UNIVERSITAT DE BARCELONA

FUNDAMENTAL PRINCIPLES OF DATA SCIENCE MASTER'S THESIS

Predicting human perceptual decisions from neural activity

Author:
Junjie JI

Supervisors:
Dr. Alexis PEREZ
Dr. Ignasi COS

A thesis submitted in partial fulfillment of the requirements for the degree of MSc in Fundamental Principles of Data Science

in the

Facultat de Matemàtiques i Informàtica

September 3, 2023

UNIVERSITAT DE BARCELONA

Abstract

Facultat de Matemàtiques i Informàtica

MSc

Predicting human perceptual decisions from neural activity

by Junjie JI

The human brain, with its intricate network of neurons and complex dynamics, presents a fascinating subject of study. In the present work we aim to investigate how humans integrate perceptual evidence in time to make decisions. To accomplish this, we conducted electroencephalogram procedures on 26 participants at the Mundet Campus of the Universitat de Barcelona in 2022. The data collection process included rigorous preprocessing steps to ensure data quality, followed by data segmentation and dataset creation for subsequent analyses. From this data, we will employ time series analyses and machine learning models to gain deeper insights into the neural processes associated with the integration of perceptual evidence and decision-making.

Acknowledgements

The former thesis portraits the entire year of knowledge acquisition and months of dedicated work throughout the master's degree.

I would like to thank both my advisors, specially Alexis Pérez-Bellido, for the continuous support since the beginning until the very end, for the tedious process of gathering the data and for the clear explanations of the neuroscience field where I knew nothing about in the very beginning.

To the master's coordinator, Jordi Vitrià, for the best academical year I've had so far.

To my family and all the friends I made along the way, who taught me, supported me and made me laugh so much.

Thank you very much.

Contents

Chapter 1

Introduction

Cognitive neuroscience is a field of study that gained a lot of attention in recent years with the application of machine learning (ML) and artificial intelligence (AI). Brain structure and neural activity is informative about multiple cognitive phenomena such as planned action movements, decision-making or even clinical diagnosis. This progress in our understanding of brain functioning is largely supported by the increasing development of new and more sophistical neuroimaging techniques like functional magnetic resonance (fMRI) which uses a magnetic field and computer-generated radio waves to create detailed images of the organs and tissues in the body with a fairly precise spatial resolution, and electroencephalography (EEG), magnetoencephalography (MEG) that registers scalp level changes in voltage induced by neural processing, and allow researchers to monitor and visualize brain activity with exceptional temporal resolution.

The continuous improvement of these technologies has allowed us to generate highly complex and multivariate datasets and using and developing new machine learning algorithms is becoming pressing in order to use all that information to increase our knowledge about brain functioning but also to develop new methods that can facilitate the life of many.

There are multiple examples of the practical application of machine learning algorithms in clinical research. For instance, it has been demonstrated that it is possible to use EEG activity in combination with machine learning techniques to diagnose depression and to identify mental states retrieved from EEG signal (Ksibi et al., 2023). Another similar example is the usage of EEG data for the purpose of classifying and gaining insights into Attention-Deficit/Hyperactivity Disorder (ADHD) (Martínez González et al., 2022).

The near future is full of machine learning and neuroimaging integration that aims to change our way to interact with the world. For example, Neuralink is a visionary venture that aims to seamlessly merge the power of artificial intelligence with our biological capabilities, opening up a world of unprecedented possibilities (Musk, 2019). With Neuralink, individuals could overcome physical limitations and regain lost sensory functions through advanced neuroprosthetics. Furthermore, this technology could pave the way for enhanced cognitive abilities, restoring perceptual losses, allowing for faster learning, improved memory retention, and even heightened creativity.

