MARENGO Matteo
matteo.marengo@ens-paris-saclay.fr

EXAM MEEG
MVA 2023-2024

02/21/24

First, we load the raw data for EEG for a subject. We extract the epochs that are data segments corresponding to events for the tasks ”hands” and ”feet”. For example, the EEG signal for the channel Fc5 is displayed in Fig 1. Indeed, as a recall, when doing an EEG measurement the channels are displayed like in Fig 2. In this practical, there will be a study of patients doing MI vs rest. MI is Motor Imagery where patients imagine performing a movement without actually executing it. It is a cognitive process that can trigger activation in the motor cortex similar to actual movement.

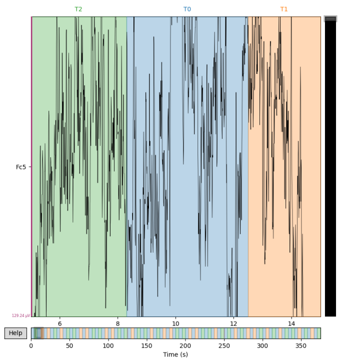

Figure 1: EEG signal for the FC5 channel over time.

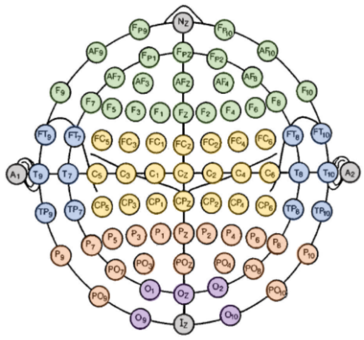

Figure 2: Different channels for EEG instrumentation.

1 Part 1 - Connectivity and Networks

1.1 Task 1: Connectivity matrices

The first task is **to compute and plot the connectivity matrices based on the imaginary coherence averaged over the mu band and across the epochs**. The results for both conditions (hands and feet) are displayed in Fig 3 and Fig 4.

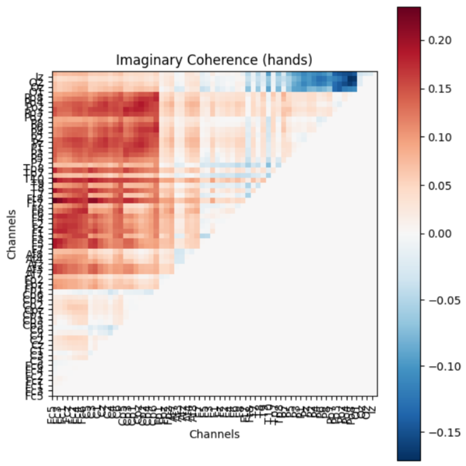

Figure 3: Connectivity Matrix for Hands

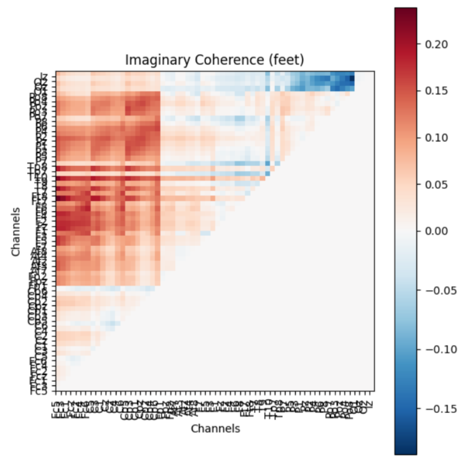

Figure 4: Connectivity Matrix for Feet

This method focuses on measuring the imaginary part of coherence between signals from different brain regions providing a metric for assessing **functional connectivity**. Coherence is a frequency-domain measure that quantifies the degree of correlation between two signals at specific frequencies (in our case mu band).

$$\text{coh}_{xy}(\omega) = \frac{\frac{1}{n}\sum_{k=1}^{n} A_x(\omega,k)A_y(\omega,k)e^{i(\phi_x(\omega,k)-\phi_y(\omega,k))}}{\sqrt{\left(\frac{1}{n}\sum_{k=1}^{n} A_x^2(\omega,k)\right)\left(\frac{1}{n}\sum_{k=1}^{n} A_y^2(\omega,k)\right)}} \tag{1}$$

The imaginary part of coherence is particularly interesting for brain connectivity studies as it is insensitive to zero-lag correlations. The mu band is between 8 and 13 Hz.

