



Classificação de emoções em eletroencefalogramas com aprendizado de máquina

Emotion classification in electroencephalograms using machine learning

Otávio Pigozzo Martelli *

Jefferson Tales Oliva†

RESUMO

Compreender o comportamento humano é fundamental para entender as necessidades e vontades individuais e coletivas dos humanos. Nesse contexto, o estudo das emoções tem papel indispensável na compreensão do comportamento humano, pois elas influenciam as ações e reações de cada indivíduo. Este trabalho tem como objetivo classificar emoções a partir de segmentos de eletroencefalograma (EEG), utilizando métodos de aprendizado de máquina. Para a classificação foram utilizados sinais de EEG de 20 voluntários. Esses sinais foram segmentados entre os intervalos de tempo em que as emoções induzidas aconteciam. Dessa forma, foram analisadas oito emoções (temor, frustração, alegria, raiva, felicidade, tristeza, amor e excitação). Em uma avaliação preliminar, essas emoções foram divididas em dois conjuntos (positivas e negativas) para classificação binária. Em seguida, para cada segmento de EEG, características do domínio do tempo foram extraídas. Feito isso, foram criados modelos de classificação por meio dos métodos árvore de decisão, k -vizinhos mais próximos (KNN) e floresta aleatória. Subsequentemente, os modelos foram avaliados utilizando validação cruzada e testes estatísticos de hipótese. Como resultado, foram atingidas valores de acurácia entre 28,5% e 83,9% considerando todos os classificadores. Adicionalmente, foi constatada diferença estatisticamente significativa entre os modelos floresta aleatória e KNN para oito classes.

Palavras-chave: Classificação de Emoções. Eletroencefalograma. Aprendizado de Máquina.

ABSTRACT

Understanding human behavior is fundamental to understanding the individual and collective needs and wants of humans. In this context, the study of emotions has an indispensable role in understanding human behavior, as they influence the actions and reactions of each individual. This work aims to classify emotions from electroencephalogram (EEG) segments using machine learning methods. For the classification, EEG signals from 20 volunteers were used, segmenting the signals between the time intervals in which the induced emotions occurred. Thus, eight emotions (fear, frustration, joy, anger, happiness, sadness, love, and excitement) are classified. In a preliminary evaluation, these emotions were divided into two sets (positive and negative) for binary classification. Then, feature extraction techniques were applied to the EEG segments, obtaining time-domain attributes. Next, classification models are built using the decision tree, k -nearest neighbors (KNN), and random forest methods. Subsequently, these models are evaluated using cross-validation and statistical hypothesis tests. As a result, accuracy values between 28.5% e 83.9% were obtained considering all classifiers. Additionally, a statistically significant difference was found only between the random forest and KNN models for eight classes.

Keywords: Emotion Classification. Electroencephalogram. Machine Learning.

* Departamento Acadêmico de Informática (DAINF), Engenharia de Computação; ✉ martelli@alunos.utfpr.edu.br.

† Departamento Acadêmico de Informática (DAINF), Engenharia de Computação; ✉ jeffersonoliva@utfpr.edu.br.



1 INTRODUÇÃO

Os computadores são incapazes de demonstrar emoções humanas genuínas, mas, eles são capazes de identificá-las, isso se tornou possível devido aos avanços tecnológicos em inteligência artificial (MURUGAPPAN, 2011). A identificação de emoções humanas tem empregabilidade em diversas áreas, como determinação de satisfação do cliente, apontamento de engajamento em redes sociais, identificação de doenças psicológicas e reconhecimento de humor em interrogatórios. Para isso, existem diversas formas de identificação, por exemplo, reconhecimento facial, padrão de escrita e áudio (HUTTO; GILBERT, 2014; LEON et al., 2005; FAIRHURST; E.; LI, 2014).

Emoção trata-se da forma como o indivíduo "percebe" o mundo ao ser redor, a influência de suas próprias memórias e suas motivações são fatores decisivos para alteração do estado emocional interpretado pelo sistema límbico e enviados aos neurotransmissores (LEON et al., 2005), os quais produzem respostas fisiológicas, alterando a percepção e as ações do indivíduo em relação ao meio em que vive (ESPERIDIÃO-ANTONIO et al., 2008). Com relação a atividade cerebral, é possível utilizar métodos de aprendizado de máquina para reconhecer e classificar as emoções no sistema límbico?

Diversos sentimentos podem ser detectados em eletroencefalogramas (EEG)¹. Portanto, este trabalho tem como objetivo utilizar métodos de aprendizado de máquina para reconhecer padrões e classificar emoções humanas a partir de segmentos de EEG. (ESPERIDIÃO-ANTONIO et al., 2008).