In this project we sought to extract information of healthy human observers in a perceptual decision making task. Perceptual decision making is the process by which sensory information is used to guide behavior toward the external world. One of the humans remarkable cognitive skills is the ability to make complex and informed decisions. Nevertheless, the complexity of this cognitive process is still a subject of research (Gazzaniga, 2022), as it depends on multiple factors, for instance, we need to consider both the information we need to assess and the prior knowledge we possess related to that information. We aimed to explore the rapid neural dynamics involved in perceptual decision-making and how they are influenced by prior perceptual experiences

To investigate these aspects, we decided to use electroencephalography, a technique that allows us to sample changes in activity voltage near the brain at high temporal resolutions (exceeding 500Hz). As previously mentioned, electroencephalography refers to the measurements of electrical activity recorded from the scalp using a set of electrodes strategically placed at specific locations. This non-invasive technique allows researchers to capture and analyze the electrical signals generated by the collective activity of billions of synchronized neurons in the brain. Although EEG has very limited spatial resolution, it allows to measure neural activity from cortical regions with high temporal resolution.

Recall that all the coding that has been done throughout the thesis is written in Python and available on Github.

Chapter 2

Experiment

2.1 Objectives

The main goal for this thesis is to understand how humans integrate perceptual evidence in time to make decisions. Specifically, we want to investigate the neural basis of the perceptual decision making process by applying different multivariate approaches on multielectrode in order to extract information. Human decision-making is influenced by a series of past experiences, choices, and their associated outcomes. This cognitive process involves integrating information from previous encounters with similar situations to inform and guide future decision-making. (Salinas and Izquierdo, 2019, St. John-Saaltink et al., 2015, Talluri, Braun, and Donner, 2021 and Urai et al., 2019)

But most of the research in this respect has been generated behaviorally (St. John-Saaltink et al., 2015, Talluri, Braun, and Donner, 2021 and Urai et al., 2019). So in this study we sought to characterize the neural changes from a neuroimaging approach. To do that, we will take advantage of different machine learning algorithms that will allow us to extract information from the stimuli encoded in neural activity patterns, the internal decision of the observer and their manual response.

In addition, we will explore how neural representations of these different processes change as a function of participants' previous choices.

2.2 Participants

26 healthy humans participated in the experiment. Participants were informed about the possible risks and inconveniences associated with EEG prior to the beginning of the experiment, and received 15 euros after the experiment completion. The experiment was approved by the bio-ethical committee of the University of Barcelona.

2.3 Experimental Design

In this experiment participants were exposed to a fast sequence of 6 oriented gratings (Figure 2.1 A). Each grating was presented during 250ms and the orientation sequences were sampled from a lookup table in which sequences were sorted by their average mean. Participants were instructed to estimate the mean orientation of the sequence and categorize it as closer to the diagonal or the cardinal axis using the mouse.

The same sequence was presented three times in each trial and participants had to respond after each sequence presentation (Figure 2.1 C). The total duration of the experiment was around 1 hour and 30 minutes (including the time dedicated to the electrodes placement, around 30 minutes)

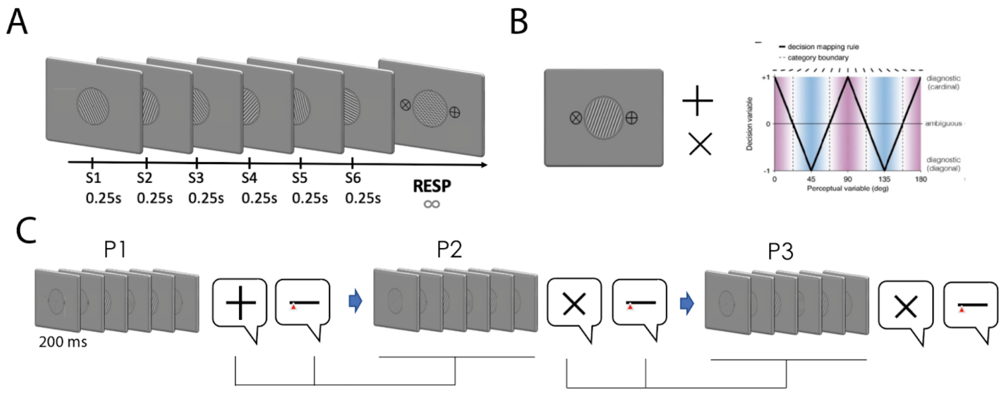


FIGURE 2.1: A) In each sequence presentation, 6 gratings were presented in rapid succession. B) Participants had to categorize the mean of the orientation sequence as closer to the cardinal or diagonal axis. This manipulation orthogonalizes the stimulus orientation and the decision variable (i.e. perfectly opposed orientations corresponded to the same decision. For instance 0 and 90 degrees are physically different but both are cardinal) See orientation to decision variable mapping at right inset (extracted from Wyart, Nobre, and Summerfield, 2012). C) In each trial, participants were presented with 3 repetitions of the same stimuli sequence.

