

# CLASSIFICAÇÃO DE SINAIS DE EEG COM MODELOS DE REGRESSÃO FUNCIONAL

Rodrigo Marcel Araujo Oliveira

Orientadora: Florencia Graciela Leonardi

Instituto de Matemática e Estatística - Universidade de São Paulo

e-mail: rodrigo.marcel.oliveira@alumni.usp.br; florencia@usp.br

## Objetivos

A eletroencefalografia (EEG) é um dos melhores métodos para avaliar a atividade elétrica cortical. O sinal do EEG pode ser resultado da atividade espontânea do cérebro ou pode estar relacionado com eventos cerebrais sensoriais, motores e cognitivos [1]. Esse presente projeto de pesquisa tem como objetivo estudar técnicas de processamento de sinais, tais como as transformadas de *Fourier* e transformadas de *Wavelet*, para decomposição do sinal do EEG, e avaliar o desempenho de modelos de Regressão Funcional para predição de novos dados.

## Métodos e Procedimentos

A análise espectral permite a identificação de fontes de interferência e proporciona uma forma rápida e eficiente de identificar as componentes de um sinal. As bases da análise de *Fourier* são ondas senoidais, e, portanto, o sinal é analisado como um todo. As *Wavelets* decompõem o sinal em versões escalonadas e deslocadas de sua *Wavelet* original, elas tendem a serem irregulares e assimétricas [2].

Na regressão logística funcional [3] a probabilidade  $p_i$  da ocorrência de um evento binário cujo  $Y_i = 1$  condicional a um preditor funcional  $X_i(t)$  e coeficiente funcional  $B(t)$  é expressa conforme a equação 1:

$$P(Y_i = 1 | X_i(t): t \in T) = \frac{e^{\alpha + \int_T X_i(t)B(t) dt}}{1 + e^{\alpha + \int_T X_i(t)B(t) dt}} \quad (1)$$

com  $i = 1, \dots, n$ .

O uso de funções de base e regularização em  $X_i(t)$  contribui para reduzir o viés induzido por erro de medição [4].

A avaliação quantitativa da pesquisa consiste na aplicação de quatro métodos diferentes. As metodologias 1 e 2 são seguidas respectivamente por: transformada de *Wavelet*; aplicação da função de base; regressão logística funcional. Já para as metodologias 3 e 4 a transformada de *Fourier* é adotada no lugar da *Wavelet*. A primeira e a terceira metodologia consideram a função de base de *Fourier*, já a segunda e a quarta são aplicadas a função de base de *Spline*. Essas abordagens foram aplicadas em dois conjuntos de dados de EEG, relacionados a problemas de reconhecimento de emoções e detecção de movimentos motores.

A base *Emotion* [5] consiste em analisar sinais de EEG de participantes enquanto eles jogavam rodadas de jogos de azar. O subconjunto fornecido contém observações agregadas de 23 participantes. Para nosso estudo o objetivo será explicar os potenciais no sinal EEG para condutividade da meta correspondente ao resultado monetário (ganho ou perda) no final de cada rodada do jogo. A base contém uma dimensão de (184, 384) para as covariáveis e (1, 184) para as variáveis categóricas. Para modelagem, separamos a base em treino (70 %) e teste (30 %).

Para base *SelfRegulationSCP1* [6] o experimento deste conjunto de dados consiste em avaliar se o sujeito está aumentando ou diminuindo sua lentidão cortical potencial, isto é, se o sujeito moveu o cursor para cima ou para baixo. As gravações foram feitas com 6 canais

de EEG em 256 Hz, o que resultou em 896 amostras por canal para cada tentativa. A base de treino contém uma dimensão de (268, 896) para covariáveis e (1, 268) para variável resposta, já para base de teste temos respectivamente, (293, 896) e (1, 293).

Para avaliar o desempenho dos modelos ajustados foram utilizadas algumas métricas derivadas da matriz de confusão, como a sensibilidade, especificidade e acurácia [7].