There are few modifications between the two connectivity matrices even if the upper channels appears less activated in feet than in hands. Some potential intensities are changing but there is not an important trend. Overall F channels have a high intensity whereas parietal ones have a lower one. Indeed PO or parietal occipital are more involved in processing **visual information and sensory input**. It is also a good result as it could prove that even with different tasks the same area of the brain will be activated as hands and feet are close. Indeed, the motor areas responsible for controlling hand and foot movements are both located in the primary motor cortex.

1.2 Task 2: Node strength

The second task is to compute the **associate node strength across the epochs**. Node strength refers to the sum of the weights of all connections linked to a particular node. It is a measure of how connected a particular node is within the network. Higher values indicate that the node is more central within the mapped network of brain activity. This is represented for the feed and hand in Fig 5. When looking at it we can conclude that the potential is similar in the channels in the two cases as suggested by the imaginary coherence matrices.

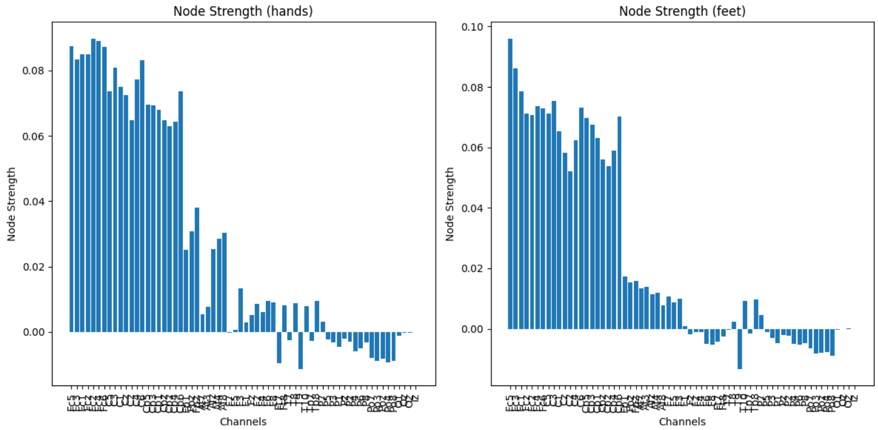

Figure 5: Node strength for Hands and Feet

Moreover, we can analyze that the channels with the highest node strength are the ones that were studied during the lectures as shown in Fig 6. We can analyze that channels like CF3 or C3 have a high node strength whereas channel like P09 has a low node strength. It is coherent as electrodes C3 and C4 are located in the sensorimotor areas.

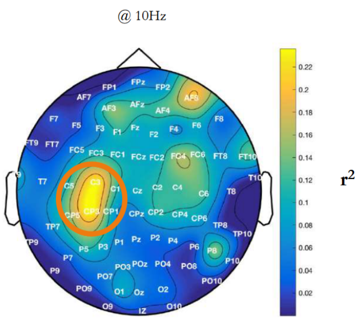

Figure 6: Active zones during MI. Retrieved from the lectures.

1.3 Task 3: Statistical Difference

In Fig 5 there is the plot of the statistical difference between MI and Rest conditions obtained from imaginary coherence (left) and the results obtained with the node strength (right). We observe that it is neurophysiologically meaningful as nearly the same areas of the brain are connected, the ones that are related to the sensorimotor areas (C3 and C4). However during MI FC channels or fronto central channels are activated. Indeed, it is known that they are predominant in cognitive tasks or motor execution. On the other hand, at rest PO channels are activated as it is more focused on sensory inputs and visual ones.

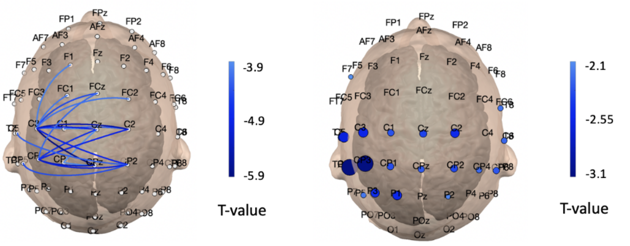

Figure 7: Statistical Difference MI vs Rest obtained from imaginary coherence and node strength.