Following each sequence, participants were required to move the mouse to indicate their response regarding the sequence's cardinality or diagonality. This answer was constrained by the fact that at every trial the diagonal and cardinal buttons were randomly positioned introducing a variable element that affected the decision making process, in order to recude the impact of motor responses bias (e. g. a tendency to respond left) in the final performance.

After their choice, they had to rate their confidence in their decision using a sliding triangle over a response bar. Only at the end of the three presentations participants received feedback about their performance.

Chapter 3

The Data

Our data consists of the recordings in EEG experiments retrieved from the 26 participants. The experiments were conducted at the Mundet Campus from the University of Barcelona in the summer of 2022.

During the experiment, the participants wore a 32 active electrodes actiCAP from BrainVision (BrainVision, 2022). The electrodes were placed according to the 10-20 system (Oostenveld and Praamstra, 2001) as shown in Figure 3.1A. More information regarding these electrodes' names and their respective code labels can be found at A. The scalp's electric activity is amplified using an amplifier also from BrainVision and recorded at 500hz on an additional supporting computer (Figure 3.1B). An example of these EEG recordings can be found in Figure 3.1C, which shows raw data before the preprocessing.

3.1 Preprocessing

The electrodes registered subtle voltage fluctuations at scalp-level resulting from synchronized postsynaptic spiking neural activity. The EEG recorded signal is corrupted by external (not-neural) electrical activity perturbations like 50 hz line noise, electrical muscle activity and eye-blinks. The data set was preprocessed applying a broadband frequency filter with a range from 0.2 Hz to 30 Hz. This temporal filter removed slow wave components associated with signal-drift (e.g. sweating) and high frequency modulations related to muscle activity (e.g. neck and jaw tension or head movements). A 50Hz notch-filter was applied to remove line-noise activity. Moreover, we used an Independent Component Analysis (ICA) on the whole data set, proven effective in the literature (Bhimraj and Haddad, 2017, Sun, Liu, and Beadle, 2005). By visually inspecting the topological and temporal distribution of the different ICA components, we detected those associated with eye-blinks and rhythmic heartbeat, and we regressed them out from the data set.

Subsequent experiments revealed notable discrepancies in the neural activity of the eye channel, specifically the 25th electrode, when compared to the other electrodes. The eye channel exhibited substantially higher signal output. Hence we considered it an outlier and proceeded with the removal. Finally, we used the triggers sent from the experimental computer in order to identify the onset of those temporal data segments that were associated with the processing of the stimuli.

Finally, throughout the entire work, a frequency re-sampling was done to all the data, from initial 500Hz to 50Hz in order to reduce the duration of the experiments.

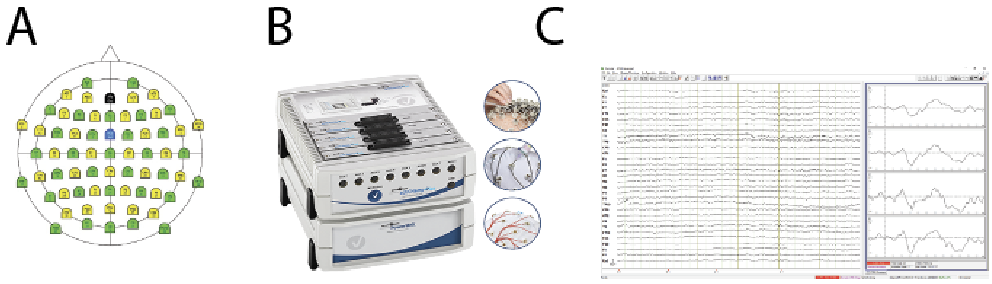


FIGURE 3.1: A) Electrodes topographic distribution. We used a 32 electrodes setup (only the electrodes labeled in green in the example layout). B) Scalp electric activity is amplified using a Brainvision amplifier and recorded at 500hz on an additional supporting computer. C) Example of the EEG recordings visualization during the testing session.

