

CLASSIFICAÇÃO DE SINAIS DE EEG COM MODELOS DE REGRESSÃO FUNCIONAL

Discente: Rodrigo Marcel Araujo Oliveira

Orientadora: Florencia Graciela Leonardi

AGENDA

Introdução

- EEG

Objetivos

Metodologia

- Análise espectral
- Análise de dados funcionais
- Regressão Logística Funcional
- Avaliação dos modelos
- Conjunto de dados
- Estratégias

Resultados

Conclusões

Considerações

Referências

Contato

- O eletroencefalograma (EEG) é o exame que registra a atividade elétrica cerebral [1].
- Técnica muito importante para avaliação neurofisiológica de pacientes com distúrbios do sono, morte cerebral, tumores, infecções cerebrais, epilepsia, etc.

Figura 1 - EEG

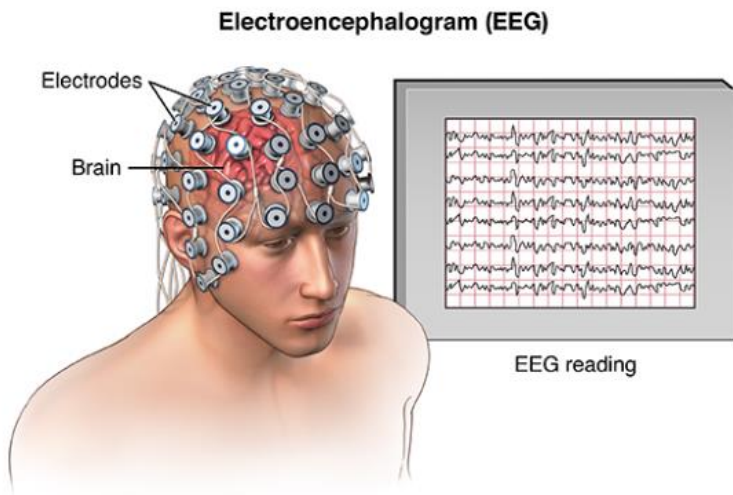


Figura 2 - Sistema 10-2

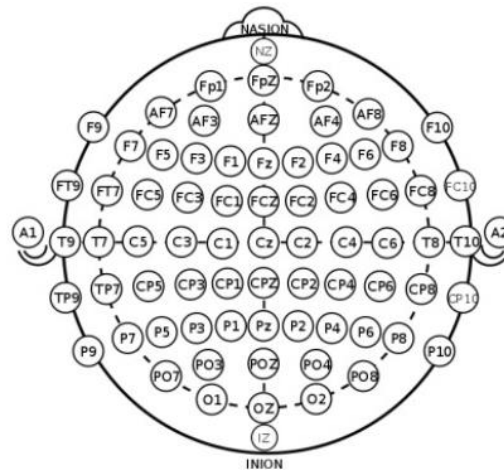
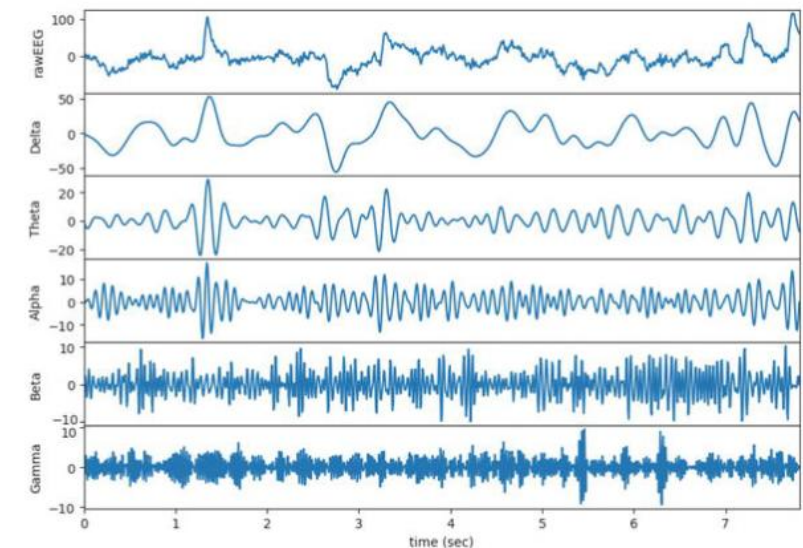