## Resultados

A comparação do desempenho dos modelos está sintetizada nas Tabelas 1 e 2.

Tabela 1: Resultados no conjunto para base *Emotion*

Metodologia	Emotion			Acurácia IC 95%
	Sensibilidade	Especificidade	Acurácia	
1	0.77	0.58	0.68	(0.54, 0.80)
2	0.48	0.70	0.59	(0.45, 0.72)
3	0.39	0.46	0.43	(0.30, 0.57)
4	0.43	0.50	0.46	(0.33, 0.60)

Tabela 2: Resultados no conjunto *Self Regulation*.

Metodologia	Self Regulation			Acurácia IC 95%
	Sensibilidade	Especificidade	Acurácia	
1	0.88	0.88	0.87	(0.82, 0.91)
2	0.91	0.86	0.88	(0.84, 0.92)
3	0.52	0.51	0.52	(0.46, 0.57)
4	0.54	0.52	0.53	(0.47, 0.59)

É notável que, o melhor desempenho para o conjunto de dados *Emotion* na base teste foi utilizando a metodologia 1, isto é, aplicando a transformada de *Wavelet* com a função de Base de *Fourier* (composta por 8 componentes), obteve-se um valor de acurácia de 68 %. Para o conjunto de dados do *Self Regulation* a metodologia 2 destaca-se com o melhor desempenho obtido, isto é, aplicando a transformada de *Wavelet* com a função de Base de *Fourier* (composta por 4 componentes) e ajustando a Regressão Logística Funcional para os canais 1, 2, 3, 4 e 5. A acurácia do modelo foi de 88 % na base de teste.

## Conclusões

A hipótese inicial do projeto de pesquisa é avaliar o desempenho de modelos de regressão funcional para predição de novos dados, utilizando técnicas de processamento de sinais, tais como as transformadas de *Fourier* e transformadas de *Wavelet*, para decomposição do sinal do EEG. Nesse trabalho vimos que os modelos de regressão funcional com auxílio da transformada de *Wavelet* e transformada de *Fourier* são técnicas estatísticas e matemáticas promissoras para avaliação de dados de EEG.

Comparando os resultados do conjunto *Self Regulation* com os obtidos em [2] utilizando modelos de aprendizado de máquina nota-se um incremento de 2 pontos percentuais no resultado da acurácia, o que indica que a abordagem de trabalhar com dados funcionais aplicando os modelos de regressão funcional para sinais de EEG é satisfatório.

## Referências Bibliográficas

- [1] I. Gannaz, 'Classification of EEG recordings in auditory brain activity via a logistic functional linear regression model.', pp. 125–130, Jun. 2014, Accessed: Aug. 26, 2022. [Online]. Available: <https://hal.archives-ouvertes.fr/hal-00830313>
- [2] L. Alípio, 'Unraveling the Brain: a Quantitative Study of EEG Classification Techniques', 2021.
- [3] M. Febrero-Bande and M. O. de la Fuente, 'Statistical Computing in Functional Data Analysis: The R Package *fda.usc*', *J Stat Softw*, vol. 51, no. 4, pp. 1–28, Oct. 2012, doi: 10.18637/JSS.V051.I04.
- [4] J. S. Morris, 'Functional Regression', Jun. 2014, Accessed: Aug. 26, 2022. [Online]. Available: <http://arxiv.org/abs/1406.4068>
- [5] 'emotion: EEG and EMG recordings in a computerised gambling study in *fdboost*/*FDboost*: Boosting Functional Regression Models'. <https://rdr.io/github/fdboost/FDboost/main/emotion.html> (accessed Aug. 26, 2022).
- [6] 'Time Series Classification Website'. <http://www.timeseriesclassification.com/description.php?Dataset=SelfRegulationSCP1> (accessed Aug. 26, 2022).
- [7] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, 'An improved method to construct basic probability assignment based on the confusion matrix for classification problem', *Inf Sci (N Y)*, vol. 340–341, pp. 250–261, May 2016, doi: 10.1016/J.INS.2016.01.033.