



## 2

# Interface Cérebro-Computador

**Objetivo:** Compreender o modelo do *perceptron* e seus parâmetros. Utilizar técnicas clássicas de processamento de sinais para analisar o sinal no tempo e na frequência, extraindo as informações de interesse.

As técnicas de processamento de sinais e *Machine Learning* podem desempenhar um papel fundamental na identificação das atividades neurológicas, possibilitando o controle de dispositivos e aplicações por meio de sinais transmitidos diretamente do cérebro.

Neste projeto, você deve desenvolver um sistema capaz de distinguir para qual fonte luminosa o indivíduo está direcionando sua atenção.

Para isto, vamos empregar uma base de dados pública, descrita no artigo: *Benchmark Dataset for SSVEP-Based Brain-Computer Interfaces*. As informações mais relevantes para o desenvolvimento deste projeto são:

- i. O sinal cerebral foi adquirido utilizando 64 sensores empregando uma taxa de amostragem de 250 Hz.
- ii. Para cada indivíduo há 6 s de coleta de sinal cerebral, resultando em 1500 amostras. Destes, os primeiros 500 ms são de pré-estímulo e os 500 ms finais são de pós estímulo, estas amostras devem ser descartadas neste projeto.
- iii. Foram realizadas 6 repetições para cada um dos 40 estímulos visuais.
- iv. As frequências dos 40 estímulos vão de 8 Hz a 15.8 Hz, com um salto 0.2 Hz, ou seja, 8 Hz, 8.2 Hz, 8.4 Hz, ..., 15.8 Hz.

## 2.1 DESENVOLVIMENTO E QUESTÕES

1. Baixe os sinais cerebrais de um indivíduo disponível em: <http://bci.med.tsinghua.edu.cn/download.html>.

A base de dados é organizada em um array 4-D de dimensão [64, 1500, 40, 6], sendo:

- **Dimensão 1** faz a identificação do eletrodo. Você deve fixar a leitura no eletrodo 61 (Oz).
- **Dimensão 2** apresenta as amostras do sinal cerebral. Você deve trabalhar com as amostras de 125 a 1375, descartando os momentos pré e pós estímulo.



- **Dimensão 3** se refere a frequência do estímulo visual. Escolha duas frequências inteiras de estimulação para representar as duas possíveis classes do problema. O mapeamento para as frequências é:

$0 \rightarrow 8 \text{ Hz}$ ,  $1 \rightarrow 9 \text{ Hz}$ ,  $2 \rightarrow 10 \text{ Hz}$ ,  $3 \rightarrow 11 \text{ Hz}$ ,  $4 \rightarrow 12 \text{ Hz}$ ,  $5 \rightarrow 13 \text{ Hz}$ ,  $6 \rightarrow 14 \text{ Hz}$  e  $7 \rightarrow 15 \text{ Hz}$

- **Dimensão 4** se refere as 6 repetições coletadas.

2. Organize sua matriz de dados de entrada  $\mathbf{X}$  e seu vetor de rótulos  $y$  da seguinte maneira:

$$\mathbf{X} = \begin{bmatrix} x[61, amostras, f_1, 0] \\ x[61, amostras, f_1, 1] \\ \vdots \\ x[61, amostras, f_1, 5] \\ \vdots \\ x[61, amostras, f_2, 0] \\ x[61, amostras, f_2, 1] \\ \vdots \\ x[61, amostras, f_2, 5] \end{bmatrix}$$

e

$$y = \begin{bmatrix} +1 \\ \vdots \\ -1 \end{bmatrix}$$

com  $\mathbf{X} \in \mathbb{R}^{12 \times 1250}$  e  $y \in \mathbb{R}^{12 \times 1}$ , com o rótulo '+1' para as entradas da classe  $f_1$  e rótulo '-1' para as entradas da classe  $f_2$ .

3. Implemente o *perceptron* que receba em sua entrada a matriz de dados  $\mathbf{X}$  e retorne o valor de  $y$ , indicando a frequência de estimulação  $f_1$  ou  $f_2$ .

Separe aleatoriamente 4 amostras de cada classe para treinar o *perceptron* e o restante das amostras para validar o sistema.

4. Como os valores de inicialização do vetor de pesos, taxa de aprendizado e demais parâmetros do modelo impactam o desempenho do sistema? Ilustre apresentando algumas configurações do modelo e o desempenho obtido.
5. Analisando o desempenho do modelo nas simulações realizadas, o problema tratado é linearmente separável? Justifique.
6. Apresente a matriz de confusão da considerando somente as amostras do conjunto de validação do melhor modelo perceptron obtido.
7. Plote uma das linhas da matriz de entrada para cada classe e sua respectiva magnitude da transformada de Fourier (FFT). O que é possível observar no tempo e na frequência em cada um dos gráficos obtidos?
8. Proponha uma técnica de extração de características, de modo que não seja necessário empregar todas as 1250 amostras do sinal realizada em cada coleta, mas que seja possível reduzir a dimensão da matriz de dados  $\mathbf{X}$  sem perder a informação relevante.