Neural Signals and Signal Processing Project 2

Jean Ciardo, Romain Frossard, Hugo Jeannin, Mathieu Verest
École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Introduction

Electromyography (EMG) data is crucial for advancing assistive technologies, particularly in intention decoding for robotic prostheses and devices supporting amputees and stroke patients. This project explores the development of predictive models using EMG signals to enhance functionality and usability. Two objectives guide this work: first, classifying movement classes to evaluate subject-specific performance and generalization across individuals; second, implementing regression algorithms to predict joint angles for precise robotic hand movements. By addressing these challenges, this study aims to improve the adaptability, accuracy, and robustness of human-machine interaction systems.

NinaPro Dataset 1 Description

The dataset used for Parts 1 and 2 of this project comes from the NINAPRO database. As described by Atzori et al. [2015], this dataset includes EMG signals on which our analyses will focus. See the Appendix for further information.

Part 1: Single subject classification

Question 1

The dataset for subject 2 comprises EMG signals recorded across 10 channels with 101,014 time points. Labels for 12 movements and 10 repetitions are provided for each time point. To assess signal quality, we qualitatively analyzed the time-domain EMG signals and their Power Spectral Density (PSD), confirming that the preprocessing steps performed during data acquisition, including filtering and rectification, were adequate.

To evaluate signals for each trial, as we will use them separately, we segmented the data by movements and repetitions, and a moving average filter was applied to extract signal envelopes. This enabled qualitative signal assessment per trial. As no abnormalities were detected, all trials and channels were retained for further analysis.

Question 2

The data is split into training, validation, and testing sets to ensure robust model development and evaluation. The training set allows the model to learn patterns in the data, building its predictive capabilities. The validation set is used during training to fine-tune hyperparameters and assess the model's performance, helping to aboid overfitting to the training data. Finally, the testing set contains unseen data and is used to evaluate the model's generalization ability. This separation ensures that the model performs well not only on the training data but also on new unseen data, making it reliable for practical applications.

Question 3

We decided to extract six time-domain features from the EMG signals to capture meaningful information for movement classification. The chosen features are Mean Absolute Value (MAV), Standard Deviation (STD), Maximum Absolute Value (MaxAV), Root Mean Square (RMS), Waveform Length (WL), and Slope Sign Changes (SSC).These features represent key aspects of the signal, such as amplitude, variability, energy, and frequency content. Detailed descriptions and mathematical definitions of these features are provided in the Appendix.

We extracted these features for each trial and visualized them as heatmaps (1 heat map per type of movement) displayed in 1. In these heatmaps, rows represent the activity recorded by each channel in response to repeated stimuli, while columns correspond to each repetition of the movement.

We observe that the rows generally appear consistent in intensity for most features, indicating that the same movements recorded by the same channel are reproducible and stable. This is expected, as the same movement should induce similar muscle activity. However, regarding the columns, the patterns are less evident. This can be explained by the fact that different muscles are not equally recruited during specific movements (e.g., agonist and antagonist muscles), the impedance of each channel may vary, and not all muscles have the same intensity potential due to differences in muscle size, fiber composition, and depth relative to the electrodes. In summary, the stability observed in the rows reflects consistent muscle activation for identical movements, whereas the variability in the columns highlights individual muscle recruitment differences and technical factors like electrode impedance