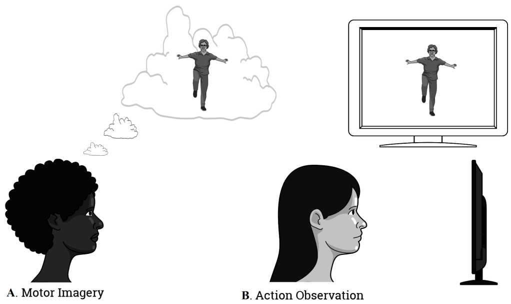

Figure 8: MI vs Rest

2 Part 2 - Features in BCI

This part displays Event-Related Desynchronization/Synchronization (ERDS) maps from EEG data during the task of the hands (Fig 9) and the task of the feet (Fig 10). ERDS maps are used to visualize changes in brain activity over time and frequency (in our case hands and foot movement).

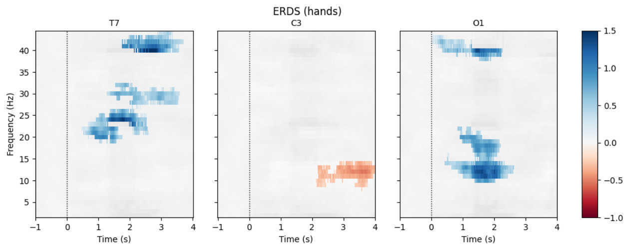

Figure 9: ERDS hands

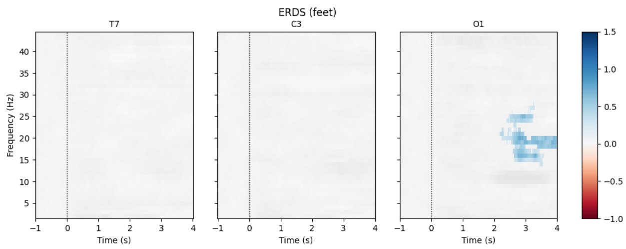

Figure 10: ERDS feet

2.1 Observations from the maps

O1 channel during Feet Imagery task: There is a clear Event-Related Synchronization (ERS, shown in blue) in the mu band. This is likely reflecting a slowdown in the visual cortex as O1 is over the occipital lobe and the task is not visually driven. However C3 has not an ERD which means that maybe for feet movement the frequency band is an other one.

C3 Channel during Hands Imagery Task: The ERD here suggests activation of the sensorimotor cortex consistent with imagining hand movements. The ERS in the alpha band over the occipital cortex during the hands imagery task could indicate a relaxed state as the visual cortex is not engaged.

2.2 BCI performance

Mu band ERD around the sensorimotor cortex informs on BCI performance for MI vs Rest tasks. As with these patterns detected, we can conclude on the usage of the relation between the human and the computer.

2.3 If I was the experimenter ...

Given the ERD observed over C3 during the hand's imagery task, this electrode and the mu band frequency range have to be selected. Even if the activity over O1 during the feet imagery task is not directly related to motor imagery, it could be used as a baseline.

3 Part 3 - Machine Learning & BCI

Two classification pipelines **(CSP+LDA: Common Spatial Patterns + LDA, RG+LR: Riemannian Geometry + Logistic Regression)** have been defined to classify a BCI dataset. Their performances from a dataset of two subjects have been plotted in Fig 11 and in Fig 12.