2 TRABALHOS RELACIONADOS

Murugappan (2011) constata a importância das emoções como característica dos humanos e apresenta a classificação das emoções humanas usando sinais de EEG. Os sinais do EEG foram coletados de 20 voluntários. Os voluntários foram submetidos a estímulos audiovisuais para induzir emoções específicas e, após a aquisição dos sinais de EEG, os sinais foram pré-processados, utilizando uma filtragem de superfície Laplaciana e através da Transformada de Wavelet, o sinal foi decomposto em 5 diferentes bandas (delta, teta, alfa, beta e gama). Em seguida, foi aplicada a validação cruzada de 5 *folds*, com o método de classificação *K*-vizinho mais próximo. Como resultado, obteve-se uma taxa de acurácia média de 82,87% para o conjunto com 62 canais e de 78,57% no conjunto com 24 canais.

Bhardwaj et al. (2015) argumenta que a detecção de emoções tem sido um grande campo de pesquisa nos últimos anos e apresenta a classificação de emoções humanas usando sinais de EEG. O uso do *Independent Component Analysis* (ICA) e técnicas de aprendizado de máquinas, tais como o *Support Vector Machine* (SVM) e *Linear Discriminant Analysis* (LDA), foram usados para classificar sete emoções humanas, a média das acurácia dos modelos foi de 74,13% utilizando o modelo SVM e 66,50% utilizando o modelos LDA. A média das acurácia foram obtida após ser calculado a acurácia dos 4 *folds* da validação cruzada.

Chakladar e Chakraborty (2018) também destaca a importância que detecção de emoções tem no estudo sobre a interface cérebro-computador. Fazendo uso de sinais de EEG obtidos por meio de múltiplos eletrodos, foi aplicado a seleção de um subconjunto baseado em correlação para a redução de dimensionalidade. Em seguida, foi utilizado as características de *High Order Statistic* (HOS) sobre o conjunto reduzido de canais para a classificação de quatro classes (Positivo, Negativo, Raiva e Harmonia). A acurácia encontrada usando esse

¹ Eletroencefalograma é o resultado do monitoramento de atividades cerebrais através de eletrodos posicionados ao redor do escalpo (FREEMAN; QUIROGA, 2012).



conjunto reduzido de canais foi de 82%.

Neste trabalho foi feito uso de três modelos preditivos para classificação com o propósito de comparação entre os mesmos. Além disso, as técnicas de extração de características foram aplicadas para obter apenas atributos no domínio do tempo. A base de dados utilizada além de ser atual (ONTON, 2021) apresenta o intervalo de tempo em que as emoções são "sentidas" de acordo com os voluntários, possibilitando a segmentação de EEG no exato momento em que a emoção começa e termina. Adicionalmente, a classificação é realizada sobre o conjunto de dados com duas classes (Emoções Positivas e Negativas) e sobre o conjunto de oito classes de emoções. Após isso, foram realizados testes estatísticos de hipótese para verificar a existência de diferença estatisticamente significativa entre os modelos.

3 MATERIAIS E MÉTODOS

3.1 Base de dados

Foram utilizados EEG de 20 voluntários da base de dados da *Imagined Emotion Study* (ONTON, 2021). Esses segmentos foram coletados considerando os seguintes procedimentos: os voluntários escutam gravações de voz que sugerem uma determinada emoção, esses voluntários são instruídos a imaginar ou relembrar uma experiência emocional que reflete a mesma emoção sugerida pela gravação; e quando os sujeitos são tomados pela emoção indicada, os mesmos, pressionam um botão e mantêm pressionados até que o sentimento acabe. Os indivíduos foram instruídos a "sentir" as emoções por um período de tempo de 3 a 5 minutos, entretanto, não existia um indicador de tempo externo para os voluntários, dessa forma, o tempo de duração da emoção foi definido pela percepção do início e fim da emoção pelo indivíduo ao pressionar o botão.

O EEG de cada voluntário foi obtido usando o sistema internacional 10–20 (JASPER, 1958), possuindo um total de 224 canais e taxa de amostragem de 256 Hz. Além disso, foram removidos os sinais de eletrocardiograma (ECG) e o eletrooculograma (EOG), para que não influenciasse no aprendizado de máquinas. Os EEG foram segmentados no período de tempo em que cada emoção teve seu início e fim para cada indivíduo, por classe (emoção), como temor, frustração, alegria, raiva, felicidade, tristeza, amor e excitação.