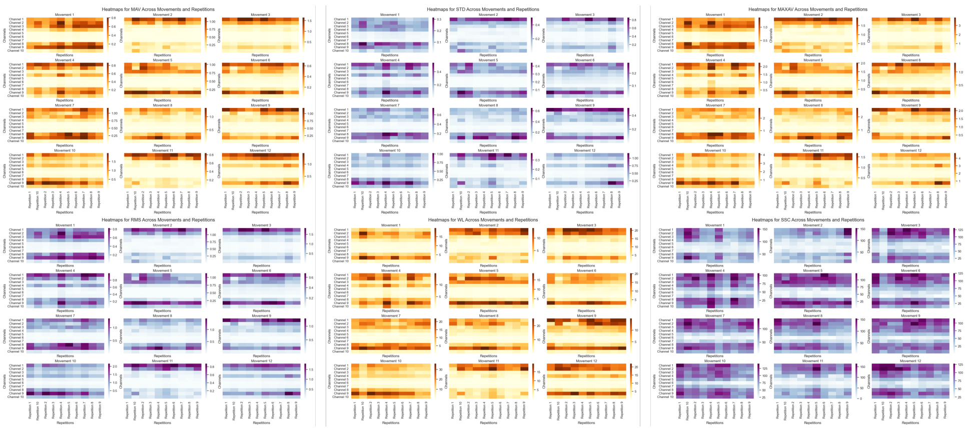


Figure 1: Heatmaps for all extracted Features Across Movement and Repetitions

Question 4

We performed EMG movement classification for Subject 2 using a Support Vector Machine (SVM) model with hyperparameter optimization. The dataset was divided into three subsets, maintaining class proportions: 60% for training, 20% for validation, and 20% for testing. To ensure consistent feature scaling, the data was standardized. Three key hyperparameters were optimized: the regularization parameter (C), kernel type, and gamma, which together control the model's complexity and decision boundary. Optimization was conducted using a 5-fold cross-validation grid search on the training set, exploring various hyperparameter combinations to identify the best configuration. The optimal parameters, which maximized the performance metric (we define in the next question), were determined to be C $= 0.1$, a linear kernel, and gamma set to scale.

Question 5

We used accuracy to evaluate our model's performance, as it measures the proportion of correctly classified samples, providing a clear indication of the model's overall predictive capability across all classes. Given that the labels in this dataset are uniformly distributed, accuracy effectively reflects the model's true performance across all classes. The accuracy is calculated as $Accuracy = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$. Using the optimized hyperparameters described earlier, the model achieved a validation accuracy of 0.92 (and a weighted F1-score of 0.91). Final evaluation on the test set yielded a test accuracy of 0.96 (and a weighted F1-score of 0.96), demonstrating the model's strong generalization ability and effectiveness in accurately inferring movements for a single subject.

Question 6

We explored two dimensionality reduction methods: k-best feature selection and Principal Component Analysis (PCA). The k-best method identified the top k features most strongly associated with the target variable (movement) using mutual information. On the validation set, selecting $k = 41$ features achieved the best accuracy of 1.00, and testing this model on the test set yielded an accuracy of 0.958. This method effectively reduced the feature set from 60 to 41 without compromising performance.

PCA, on the other hand, captured the maximum variance in the dataset by projecting it onto fewer dimensions. Using 9 principal components (explaining 92.5% of the total variance) was already enough to maximized, the validation accuracy was 0.917, but the test accuracy dropped to 0.833.

The choice between these methods depends on the objectives. The k-best method provides a moderate feature reduction while maintaining high accuracy, making it ideal for applications prioritizing performance. PCA offers a more significant reduction in dimensionality, which can be advantageous for resource efficiency, though it comes at the cost of reduced classification performance. In our case, prioritizing accuracy makes k-best feature selection the more suitable approach.

Part 2: Generalization across subjects

Question 1

The same preprocessing steps as in Part1-Question1 were performed on all 27 subjects.

Question 2

The same set of features as in Part1-Question3, was extracted. For each stimulus, we retained only the average of each feature across the 10 repetitions. Subsequently, we computed the variance across the subject dimension for each channel-stimulus combination, resulting in six heatmaps. We see that, for nearly all features, movement 5 on channel 7 varies a lot accross subjects, whereas channels 4-5-6 show similarity between subjects. The SSC is the feature with the most diversity.

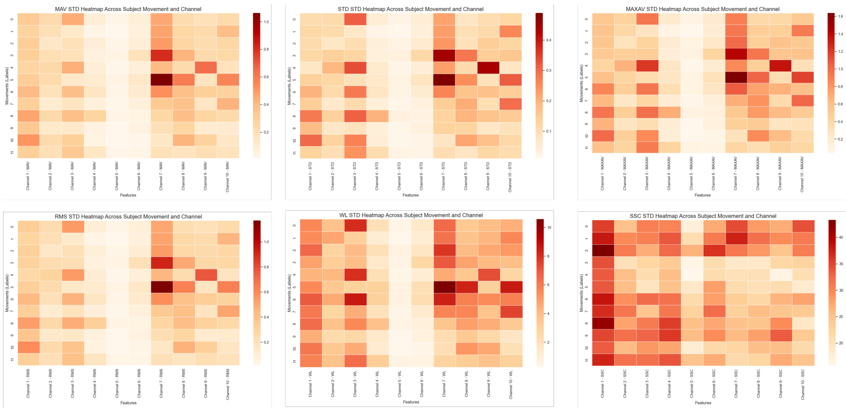


Figure 2: STD Heat Maps

Question 3

The model performs significantly better when trained and tested on the same subject (accuracy: 0.97) compared to being trained on 26 other subjects and tested on a new one (accuracy: 0.39). This indicates that the model struggles to generalize across subjects, likely due to subject-specific data characteristics dominating the feature space.