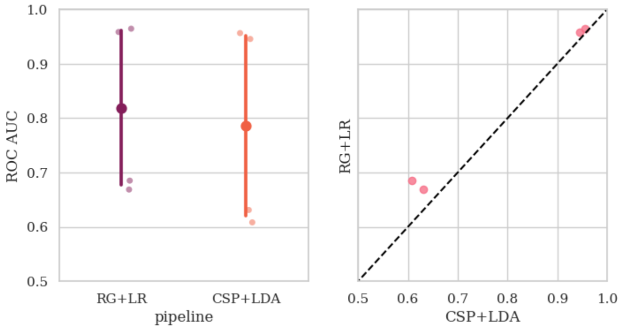

Figure 11: Performances of the pipelines

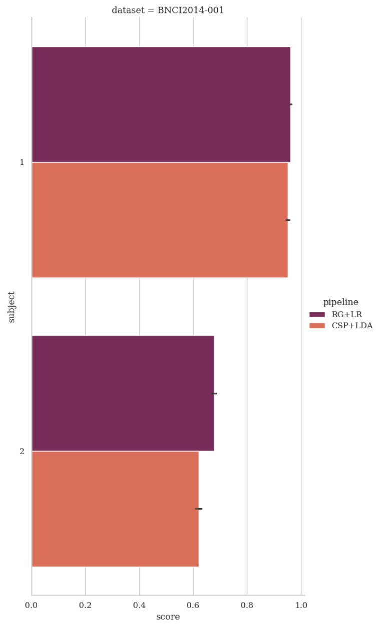

Figure 12: Individual distribution of performance

3.1 Observations from the plot

- The first plot shows the ROC AUC for both pipelines. The AUC scores for RG + LR are higher than for CSP+LDA, indicating better performance in distinguishing between classes. The variability in performance across subjects is notable. While CSP+LDA consistently outperforms RG+LR, the individual subject performance varies, with some subjects showing a high AUC close to 1 and others closer to 0.5.

- The scatter plot compares the performance of the two pipelines for individual subjects. Most points lie above the diagonal line it emphasizes the fact that RG+LR outperforms CSP+LDA.

- The bar plot shows the score for each pipeline, separated by subject. Once again, RG+LR performs better across the two subjects.

3.2 Proposed Framework for Feature Extraction, Selection, and Classification

- Time-frequency analysis to capture the **power spectral density** changes over time, which could reveal more nuanced features related to ERD/ERS phenomena. Advanced methods like Wavelet Transform could be applied to capture both time and frequency characteristics of EEG signals

- Recursive feature elimination could be applied to identify the most informative features and reduce dimensionality. Common techniques such as PCA are also an idea. Finally, unsupervised feature learning using autoencoders could be used to learn data representations.

- Ensemble methods such as Random Forests or Gradient Boosting, can capture non-linear relationships. Support Vector Machines (SVM) with non-linear kernels, which are powerful for binary classification problems typical in BCIs. Deep learning approaches, specifically Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), can automatically extract features and model temporal dependencies.

3.3 Performance Assessment

Employing K-fold cross-validation to ensure the model's robustness and generalizability across different subsets of data. Analyzing true positives, false positives, true negatives, and false negatives to understand specific areas where the classifier might be underperforming.

3.4 BONUS - new framework

Based on these ideas, a new pipeline has been implemented, one with SVM for classification instead of LDA or Logistic Regression and a pipeline based on LedoitWolf covariances merged with RandomForestClassifier. Results and comparison with other methods are displayed in Fig 13. We can conclude that finally, these two methods do not perform better than the suggested ones, to enhance our framework, maybe a more robust feature selection should be applied such as **deep learning methods**.