3.2 Extração de características

Para a representação de segmentos de EEG em formato adequado para a aplicação de métodos de aprendizado de máquina, para cada segmento de EEG foram computadas as seguintes medidas: média, desvio padrão, pico, amplitude, quartis, amplitude de quartil, curtose, assimetria, coeficiente de variação, raiz quadrática média, fator de crista, canal e emoção (OLIVA; ROSA, 2021). Desse modo, foram extraídas, no total, 16 características no domínio do tempo pelo fato de não haver necessidade de aplicar nenhuma transformada no sinal, reduzindo o custo computacional no processo de construção de modelos preditivos.

Após isso, foram selecionados os dados de 44 canais por serem comuns a todos os 20 indivíduos, foi aplicado extração de características, para cada canal, obtendo um total de 704 características.



3.3 Construção de modelos preditivos

As características extraídas (conforme a Seção 3.2), foram submetidas a três modelos de classificação, os quais são, K-ésimo vizinho mais próximo (KNN), Árvore de decisão e Floresta Aleatória.

K-Vizinho mais Próximo é um método não paramétrico que se adapta à quantidade de suavização da densidade de dados em um determinado local. O nível da suavização depende de K , que é o número de vizinhos mais próximos, de forma que K é muito menor que N , o tamanho total das amostras (ALPAYDIN, 2020).

Árvore de Decisão é uma estrutura hierárquica que é gerada utilizando a estratégia de dividir para conquistar em termos de seus atributos, até atingir o nó folha que representa a decisão. A árvore de decisão é um método não paramétrico eficiente, usando tanto para regressão quanto para classificação (ALPAYDIN, 2020).

Floresta Aleatória é um método de aprendizado tanto para classificação como para regressão, que cria várias Árvores de Decisão por meio da seleção aleatória de atributos (HO, 1995).

Duas configurações foram consideradas para os três métodos apresentados, de forma que, ao todo, foram construídos seis modelos preditivos:

- 2 Classes: as emoções foram rotuladas entre emoções boas e emoções ruins.
- 8 Classes: todas as emoções registradas na base de dados.

Além disso, durante a criação dos modelos, foi realizado a otimização de hiper-parâmetros para os modelos preditivos utilizando validação cruzada com 5 *folds* para obter os hiper-parâmetros dos modelos que apresenta a melhor acurácia média.

3.4 Avaliação de modelos

O desempenho dos modelos preditivos foi avaliado utilizando validação cruzada k -fold, que consiste na divisão dos dados em 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 et al., 1995).

Neste trabalho foi considerada a validação cruzada para 5 *folds*. Como resultado, foram obtidos os valores da média e desvio padrão das acurácias calculadas para cada *fold*. Além disso, foram aplicados testes estatísticos de hipótese de *Friedman* sobre os dados pareados para verificar se existe diferença estatisticamente significativa entre os modelos preditivos.

Para desenvolvimento deste trabalho, as seguintes ferramentas utilizada: Python, uma das linguagens mais utilizadas para aprendizado de máquina e possui grande quantidade de bibliotecas para auxiliar no treinamento dos modelos; e bibliotecas como NumPy, Pandas, Scipy e Sklearn para a extração de características e aplicação de métodos de aprendizado de máquina.

4 RESULTADOS

Na Tabela 1, são apresentados os valores de Acurácia Média (AM) e Desvio Padrão (DP) referentes ao desempenho dos modelos preditivos, para os quais foram considerados dois tipos de problemas: duas e oito classes.

**Tabela 1 – Acurácia dos modelos utilizados**

	Árvore de Decisão		Floresta Aleatória		KNN	
	AM (%)	DP (%)	AM (%)	DP (%)	AM (%)	DP (%)
2 Classes	80,5	12,0	83,9	11,7	76,0	9,3
8 Classes	38,0	6,2	44,5	3,4	28,5	2,7

Fonte: Autoria própria (2021).

Para o problema de duas classes, o modelo Floresta Aleatória atingiu a maior AM (83,9%). Por outro lado, o classificador kNN obteve o menor AM (76%) e de DP (9,3%). Para o problema de oito classes, o modelo Floresta Aleatória atingiu a maior AM (44,5%). Por outro lado, o classificador kNN obteve o menor AM (28,5%) e de DP (2,7%).

O modelo de Floresta Aleatória por aplicar múltiplas árvores de decisão com atributos aleatórios, alcançou maiores valores de AM em comparação com outros métodos.

Os valores de acurácia de modelos construídos para a diferenciação entre oito classes são menores em comparação com os classificadores binários. O desempenho considerado inferior ocorreu por diversos motivos, por exemplo a quantidade de dados por classe é menor e os atributos obtidos não são suficiente para a classificação de oito classes.

Com os resultados das classificações, foi necessário investigar a existência de diferença estatisticamente significativa entre os desempenhos dos modelos preditivos. Sendo assim, como os dados das amostras são pareados, foi aplicado o teste de Friedman para as duas configurações considerando o nível de significância em 5%, cujos resultados são apresentados na Tabela 2.

