possuírem diferença estatística significativa entre si e terem alcançado acurácia média maior que 80% em todos os estágios.

Como trabalho futuro podem ser implementados métodos de classificação em tempo real, de maneira em que se possa comparar o tempo médio gasto na rotulação de cada segmento usando diferentes classificadores. Adicionalmente pode ser implementado métodos de seleção de atributos de modo a reduzir o tempo necessário para extração de características, treino dos modelos e testes.

AGRADECIMENTOS

Ao PIBIC/Fundação Araucária pelo apoio financeiro através da concessão de bolsa de iniciação científica.

CONFLITO DE INTERESSE

Não há conflito de interesse.

REFERÊNCIAS

ALPAYDIN, E. **Introduction to Machine Learning**. 3. ed. Londres, England: MIT Press, 2014.

ALTMAN, N. S. An introduction to kernel and nearest-neighbor nonparametric regression. **The American statistician**, v. 46, n. 3, p. 175–185, 1992.



HASAN, M. J. et al. Sleep state classification using power spectral density and residual neural network with multichannel EEG signals. **Applied sciences (Basel, Switzerland)**, v. 10, n. 21, p. 7639, 2020.

HO, T. K. **Random decision forests**. Proceedings of 3rd International Conference on Document Analysis and Recognition. **Anais...IEEE Comput. Soc. Press**, 2002.

HOLLANDER, M.; WOLFE, D. A. **Nonparametric Statistical Methods**. 2. ed. Nashville, TN, USA: John Wiley & Sons, 1999

KEMP, B. et al. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. **IEEE transactions on bio-medical engineering**, v. 47, n. 9, p. 1185–1194, 2000.

KOHAVI, R. **A study of cross-validation and bootstrap for accuracy estimation and model selection**. Proceedings of the 14th international joint conference on Artificial intelligence - Volume 2. **Anais...San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.**, 1995. Acesso em: 14 ago. 2022.

LAJNEF, T. et al. Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines. **Journal of neuroscience methods**, v. 250, p. 94–105, 2015.

LIN, N.-H. et al. Detecting rapid eye movement sleep using a single EEG signal channel. **Expert systems with applications**, v. 87, p. 220–227, 2017.

Naive Bayes. Disponível em: <https://scikit-learn.org/stable/modules/naive_bayes.html>. Acesso em: 14 ago. 2022.

OLIVA, J. T.; ROSA, J. L. G. Binary and multiclass classifiers based on multitaper spectral features for epilepsy detection. **Biomedical signal processing and control**, v. 66, n. 102469, p. 102469, 2021.

OXFORD - Sleep Definition, pictures, pronunciation and usage notes. Disponível em: <https://www.oxfordlearnersdictionaries.com/us/definition/american_english/sleep_2>. Acesso em: 14 ago. 2022.

Ridge regression and classification. Disponível em: <https://scikit-learn.org/stable/modules/linear_model.html>. Acesso em: 14 ago. 2022.

SEKKAL, R. N. et al. Automatic sleep stage classification: From classical machine learning methods to deep learning. **Biomedical signal processing and control**, v. 77, n. 103751, p. 103751, 2022.