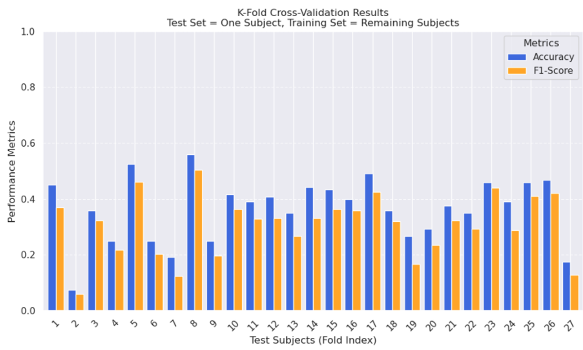
Table 1: Training comparison on subject 18

	On the same subject	**On the 26 others**
k best features	36	47
Accuracy	0.97	0.39

Question 4

The cross-validation results (Figure 3) indicate poor generalization of the model across subjects. The mean accuracy is **0.36 ± 0.11**, and the mean F1-score is **0.31 ± 0.11**. Performance varies significantly, with a minimum accuracy of **0.075** and F1-score of **0.060**, and maximum accuracy of **0.558** and F1-score of **0.503**.

These variations suggest that the model struggles to capture subject-invariant patterns, leading to inconsistent performance when applied to unseen subjects. This highlights the need for improved strategies to handle inter-subject variability and enhance model robustness.



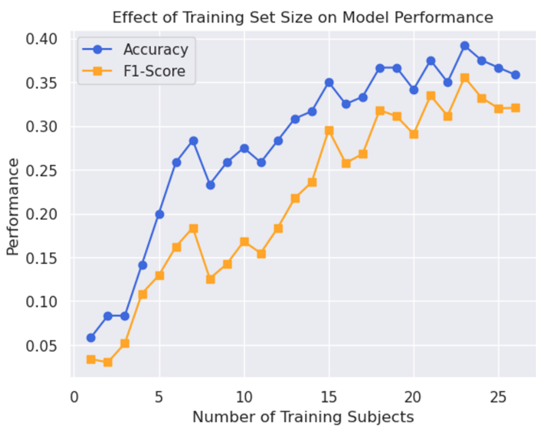
	Mean ± Std	**Min**	**Max**
Accuracy	0.36 ± 0.11	0.075	0.558
F1-Score	0.31 ± 0.11	0.060	0.503

Figure 3: Cross-Validation Performance Across Subjects

Question 5

The results (Figure 4) for Subject 18 demonstrate that increasing the size of the training set improves classification performance. Accuracy rises from a minimum of 0.058 (with 2 training subjects) to a maximum of 0.392 (with 23 training subjects). However, despite this improvement, the overall performance remains relatively low, with a mean accuracy of $\mathbf{0.28 \pm 0.10}$ and a mean F1-score of $\mathbf{0.22 \pm 0.10}$.

These findings suggest that while adding more training subjects slightly enhances generalization, the gains eventually plateau. This trend highlights the model's limitations in effectively learning subject-invariant features and achieving robust performance across a diverse range of subjects.



	Accuracy	Size of training set
Min	0.058333	1
Max	0.391667	23

Figure 4: Impact of Training Set Size on Performance (Subject 18)

Part 3 : Regression for joint angles

In this section, we implemented a regression algorithm to predict joint angles using EMG data from a single subject. The objective was to improve robotic hand control, enabling precise movements such as achieving half-closed positions.
The dataset used for Part 3 comes from the NINAPRO database. This dataset includes EMG signals and glove signals (for joint angles) on which our analyses will focus. See the Appendix for further information.

Question 1

We began by extracting data by selecting the variables of interest, i.e. EMG and glove, and the joint angles of interest, i.e. joint angles number 3,6,8,11,14 .

