MINI PROJECT 3

For this mini-project, we propose 3 variants and you will do only the one your group was assigned to. Each of the 3 variants contains tasks that you should address and integrate in your report/presentation.

Groups AA, AC, C, E, H, K, X will present on Thursday 14/12 during the exercise session.
Presentations should last **10 min + 5 min of QAs**.
The grade within the group might vary by max ¼ according to the presentation skills.

All the other groups should provide us with **a report of 3 pages max** (including images), single column, font-size 11px in Arial. Make sure to shape the structure in a paper-like format: Introduction, Methods, Results and Discussion sections + **the code** you used for solving the mini-project. Both should be uploaded on Moodle.

Report+Code are due on **Thursday 14th December at 16:00**.
Note that this is a sharp deadline, no extensions are allowed.

Everyone is encouraged to come to the presentations.

NX-421: Neural signals and signal processing

Variant 1

Groups: AA, E, X, AB, F, J, N

Description

This mini-project is focused on the development of a regression-based model to analyze EMG data for hand movement prediction. The objective is to construct a model employing regression methods for precise mapping of EMG signals to specific hand gestures and movements. The significance of this research lies in its application in prosthetic and rehabilitative technology. Accurate interpretation of EMG signals is crucial for enhancing the functionality of prosthetic limbs.

Dataset:

The dataset NinaPro Dataset 8 (you can download it and see a detailed description here: (https://ninapro.hevs.ch/instructions/DB8.html) is designed for the estimation and reconstruction of finger movements, focusing on decoding finger positions from contralateral EMG (electromyography) measurements using regression algorithms, rather than for motion or grip classification. The dataset includes data from ten able-bodied individuals and two right-hand transradial amputee participants. These subjects were asked to repeat different hand movements using both hands, which were mirrored bilaterally. The experiment included nine different movements, ranging from single-finger movements to more complex functional grips, such as cylindrical and tripod grips. Participants performed these movements following visual cues, with each movement lasting between 6 to 9 seconds, interspersed with 3 seconds of rest.

Data collection involved EMG and inertial measurement units (IMUs) using 16 active double-differential wireless sensors from a Delsys Trigno IM Wireless EMG system. These sensors, which included EMG electrodes and 9-axis IMUs, were positioned around the participants' right forearm. Hand kinematic data was recorded using a Cyberglove 2 worn on the left hand, contralateral to the arm with EMG sensors. The Cyberglove signals were then synchronized with the EMG and IMU data.

Tasks

1. Visualize the data and preprocess them accordingly. Split the dataset into training, validation and test sets. Why do we need the different datasets? How do you decide where to split them?

2. Perform sliding windows (choose a reasonable window width and sliding step) and explain your choice.

3. Extract a set of features, normalize them and visualize the correlation between them. What do you observe?

4. Perform a regression on the kinematics (use a method of your choice). Visualize and comment on the performance of the regressor.

5. Repeat task 3 and 4 after performing feature selection. Use at least 2 different methods for feature selection and compare the performances. Discuss the obtained results. Use the selected features for the following tasks.
6. Perform again the regression by adding to the selected window features the predicted kinematic data from the previous window. Does the regression performance improve? Compute feature importance, are the past kinematic data important?
7. Is the regression performance stable across the different DataGlove trackers? If you observe any substantial differences, can you try to explain why?

NX-421: Neural signals and signal processing

Variant 2

Groups: AC, H, B, G, L, O, W

Description
Classifiers, a subset of machine learning algorithms, enable researchers to predict an output state by analyzing input data. Specifically for amputees, these classifiers can interpret electromyography (EMG) data to predict intended actions. This is crucial in controlling robotic prosthetic arms, offering amputees the potential to regain certain functional abilities.

Dataset:
In this mini-project, you will use NinaPro Dataset 1 https://ninapro.hevs.ch/instructions/DB1.html) to investigate the **generalization of movement classification from EMG signals across different subjects**. More details could be found on the website. Briefly, the participants are tasked to replicate the movement shown on the screen. sEMG signal is recorded while the participants are performing the tasks.

Tasks

1. Visualize and preprocess the data. Split the data into training, validation, and testing sets for each subject. Why do we need the different datasets?

2. Perform sliding windows (choose a reasonable window width and sliding step) and explain your choice.

3. Extract the same set of features across 10 different subjects. Look at the typical values of those features across the same set of movements for different subjects. What do you see? Are there any regularities between the different subjects? What are some possible reasons for similarity/dissimilarity?

4. Extract the same set of features across 10 different subjects. Look at the typical values of those features across the same set of movements for different subjects. What do you see? Are there any regularities between the different subjects? What are some possible reasons for similarity/dissimilarity?

5. Perform classification (use a method of your choice) on different subjects **separately**. Perform analysis to determine the importance of the features to the classification. Compare this ranking across the different subjects. Are the features stable?

6. Train a classification model on a set of subjects and test it on a subject that does not belong to that set. Evaluate the performance. How does it compare to training on that subject and testing on the same subject directly?

7. Repeat task 5 by varying the number of subjects in the training set. Discuss how the number of subjects used in the training set could affect the classification performance.

NX-421: Neural signals and signal processing

Variant 3

Groups: C, K, D, I, M, P, Z

Description

Blind source separation (BSS) methods are used to isolate individual signals from a mixed set without prior knowledge of the sources. Applied to electromyography (EMG), BSS helps in dissecting EMG data into separate motor unit potentials. This is helpful for studying EMG synergy stability, allowing us to investigate muscle coordination and control, which is important for developing better prosthetics.

Dataset:

The dataset is available at: https://doi.org/10.6084/m9.figshare.c.5090861.v1. Briefly, EMG signals were captured from 20 able-bodied volunteers from 128 channels, placed on their forearms. These participants executed 65 distinct movements in an isometric fashion. An automated recording protocol was employed to meticulously time the hand movements, ensuring simultaneous recording of both the EMG signals and the forces exerted by the hand joints.

Tasks

1. Visualize and preprocess the data.

2. Compute the **spatial** synergies using different blind source separation algorithms (e.g. PCA, ICA, NMF). Are the synergies stable?

3. Compute the spatial synergies corresponding to different subsets of the data samples. Are the synergies stable across the different subsets?

4. Compute the spatial synergies for different subjects. Is it now stable across the different subjects?

5. Try to vary the frequencies used in the preprocessing step. Compute the spatial synergies again. Does the choice of filtering frequencies affects the stability of the synergies?

6. Evaluate the stability of the synergies across the different BSS algorithms with respect to the tasks 3-4-5. Which algorithm performs better, and which one is least stable. Discuss the pros and cons of the different algorithms.