3.2 Data Segmentation and Data Set Creation

We created a data set in which EEG activity was epoch-ed [1] taking the onset of each sequence presentation (-0.5 to 5 s relative to the onset of the sequence) in order to understand how Event-Related Potential (ERP)[2] activity unfolded during the whole trial (Figure 3.2A and Figure 3.2B). The data from our EEG recordings was saved in FieldTrip FIF formats for its ease of use and compatibility with the MNE library (Oostenveld et al., 2011, Gramfort et al., 2013).

For study-specific purposes, a smaller data set was created by separating the 6 oriented gratings of each sequence. We will refer to this smaller data set as *Stim data set* and the original data set as *Main data set*.

Moreover, for both data sets, we also have their corresponding *Metadata data set*. This data set consists of the variables used to conduct each trial, such as, participant ID (known as *subj*), degree of the orientations given to the participant, decision of the previous trial, confidence level of their decision, etc. More information about the Metadata data set can be found at B.

The Main data set results in a 4-dimensional array, composed by 26 participants, 31 electrode channels, around 250 trials and 2875 time points (288 time points when down-sampled).

[1]EEG epoch data refers to a segmented or divided portion of continuous EEG recordings, which are used in the analysis of brain activity.

[2]Event-Related Potentials are electrical patterns or brain responses that are measured from the scalp of a person in response to a specific sensory, cognitive, or motor event or stimulus

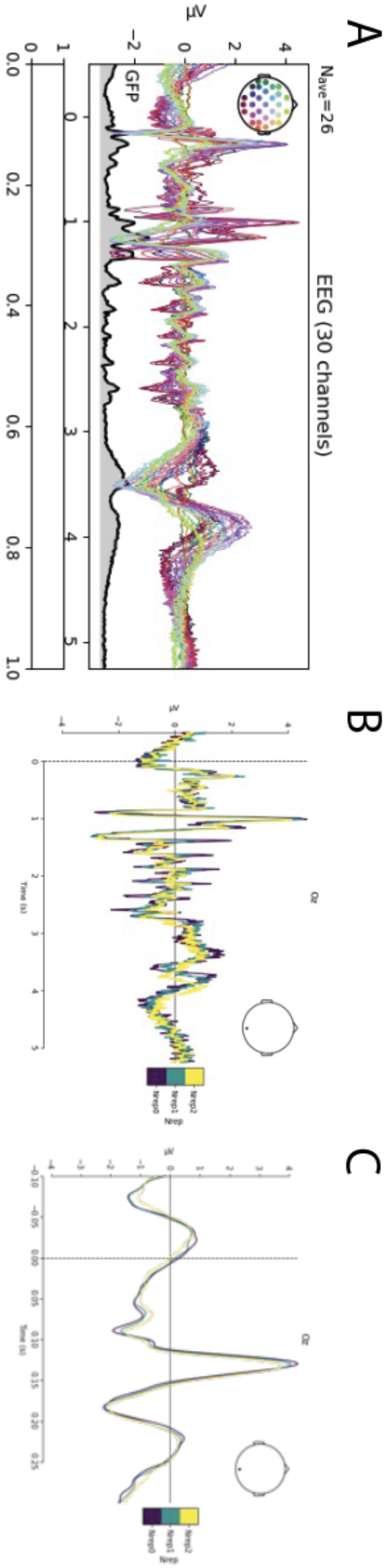


FIGURE 3.2: A) Trial epochs: Mean ERP activity over time and across trials for each electrode and its topographic distribution (different color for each electrode). B) Mean ERP activity at the Oz (occipito-central) electrode for each presentation. C) Stimuli epochs: Same as in B but ERP activity evoked by each individual stimulus in the sequence.

Chapter 4

Data Analysis and Methods

In perceptual decision making we can differentiate three different processing stages. The first one is the transduction and encoding of the physical signal into neural activity in early sensory areas (e.g. visual cortex). A second phase in which the sensory encoded information is mapped and integrated into a latent decision variable that will be finally transformed in a motor response. In this experiment participants should encode different orientations (0 to 180o) and map them onto two categories (cardinal or diagonal) and finally select a response option (left or right).