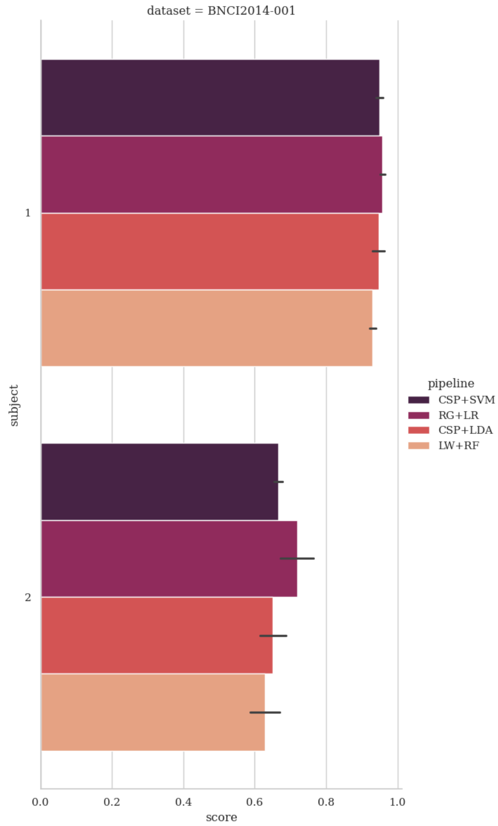

Figure 13: Scores with the implemented pipelines

4 Part 4 - Experimental considerations

4.1 New protocol based on EEG acquisitions

4.1.1 Types of Artifacts in EEG

The two main types of artifacts in EEG data are physiological and non-physiological artifacts.**Physiological Artifacts** originate from the subject's own body functions. One common example is eye movement or blink artifacts. Eye movements can create large voltage fluctuations in the EEG signal. **Non-Physiological Artifacts** are caused by external sources. A typical example is line noise, which is a constant electrical signal coming from the power supply. Some examples are displayed in Fig 14.

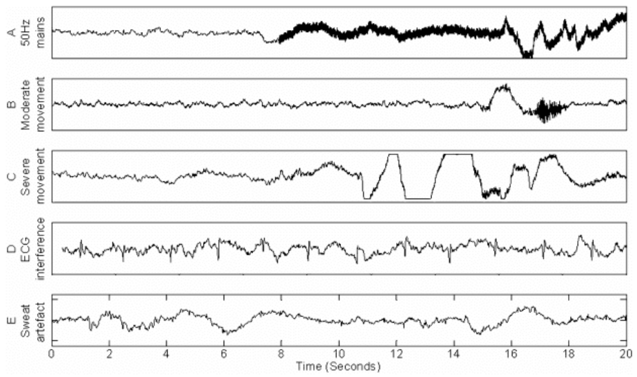

Figure 14: Some effects of artifacts in EEG.

4.1.2 Main Steps in an EEG Processing Pipeline

An EEG processing pipeline has multiple steps. First, the data acquisition. Then, the pre-processing of the data (filter, epoching to segment the continuous EEG signal into relevant time windows around events that are worth analyzing such as MI). Feature Extraction is important (which channel, which frequency band for example). Then we extract the feature with machine learning methods. And then we do the classification. In the lectures the pipeline was defined a bit differently with data inspection / artifacts removal / source reconstruction and finally advanced analysis. The proposed pipeline in lectures is the one in Fig 15.

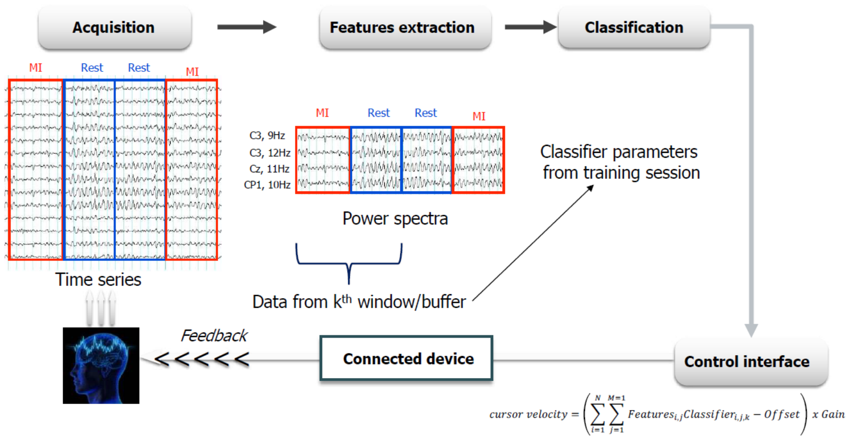

Figure 15: Pipeline EEG processing

4.2 Conducting an experimental protocol in BCI

Given that the subject shows a drop in performance after four training sessions, here are some suggestions. First, we can make the instructions clearer to the subject and be confident that he understands what is the task he has to do. It goes along an extensive communication, even during the experiment to adjust some parameters. Discuss with the subject their mental strategy for motor imagery and experiment with different strategies to see if performance improves. Although the same features have been used consistently, it might be beneficial to reassess the feature set or change the frequency range. Finally, we have to be sure that the patient is not tired or that an event is not stressing him out.

5 Conclusion

To sum up, this lab has been to visualize results studied in the lectures and implement ourselves some processing steps for EEG measurements. It helps us to work on real data to understand what steps we should take to work on this with care and tediousness.