Tabela 2 – P-valores do Teste de Friedman

	Teste de Friedman
2 Classes	0.2096
8 Classes	0.0083

Fonte: Autoria própria (2021).

De acordo com a Tabela 2, para a configuração utilizando 2 classes, não foi encontrada diferença estatisticamente significativa entre os resultados. No entanto, para a classificação utilizando 8 classes, foi evidenciada diferença estatisticamente significativa. Sendo assim, foi aplicado o pós-teste de Nemenyi para determinar quais pares de classificadores apresentam essa diferença.

O resultado do teste de Nemenyi mostrou que a diferença estatisticamente significativa foi constatada apenas entre os modelos Floresta Aleatória e kNN, em cuja comparação foi obtido um p -valor de 0,007. Nesse caso, podemos afirmar, com 95% de certeza, que a Floresta Aleatória atingiu o melhor desempenho em comparação com o kNN para a diferenciação entre oito classes.

5 CONCLUSÕES

Este trabalho tem como objetivo utilizar métodos de aprendizado de máquina para reconhecer padrões e classificar emoções humanas por meio de segmentos de EEG. Na avaliação experimental deste trabalho foram alcançados resultados promissores para classificação binária com acurácia média mais alta de 83,9%, não demonstrando diferença estatisticamente significativa entre os modelos. Entretanto, para a classificação envolvendo oito classes, o desempenho foi considerado inferior, com acurácia média mais alta de 44,5%, pois a quantidade de dados de aprendizado para cada classe é menor. Porém, foi constatada diferença estatisticamente



significativa entre Floresta Aleatória e K -Vizinho mais Próximo.

Par a melhorar o desempenho dos modelos preditivos, como trabalho futuro, podem ser implementadas características dos domínios de frequência e de tempo-frequência. Adicionalmente, outros métodos de aprendizado de máquina podem ser utilizados.

AGRADECIMENTOS

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

REFERÊNCIAS

- ALPAYDIN, E. **Introduction to machine learning**. Cambridge, EUA: MIT press, 2020.
- BHARDWAJ, A. et al. Classification of human emotions from EEG signals using SVM and LDA Classifiers. In: SPIN. Noida, Índia: IEEE, 2015. P. 180–185.
- CHAKLADAR, D. D.; CHAKRABORTY, S. EEG based emotion classification using “Correlation Based Subset Selection”. **Biologically inspired cognitive architectures**, Elsevier, v. 24, p. 98–106, 2018.
- ESPERIDIÃO-ANTONIO, V. et al. Neurobiologia das emoções. **Archives of Clinical Psychiatry (São Paulo)**, SciELO Brasil, v. 35, p. 55–65, 2008.
- FAIRHURST, M.; E., M.; LI, C. Enhancing the forensic value of handwriting using emotion prediction. In: IWBF. Valletta, Malta: IEEE, 2014. P. 1–6.
- FREEMAN, W.; QUIROGA, R. Q. **Imaging brain function with EEG: advanced temporal and spatial analysis of electroencephalographic signals**. Nova York, EUA: Springer Science & Business Media, 2012.
- HO, T. K. Random decision forests. In: ICDAR. Montreal, Canadá: IEEE, 1995. v. 1, p. 278–282.
- HUTTO, C.; GILBERT, E. Vader: A parsimonious rule-based model for sentiment analysis of social media text. In: 1. ICWSM. Michigan USA: AAAI, 2014. v. 8.
- JASPER, Herbert. Report of the committee on methods of clinical examination in electroencephalography. **Electroencephalogr Clin Neurophysiol**, v. 10, p. 370–375, 1958.
- KOHAVI, R. et al. A study of cross-validation and bootstrap for accuracy estimation and model selection. In: 2. IJCAI. Montreal, Canadá: IEEE, 1995. v. 14, p. 1137–1145.
- LEON, E. et al. Real-time Physiological Emotion Detection Mechanisms: Effects of Exercise and Affect Intensity. In: ACEMB. Shanghai, China: IEEE, 2005. P. 4719–4722.
- MURUGAPPAN, Muthusamy. Human emotion classification using wavelet transform and KNN. In: ICPAIR. Kuala Lumpur, Malásia: IEEE, 2011. v. 1, p. 148–153.
- OLIVA, J. T.; ROSA, J. R. G. Binary and multiclass classifiers based on multitaper spectral features for epilepsy detection. **Biomedical Signal Processing and Control**, Elsevier, v. 66, p. 102469, 2021.
- ONTON, J. **"Imagined Emotion Study"**. Califórnia, EUA: OpenNeuro, 2021.