Using different machine learning algorithms we evaluated whether we can predict information in each one of the phases based on neural activity patterns and whether the decoded information is affected by the previous choices of the participants.

4.1 Orientation Decoding

Recent studies in humans have been able to measure orientation selectivity in the primary visual cortex using non-invasive neuroimaging methods like fMRI (Kamitani and Tong, 2005) and EEG (Garcia, Srinivasan, and Serences, 2013).

In our experiment, participants were exposed to a fast sequence of 6 oriented gratings. In our first analysis we will decode the orientation of each of the sequence samples using a forward encoding model.

A forward encoding model is a type of computational model that breaks down complex stimuli into distinct sensory features. In our case, we capitalize on the orientation tuning of visual cortex neurons in order to filter each stimulus through different orientation channels. An orientation channel is a mathematical function that transforms each stimulus orientation function in the hypothetical response profile of the orientation tuned neurons in the visual cortex.

Let X be a vector of size N, where N is the number of trials that contains the orientations data, let Y be a matrix of size $F \times N$, where F is the number of features, that contains the EEG data, let b be the vector of coefficients and ϵ the noise variable. (Essentially X is given by the orientation column in the Metadata data set and Y is given by the Stim or Main data set)

Mathematically, the linear equation is

$$Y = X^T \beta + \epsilon$$

and its known that the optimal β coefficients for the least squares approach is,

$$\beta = (X^T X)^{-1} X^T Y$$

estimating the noise covariance letting the columns of ϵ be the variables,

$$S = \text{Cov}[\epsilon_{ij}]$$

finally our weights are given by,

$$w = \beta S$$

In practice we add a regularization parameter to the noise covariance and also a demean parameter when needed. Also by inverting the weights we can recover the orientations from predicted activity patterns.

These models are extensively utilized in neuroscience (Garcia, Srinivasan, and Serences, 2013, Kay et al., 2008) to gain insights into how the brain processes sensory information. Furthermore, such models can help identify relevant features and channels that play key roles in shaping neural responses, providing valuable insights into the complex interplay between sensory stimuli and neural activity.

Regarding the analyses, we initiated the process by discretizing the orientation values into distinct bins, effectively creating 8 channels. This categorization mapped values in the range of 0 to 180 degrees into their respective channel representation, this is shown at Figure 4.1A. Subsequently, we extracted the stimulus features from the epoch data, enabling us to obtain theoretical channel responses based on the orientation presented. In the following step, for each time point of the data, we executed a cross-validation procedure, where each fold of the data set served as a training set for a forward encoding model. We store the weights of the decoded information as the values of the orientation channels, showcased at Figure 4.1B. Finally, the study computed the mean values of the orientation channels and corresponding channel responses across all participants. These mean values were then utilized to create visual representations alongside their respective confidence intervals, aiding in the interpretation and analysis of the results.

Additionally, in order to assess how current decisions change as a function of previous decisions, the exact same procedure was taken but grouping trials by their corresponding number of repetition, resulting in three data sets based on the first, second and third presentations (P1, P2 and P3 respectively) experiment group that the participants took. More insights regarding the results and following procedures will be explained later in the results section for this particular study.

Lastly, to see if the participants were biased to the decisions that they took on the previous trial, we also conducted the study by splitting the data set based on whether the decisions that the participants took on their previous trial was cardinal or orthogonal.

Essentially, we performed three studies. The first one regarding orientation decoding, the second one regarding the repetition within the trial and the final one regarding the decision taken on the previous trial.

Our approach to this model uses the Matlab toolbox by Pim Mostert, found here Dr. Pérez-Bellido did a Python adaptation of this code.

More insights regarding the results and following procedures will be explained later in Chapter 5 for these studies.

