NX-421: Neural signals and signal processing

MINI PROJECT 2

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

The hand-in format is that of recorded presentations (video). Recordings should last a **maximum of 5 min**. The deliverables should also include a **written report** and a **code repository**.

Recordings Guidelines:

- Use zoom recording option
- The speech needs to be shared evenly across group members (i.e. each member should speak)
- Presenting flow:
  - Briefly present the dataset and the motivation of the mini project
  - Answer succinctly each point and question

Written reports Guidelines:

- **5 pages max** (including images), single column, font-size 11px in Arial, single line spacing.
- Optional **appendix 1 page** (same font size and type)
- **Submit the code** you used for solving the mini-project.
- Please, do not upload datasets, images or any additional content

All deliverables should be submitted on Moodle.

Deliverables are due on **Thursday 19th December at 16:00**.
Note that this is a sharp deadline, no extensions are allowed.

NX-421: Neural signals and signal processing

Description

Electromyography (EMG) data is often used to predict the intended action of amputees and stroke patients. A good intention decoder is crucial in controlling robotic prostheses and other assistive devices. Through this project, students will apply knowledge obtained from the class to process and analyze EMG signals.

Part 1: Single subject classification (subject 1)

Suggested time allocation 29th November – 5th December

In this part of the mini-project, you will use NinaPro Dataset 1 https://ninapro.hevs.ch/instructions/DB1.html) to classify and predict the movement classes associated to the different EMG signals for a single subject. More information regarding the tasks and experimental setup can be found on the website. Briefly, the participants are tasked to replicate the movement shown on the screen. Surface EMG signals are recorded while the participants perform the tasks.

Please complete the following tasks and answer the questions:

1. Visualize and preprocess the data for subject 2 (use only the first exercise set: **S2_A1_E1** with 12 actions). Does the data look reasonable? Did you need to remove any trials?

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

3. Extract features from the trials (at least 5 different ones). Look at the typical values of the features across the different trials. What do you see? Are the values similar between repetitions and between channels? Explain the possible reasons for similarity/discrepancies.

4. Perform **classification** on the data of subject 1. Predict the action of the subject based on the EMG signals. Use hyperparameter optimization to increase your models' performance. You will use **one of the algorithm families** throughout the whole project (meaning for part 1, 2 and 3) **depending on your project variant**:

Project Variant	Algorithm Family
A	Support vector machine
B	Random forest
C	Gradient boosting

5. Evaluate the performance using a metric of your choice. Justify why the metric is suitable for this task and whether the performance is satisfactory.

6. Perform feature selection / dimension reduction using **2 methods** of your choice. Evaluate the performance using the same metric as point 5. Is there an improvement in the performance?

Part 2: Generalization across subjects

Suggested time allocation 5th December – 12th December

In this part of the mini-project, you will continue working with the NinaPro Dataset 1 (https://ninapro.hevs.ch/instructions/DB1.html) to classify and predict the movement classes associated with different EMG signals. The focus will now shift to evaluating whether the prediction model can generalize across different subjects. Specifically, the objective is to assess if training the model on data from a group of individuals improves its performance on new subjects. This is important because it helps develop models that accommodate individual differences, ensuring robust performance without requiring extensive retraining for each person.

1. Consider now 27 subjects (only exercise set 1 as in part 1). Perform the same preprocessing pipeline as in part 1.
2. Extract the same set of features. Now look at the values of those features across the different subjects. Are there any similarity/discrepancies between the subjects?
3. Train a **classification** model (same model as part 1) on 26 subjects (aggregating the data) and test it on a subject outside the train set. Evaluate the performance. How does it compare to training and testing on the same subject directly?
4. Perform cross-validation by rotating the subject used for testing. Specifically, for each fold, select one subject as the test set and use the remaining 26 subjects to train the model. Repeat this process for all 27 subjects, ensuring each subject is used as the test set exactly once. Evaluate and compare the performance across all folds to understand how well the model generalizes to new subjects.
5. Repeat task 3 by varying the number of subjects in the training set (without performing cross-validation). For example, use subject 27 for testing, and train with subject 1, then train with subjects 1 and 2, then 1, 2, and 3, and continue to increase number of subjects used for training. Analyze how the size of the training set impacts the classification performance on the test subject. Discuss the potential effects of including more or fewer subjects in the training data, focusing on trends in generalization and accuracy.

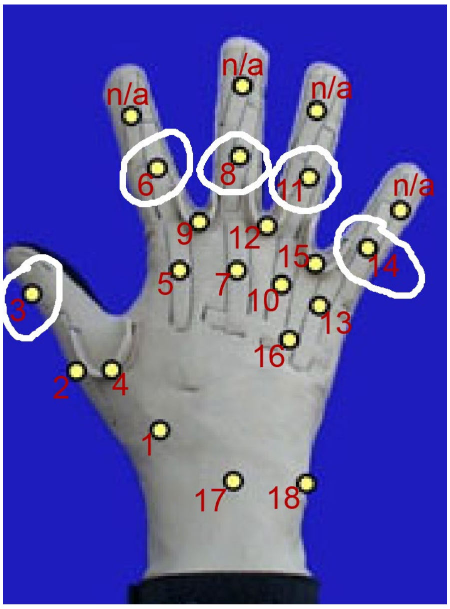
NX-421: Neural signals and signal processing

Part 3: Regression for joint angles

Suggested time allocation 12th December – 19th December

In this part of the mini-project, you will be working on **subject 1** of the NinaPro Dataset 8 ([https://ninapro.hevs.ch/instructions/DB8.html,](https://ninapro.hevs.ch/instructions/DB8.html) use only the file **S1_E1_A1.mat**). Instead of focusing on classification, you will implement a regression algorithm to predict joint angles. The goal is to enhance the control of the robotic hand, enabling more precise movements such as achieving half-closed positions.

1. Visualize the data and preprocess them accordingly. Use the joint angles 3,6,8,11,14 (refer to picture below). Split the dataset into training, validation and test sets. Take note of how the splitting is done for a time-series data.

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 the same family of regression method of the previous part). Visualize and comment on the performance of the regressor.

5. Evaluate the performance using a metric of your choice. Justify why the metric is suitable for this task and whether the performance is satisfactory.

6. Is the regression performance stable across the different finger angles? If you observe any substantial differences, can you explain why?