Figura 3 - Ondas cerebrais de acordo com sua frequência



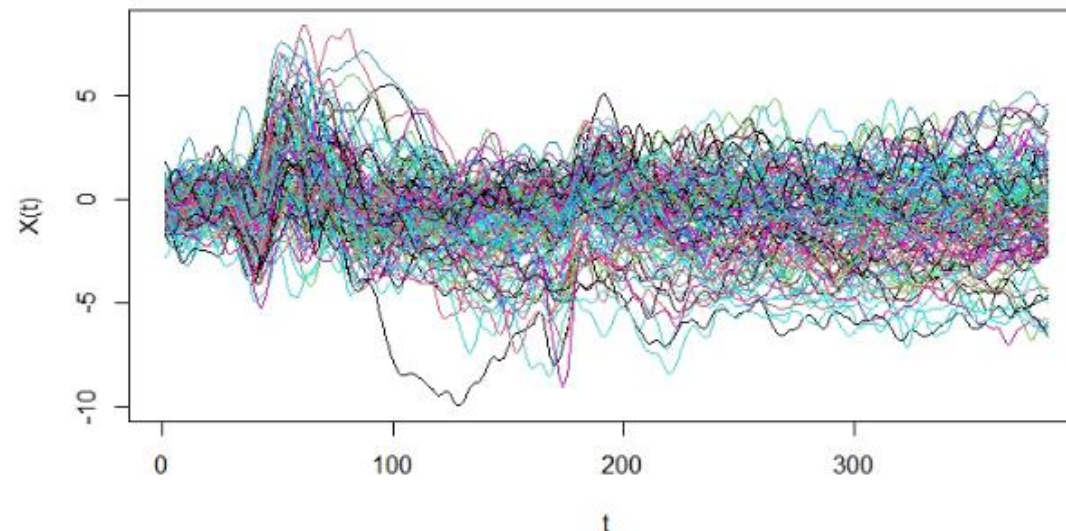
- Estudar técnicas de processamento de sinais, tais como as transformadas de Fourier e transformadas de Wavelet, para decomposição do sinal do EEG.
- Avaliar o desempenho de modelos de Regressão Funcional para predição de novos dados.

- A análise de dados funcionais (FDA) é um ramo da estatística preocupado com a análise de dados na forma de funções [3].
- Modelo Linear Funcional:

$$Y_i = B_0 + \int X_i(t)B(t)dt + \epsilon_i \quad (1)$$

- As funções de bases (Fourier e Spline) são os blocos de construção do FDA e determinam o mecanismo pelo qual a regularização é feita.

Figura 10 - Sinais de EEG ao longo do tempo.



- A análise espectral permite a identificação de fontes de interferência e proporciona uma forma rápida e eficiente de identificar as componentes de um sinal [2].

- Transformada de Fourier:

$$S(t) = a_0 + \sum_{n \in \mathbb{N}} a_n \cos n\omega t + b_n \sin n\omega t \quad (2)$$

- Transformada de Wavelet:

$$W_{s,\tau} = \frac{1}{\sqrt{s}} \int x(t) \psi^*\left(\frac{t-\tau}{s}\right) dt = \int x(t) \psi_{s,\tau}^*(t) dt \quad (3)$$

- Na Regressão Logística Funcional [4] a probabilidade p_i da ocorrência de um evento binário cujo $Y = 1$ condicional a um preditor funcional $X_i(t)$ e coeficiente funcional $B(t)$ é expressa conforme a equação 3:

$$p_i = \mathbf{P}[Y = 1 \mid X_i(t) : t \in T] = \frac{\exp\{\alpha + \int_T X_i(t)B(t)dt\}}{1 + \exp\{\alpha + \int_T X_i(t)B(t)dt\}} \quad (4)$$

com $i = 1, \dots, n$

- Técnicas de amostragem: Hold-out [7].
- Matriz cruzada [7].
- Métrica de avaliação: acurácia (ACU); sensibilidade (SEN); especificidade (ESP); etc.

Figura 12 - Matriz cruzada

		Valor Predito	
		Positivo	Negativo
Valor Real	Positivo	n_{11}	n_{12}
	Negativo	n_{21}	n_{22}

Figura 13 - Métricas de avaliação

$$SEN = \frac{n_{11}}{n_{11} + n_{12}}$$

$$ESP = \frac{n_{22}}{n_{21} + n_{22}}$$

$$ACU = \frac{n_{11} + n_{22}}{n_{11} + n_{12} + n_{21} + n_{22}}$$

Self Regulation

- 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.
- 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).

Emotion

- A base Emotion [5] consiste em analisar sinais de EEG de participantes enquanto eles jogavam rodadas de jogos de azar. O objetivo será explicar os potenciais no sinal EEG para condutividade da meta correspondente ao resultado monetário (ganho ou perda).
- 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 %).

- **Metodologia 1:** Transformada de Wavelet + Função de Base de Fourier + Regressão Logística Funcional
- **Metodologia 2:** Transformada de Wavelet + Função de Base de Spline + Regressão Logística Funcional
- **Metodologia 3:** Transformada de Fourier + Função de Base de Fourier + Regressão Logística Funcional
- **Metodologia 4:** Transformada de Fourier + Função de Base de Spline + Regressão Logística Funcional

Self Regulation: base de validação

Tabela 1 - Metodologia 1

Desempenho do Modelo			
Accuracy	0.8703	Sensitivity	0.8844
95% CI	(0.8264, 0.9066)	Specificity	0.8844
No Information Rate	0.5017	Pos Pred Value	0.8609
P-Value [Acc >NIR]	<2e-16	Neg Pred Value	0.8803
		Prevalence	0.5017
Kappa	0.7406	Detection Rate	0.4437
		Detection Prevalence	0.5154
Mcnemar's Test P-Value	0.6265	Balanced Accuracy	0.8703