4.2 Decision and Response Decoding

In this section we will explore different techniques in order to decode the decision and response variables from our multivariate EEG data. The decision variable corresponds to a binary variable in which participants gradually integrate information to

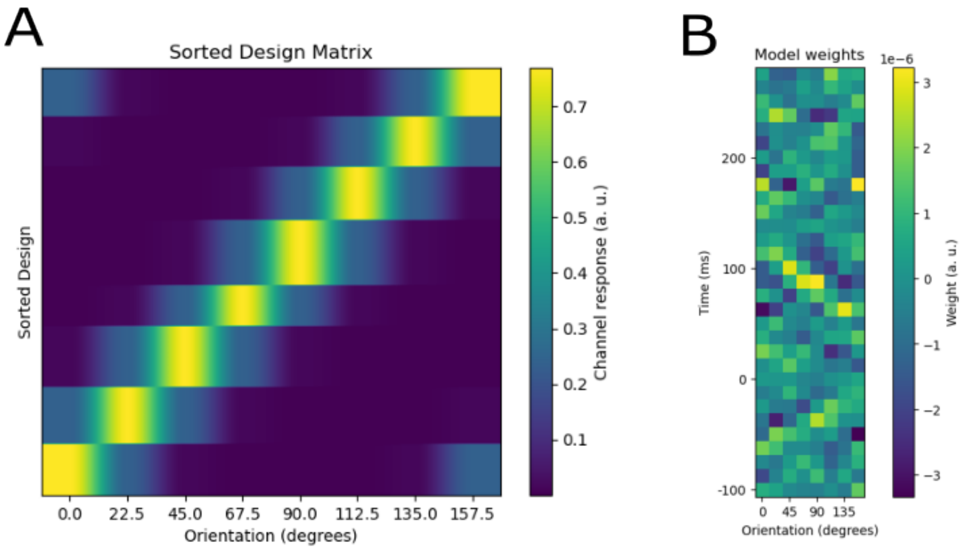


FIGURE 4.1: A) Sorted values of the design matrix that contains the hypothetical channel responses given a presented orientation. Stimuli orientations for each trial have been discretized in 8 bins along the y axis. Orientation channels responses for each stimuli are represented along the y axis. The response strength is depicted by the color scale. B) The weights of the Forward Encoding Model at every time point and orientation for one subject.

decide whether they will respond cardinal or diagonal. The response value instead, indicates whether participants will move the mouse to the left or to the right of the monitor to make the final choice.

4.2.1 Vector Autoregression

The Vector Autoregression (VAR) model is an extension of the univariate autoregressive model that allows for the simultaneous analysis of multiple time series variables. By incorporating the inter-dependencies among the variables, the VAR model captures the dynamic relationships and feedback mechanisms between them.

The general formula for the VAR of order p is

$$y_t = c + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + e_t$$

where y_{t-i} indicates that variable's value i time periods earlier and are called the "i$^{\text{th}}$ lag" of y_t, c is a vector of constants serving as the intercept, A_i is a time-invariant and e_t is a vector of error terms. These models have been use on the literature to handle EEG data (Herrera et al., 1997).

For our model, we tried different values for p, but the best performing value was p = 2. This selection was taken based on the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), both parameters were found the lowest when using p = 2. After fitting the model, we extracted the parameters and trained multiple machine learning classifiers on these parameters and the target decision and response values. However the results were essentially a random choice for all the classifiers. To address these limitations, consideration should be given to more other methods, such as machine learning or deep learning techniques.

4.2.2 Machine Learning Approaches

Given that the problem is a classification task, we seek to build a machine learning model with the EEG data to handle the binary values of the decision and response variables. These techniques are increasingly being applied to EEG data for pattern analysis, group membership classification, and brain-computer interface purposes (Saeidi et al., 2021).

Support Vector Machines

Support Vector Machine (SVM) is a machine learning algorithm used for classification and regression tasks. SVM works by finding a hyperplane that best separates different classes in the feature space while maximizing the margin (distance) between the nearest data points of each class, called support vectors. This hyperplane aims to achieve the highest generalization by effectively classifying new, unseen data. SVM can handle linear and non-linear separation by using kernel functions to transform data into a higher-dimensional space where linear separation is possible.

Random Forest

Random Forest is an ensemble learning algorithm that is often used classification tasks. It works by constructing multiple decision trees at training time and combining their outputs to make a final prediction. Each decision tree is trained on a different subset of the training data and uses a random subset of the features to split the data at each node. The final prediction is made by aggregating the predictions of all the decision trees by taking the majority vote.