Moreover, we decided to use the first dataset of the first subject for the training, the second one for the validation, and the third one for the test of the model. This choice simplifies the splitting step because it is not possible to split the dataset randomly, as we are working with time series data and need to preserve the temporal structure to ensure realistic and meaningful evaluation of the model's performance.

The raw dataset already includes some preprocessing steps (e.g., upsampling, synchronization, and calibration), but further preprocessing, such as noise filtering to remove in particular powerline interference, or normalization, might still be required for the regression model.

Question 2

A sliding windows was used to analyze time-series data by breaking it into smaller, overlapping segments. This technique helps capture temporal dynamics while reducing the complexity of processing large datasets at once. In order to capture sufficiently detailed information on finger movements while avoiding redundancy, we opted for a window size of 128 ms. The sliding step chosen corresponds to 60 % of the window size, i.e. 50 ms, thus maintaining a balance between accuracy and computational load. Too large a window might mix different phases of motion, while too large a step might miss important transitions. These parameters enable significant variations in motion to be captured while optimizing computational resources.

Question 3

A feature extraction of the EMG data was performed to reduce dimensionality without losing essential information. The features that we extracted are statistical measures such as mean, standard deviation, and maximum signal amplitude in each time window. The mean provides an overall indication of muscle activity, the standard deviation captures the variability of muscle activity, and the maximum amplitude represents the peak of muscle engagement.

Additionally, the mean of the labels is computed to represent the overall muscle activity during the window, aiding in the prediction of continuous variables in regression tasks.

The correlation matrix reveals that same-channel features are highly correlated, indicating that they capture similar signal characteristics. In contrast, features from different channels show weak or negative correlations, suggesting that they represent distinct aspects of EMG signals. Overall, this correlation matrix highlights the potential for optimization of the feature set to improve model performance.

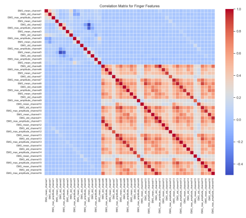


Figure 5: Correlation matrix

Question 4

To predict joint angles, we use a regression support vector machine (SVR) algorithm with a multi-output regression method to simultaneously predict multiple joint angles for each time window of EMG data.

The result of our regression (Figure 6) shows that the regressor accurately captures general movement trends. For most joints, the predicted joint angles (shown in orange) closely follow the actual joint angles (shown in blue), demonstrating the effectiveness of the model in predicting finger movements.

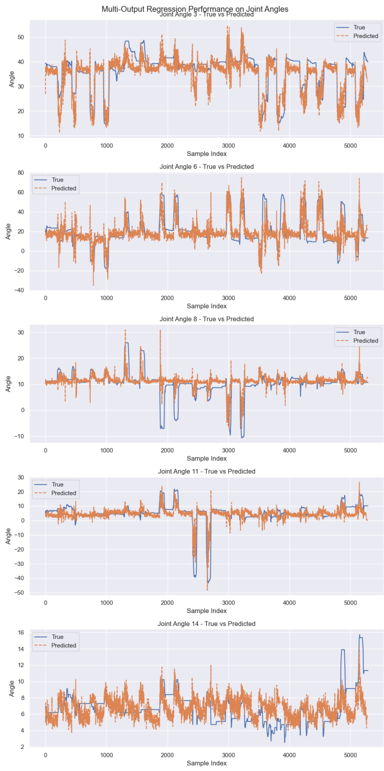


Figure 6: Regressor performance

Question 5

To assess the model's performance, we used the root mean square error (RMSE) and the mean absolute error (MAE) as metrics. The RMSE is sensitive to large errors and particularly useful when large deviations are more costly. The MAE, on the other hand, allows clear interpretation of the mean prediction error without amplifying larger deviations as the RMSE does. The two metrics thus provide complementary views of model accuracy.

As we can see in Figure 7, the performance is satisfactory for most joint angles (e.g., joints 3 and 6), where RMSE and MAE values are low, indicating good prediction accuracy. However, joint angle 14, which concerns the little finger, has a slightly higher RMSE and MAE, suggesting that predictions for this joint angle are not entirely satisfactory and need to be improved, particularly for robotic hand control which requires high accuracy.