• **Tabela 3 - Metodologia 3**

Desempenho do Modelo			
Accuracy	0.5154	Sensitivity	0.5170
95% CI	(0.4565, 0.5739)	Specificity	0.5137
No Information Rate	0.5017	Pos Pred Value	0.5170
P-Value [Acc >NIR]	0.3413	Neg Pred Value	0.5137
		Prevalence	0.5017
Kappa	0.0307	Detection Rate	0.2594
		Detection Prevalence	0.5017
Mcnemar's Test P-Value	1.0000	Balanced Accuracy	0.5154

Tabela 2 - Metodologia 2

Desempenho do Modelo			
Accuracy	0.884	Sensitivity	0.9116
95% CI	(0.8416, 0.9183)	Specificity	0.8562
No Information Rate	0.5017	Pos Pred Value	0.8645
P-Value [Acc >NIR]	<2e-16	Neg Pred Value	0.9058
		Prevalence	0.5017
Kappa	0.7679	Detection Rate	0.4573
		Detection Prevalence	0.5290
Mcnemar's Test P-Value	0.2299	Balanced Accuracy	0.8839

Tabela 4 - Metodologia 4

Desempenho do Modelo			
Accuracy	0.5324	Sensitivity	0.5442
95% CI	(0.4735, 0.5907)	Specificity	0.5205
No Information Rate	0.5017	Pos Pred Value	0.5333
P-Value [Acc >NIR]	0.1603	Neg Pred Value	0.5315
		Prevalence	0.5017
Kappa	0.0648	Detection Rate	0.2730
		Detection Prevalence	0.5119
Mcnemar's Test P-Value	0.8643	Balanced Accuracy	0.5324

Emotion: base de validação

Tabela 5 - Metodologia 1

Desempenho do Modelo			
Accuracy	0.6786	Sensitivity	0.7667
95% CI	(0.5404, 0.7971)	Specificity	0.5769
No Information Rate	0.5357	Pos Pred Value	0.6765
P-Value [Acc >NIR]	0.02122	Neg Pred Value	0.6818
		Prevalence	0.5357
Kappa	0.3472	Detection Rate	0.4107
		Detection Prevalence	0.6071
McNemar's Test P-Value	0.4795	Balanced Accuracy	0.6718

Tabela 6 - Metodologia 2

Desempenho do Modelo			
Accuracy	0.5893	Sensitivity	0.4828
95% CI	(0.4498, 0.719)	Specificity	0.7037
No Information Rate	0.5179	Pos Pred Value	0.6364
P-Value [Acc >NIR]	0.1747	Neg Pred Value	0.5588
		Prevalence	0.5179
Kappa	0.1848	Detection Rate	0.2500
		Detection Prevalence	0.3929
McNemar's Test P-Value	0.2109	Balanced Accuracy	0.5932

Tabela 7 - Metodologia 3

Desempenho do Modelo			
Accuracy	0.4286	Sensitivity	0.3929
95% CI	(0.2971, 0.5678)	Specificity	0.4643
No Information Rate	0.5000	Pos Pred Value	0.4231
P-Value [Acc >NIR]	0.8856	Neg Pred Value	0.4333
		Prevalence	0.5000
Kappa	-0.1429	Detection Rate	0.1964
		Detection Prevalence	0.4643
McNemar's Test P-Value	0.8597	Balanced Accuracy	0.4286

Tabela 8 - Metodologia 4

Desempenho do Modelo			
Accuracy	0.4643	Sensitivity	0.4286
95% CI	(0.3299, 0.6026)	Specificity	0.5000
No Information Rate	0.5000	Pos Pred Value	0.4615
P-Value [Acc >NIR]	0.7478	Neg Pred Value	0.4667
		Prevalence	0.5000
Kappa	-0.0714	Detection Rate	0.2143
		Detection Prevalence	0.4643
McNemar's Test P-Value	0.8551	Balanced Accuracy	0.4643

Tabela 9 - Comparação dos desempenhos dos modelos

Metodologias	Self Regulation				Emotion			
	Sensibilidade	Especificidade	Acurácia	Acurácia IC 95%	Sensibilidade	Especificidade	Acurácia	Acurácia IC 95%
1	0.88	0.88	0.87	(0.82, 0.91)	0.77	0.58	0.68	(0.54, 0.80)
2	0.91	0.86	0.88	(0.84, 0.92)	0.48	0.70	0.59	(0.45, 0.72)
3	0.52	0.51	0.52	(0.46, 0.57)	0.39	0.46	0.43	(0.30, 0.57)
4	0.54	0.52	0.53	(0.47, 0.59)	0.43	0.50	0.46	(0.33, 0.60)

- 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.

- Análise disponível em: <https://github.com/roaraujo/EEG-Regressao-Funcional-Wavelet-Fourier>

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.archivesouvertes.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/man/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

Obrigado

Contato: rodrigo.marcel.oliveira@alumni.usp.br