CLASSIFICATION OF EEG SIGNALS WITH FUNCTIONAL REGRESSION MODELS

Rodrigo Marcel Araujo Oliveira

Supervisor: Florencia Graciela Leonardi

Instituto de Matemática e Estatística - Universidade de São Paulo

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

Objectives

Electroencephalography (EEG) is one of the best methods to assess cortical electrical activity. The EEG signal may be a result of spontaneous brain activity or may be related to sensory, motor and cognitive brain events [1]. This present research project aims to study signal processing techniques, such as Fourier transforms and Wavelet transforms, to decompose the EEG signal, and to evaluate the performance of Functional Regression models for predicting new data.

Materials and Methods

Spectral analysis allows the identification of interference sources and provides a quick and efficient way to identify the components of a signal. The bases of Fourier analysis are sine waves, and therefore the signal is analyzed as a whole. Wavelets decompose the signal into staggered and offset versions of its original Wavelet, they tend to be irregular and asymmetrical [2].

In functional logistic regression [3] the probability p_i of the occurrence of a binary event whose $Y_i = 1$ conditional on a functional predictor $X_i(t)$ and functional coefficient $\boldsymbol{B}(t)$ is expressed according to equation 1:

$$\mathrm{P}(\boldsymbol{Y}_i = 1 | \boldsymbol{X}_i(t) : t \in T) = \frac{e^{\alpha + \int_T X_i(t)B(t)\,dt}}{1 + e^{\alpha + \int_T X_i(t)B(t)\,dt}} \qquad (1)$$

for $i = 1, \ldots, n$.

The use of basis and regularization functions in $X_i(t)$ contributes to reduce the bias induced by measurement error [4].

The quantitative evaluation of the research consists of the application of four different methods. Methodologies 1 and 2 are followed respectively by: Wavelet transform; application of the base function; functional logistic regression. For methodologies 3 and 4, the Fourier transform is adopted instead of the Wavelet. The first and third methodologies consider the Fourier basis function, while the second and fourth are applied to the Spline basis function. These approaches were applied to two EEG datasets related to emotion recognition and motor movement detection problems.

The Emotion base [5] consists of analyzing EEG signals from participants while they are playing rounds of games of chance. The subset provided contains aggregated observations from 23 participants. For our study, the objective will be to explain the potentials in the EEG signal for target conductivity corresponding to the monetary result (win or loss) at the end of each game round. The base contains a dimension of (184, 384) for the covariates and (1, 184) for the categorical variables. For modeling, we separated the base into training (70%) and testing (30%).

For *SelfRegulationSCP1* [6] the experiment with this dataset consists of evaluating whether the subject is increasing or decreasing his potential cortical slowness, that is, whether the subject has moved the cursor up or down. Recordings were made with 6 EEG channels at 256 Hz, which resulted in 896 samples per channel for each trial. The training base contains a dimension of (268, 896) for covariates and (1, 268) for the response variable, while for the test

base we have, respectively, (293, 896) and (1, 293).
To evaluate the performance of the adjusted models, some metrics derived from the confusion matrix were used, such as sensitivity, specificity and accuracy [7].

Results

The comparison of the performance of the models is summarized in Tables 1 and 2.

Table 1: Results in the Emotion set

Methods	Emotion			
	Sensitivity	Specificty	Accuracy	Accuracy IC 95%
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)

Table 2: Results in the Self Regulation set.

Methods	Self Regulation			
	Sensitivity	Specificty	Accuracy	Accuracy IC 95%
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)

It is notable that the best performance for the Emotion dataset in the test base was using methodology 1, that is, applying the Wavelet transform with the Fourier Base function (consisting of 8 components), a value of 68% accuracy. For the Self Regulation dataset, methodology 2 stands out with the best performance obtained, that is, applying the Wavelet transform with the Fourier Base function (composed of 4 components) and adjusting the Functional Logistic Regression for the channels 1, 2, 3, 4 and 5. The model's accuracy was 88% on the test basis.

Conclusions

The initial hypothesis of the research project is to evaluate the performance of functional regression models for predicting new data, using signal processing techniques, such as Fourier transforms and Wavelet transforms, to decompose the EEG signal. In this work we saw that the functional regression models with the aid of the Wavelet transform and Fourier transform are promising statistical and mathematical techniques for evaluating EEG data. Comparing the results of the Self Regulation set with those obtained in [2] using machine learning models, there is an increase of 2 percentage points in the accuracy result, which indicates that the approach of working with functional data by applying the regression models functional for EEG signals is satisfactory.

References

[1] I. Gannaz, "Classification of EEG recordings in auditory brain activity via a logistic functional linear regression model.", p. 125–130, jun. 2014, Acessado: ago. 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 e M. O. de la Fuente, "Statistical Computing in Functional Data Analysis: The R Package fda.usc", *J Stat Softw*, vol. 51, nº 4, p. 1–28, out. 2012, doi: 10.18637/JSS.V051.I04.

[4] J. S. Morris, "Functional Regression", jun. 2014, Acessado: ago. 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://rdrr.io/github/fdboost/FDboost/man/emotion.html (acessado ago. 26, 2022).

[6] "Time Series Classification Website". http://www.timeseriesclassification.com/description.php?Dataset=SelfRegulationSCP1 (acessado ago. 26, 2022).

[7] X. Deng, Q. Liu, Y. Deng, e 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, p. 250–261, maio 2016, doi: 10.1016/J.INS.2016.01.033.