Question 6

Regression performance is not stable for all finger angles. Some joint angles show consistent and accurate predictions, while others, in particular joint 14, show larger errors, especially at larger angles, indicating instability of performance. The reason for these differences can come from several sources such as data quality, features relevance or sensor sensitivity.

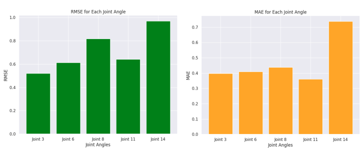


Figure 7: Evaluation of the performance using MSE and MAE

Appendix

References

Manfredo Atzori, Arjan Gijsberts, Ilja Kuzborskij, Simone Heynen, Anne-Gabrielle Mittaz Hager, Olivier Deriaz, Claudio Castellini, Henning Müller, and Barbara Caputo. Characterization of a benchmark database for myoelectric movement classification. *Transactions on Neural Systems and Rehabilitation Engineering*, 23(1): 73–83, January 2015. ISSN 1534-4320. doi: 10.1109/TNSRE.2014.2328495. URL `http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6825822`. (In press).

Advanced dataset 1 description

These signals were recorded while each subject performed 12 simple hand movements (stimulus), collected using 10 electrodes (channels) positioned uniformly around the subject's forearm and on key flexor and extensor muscles.Each movement was performed 10 times, with each repetition lasting 5 seconds, followed by a 3-second rest period. The resulting signals were sampled at a frequency of 100 Hz and preprocessed both directly by the electrodes and after acquisition. This preprocessing included amplification, band-pass filtering (to retain EMG-relevant frequencies between 15 and 500 Hz), and Root Mean Square (RMS) rectification of the raw signal, ensuring that only the components relevant for EMG analysis were retained. Additionally, the signals were synchronized with the observed muscle responses and the corresponding movement labels.

The data for each subject is organized so that each row represents a specific time point, with one column corresponding to the timestamp and 10 columns representing the signals recorded by each electrode. Additionally, the movement repetition number and the movement or rest label, synchronized with the signals, are available for each time point.

Advanced dataset 8 description

The dataset used to performed joint angles predictions is the NinaPro Dataset 8 from subject one. This dataset is focused on the estimation and reconstruction of finger movements using regression algorithms based on EMG signals. The data collection involves two types of sensor, 16 active wireless sensors were used to record sEMG signals and 9-axis inertial measurement units (IMUs) and 18-degree-of-freedom (DOF) Cyberglove that was used to record hand kinematics. The datasets contains 11 variables (subject, exercise, EMG, ACC, Gyro, Mag, Glove, Stimulus, Restimulus, Repetition, Rerepetition) but only EMG and Glove were used to perform the predictions. The sEMG signals and IMU data (accelerometer, gyroscope, magnetometer) were upsampled to 2 kHz and synchronized to ensure consistency across all modalities. Additionally, the Cyberglove data were calibrated for each participant before each session using a manufacturer-provided quick calibration procedure, aligning the glove signals with the subject's specific movements. These steps enhance the accuracy and reliability of the recorded data. The data acquisition protocol involved subjects performing a sequence of movements (e.g., thumb flexion, index finger flexion, cylindrical grip) with 5-second repetitions followed by 3-second rest periods. Each participant's data is provided in three acquisitions as MATLAB .mat files. Acquisitions 1 and 2 contain 10 repetitions of each movement for training and validation of the regresssion model, while Acquisition 3 includes only 2 repetitions per movement and is reserved for testing and performance evaluation.

EMG Extracted Features

Mean Absolute Value (MAV): $\text{MAV} = \frac{1}{N}\sum_{i=1}^{N}|x_i|$. Measures the overall signal amplitude.

Standard Deviation (STD): $\text{STD} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}$. Reflects the variability or spread of the signal around its mean.

Maximum Absolute Value (MaxAV): $\text{MaxAV} = \max(|x_i|)$. Captures the peak amplitude.

Root Mean Square (RMS): $\text{RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}x_i^2}$. Measures the energy content of the signal.

Waveform Length (WL): $\text{WL} = \sum_{i=1}^{N-1}|x_{i+1}-x_i|$. Represents the cumulative length of the signal.

Slope Sign Changes (SSC): $\text{SSC} = \sum_{i=2}^{N-1}\text{sign}((x_i-x_{i-1})(x_i-x_{i+1})) < 0$. Counts the number of slope changes in the signal.