4.2.3 Time Point Classification

The last experiment involved training Random Forest and SVM classifiers separately at each time point using the data channels as features instead. The objective was to target the decision and response values, and the same approach was applied to both variables. For both classifiers, standard hyper-parameters were used due to time constraints, and the fact that we used the same classifier several times. For each time point, an 80-20 split was used for training and testing, with the accuracy being recorded. Subsequently, the mean accuracy for all participants was calculated along with their respective confidence intervals, providing a measure of the model's overall performance.

To gain insights into the classifier's performance under random conditions, we conducted an additional set of experiments. In these randomized trials, we performed the same procedure of training and testing, but with the labels shuffled 10 times. This allowed us to establish a baseline to compare the real model performance against random chance. By comparing the classifier's accuracy to this shuffled label baseline, we might be able to determine whether the model's performance was statistically significant and not simply a result of random guessing.

4.2.4 Long Short-Term Memory Networks

Long Short-Term Memory (LSTM) networks have gained significant attention and demonstrated remarkable efficacy in analyzing and predicting time series data. Their ability to capture long-term dependencies, effectively model complex temporal patterns, and account for variable time lags make them a powerful tool in time series analysis. For instance, there is an implementation of an LSTM network for time series forecasting in the work by (Elsworth and Güttel, 2020) provide an implementation of an LSTM network for time series forecasting. In this work, they explicitly highlight the challenge that 'Machine learning methods trained on raw numerical time series data exhibit fundamental limitations such as a high sensitivity to hyper-parameters and even to the initialization of random weights.' This underscores the importance of carefully considering the model's architecture and parameters when working with time series data.

We implemented a simple LSTM network that on the first layer had the input shape matching the dimensions of the data set. Added a regularization layer and another LSTM layer and finally sigmoid activation function in order to handle the classification task. A schema of the network is shown in C. Notice that the conducted experiments encompassed two distinct data preprocessing approaches: one incorporating Principal Component Analysis (PCA) and the other omitting PCA prior to fitting the LSTM network.

The presented network, while powerful, lacks explainability due to its deep learning nature. However, the primary objective of this study was to assess the performance and capabilities of a more sophisticated deep learning approach in modeling the EEG data and decoding the human decision or response values that we failed to decode previously. By employing this advanced modeling technique, the research aimed to ascertain whether the complex neural patterns captured by the deep learning model could effectively predict and elucidate human decision-making or response processes based on EEG data.

4.3 Implementation

Details of most of the methods described in this section can be found in GitHub on the *methods.py* file.[1] The file contains functions for reading EEG signals that were previously saved in FIF formats. Distinct reading functions are used in this context since the two data sets are slightly different. These divergences significantly contribute to the ease and effectiveness of data categorization for both number of repetition or decision taken on previous trial. The file also contains functions for the entire pipeline of the methods described in this Chapter, that varied depending on the experiment and multiple other variables such as frequency down sampling (for time efficiency), number of participants, class balance on classifiers, using either the Main data set or Stim data set, etcetera. However, the methods that are not found in this file are done within their corresponding Python Notebook files alongside their analyses and results.

[1]The code repository is accessible at: `https://github.com/unthrived/master-thesis/blob/main/toolbox/methods.py`.

Chapter 5

Results

To enhance clarity and facilitate a better understanding of the data patterns, a smoothing or filtering procedure was applied to the visualizations when required, allowing for more coherent and interpretable representations. This smoothing or filtering step was solely utilized for visualization purposes and did not impact the underlying statistical analyses or interpretations of the findings. Additionally, it is important to recall that a frequency down-sampling was conducted in order to reduce the duration of the experiments. Consequently, the results may not be as qualitatively robust as they could be.

5.1 Behavioral Results

We calculated the proportion of correct decisions in each presentation. As expected, participants' performance improved with more repetitions of the sequence (Figure 5.1A). To describe how participants' responses changed as a function of the information in each trial, we plotted the proportion of diagonal responses against the mean decision variable information in each trial (negative values represent cardinal trials, and positive values represent diagonal trials). The participants' performance is well described by a logistic regression curve (Figure 5.1B). Consistent with the results in Figure 5.1A, the slope of the logistic regression is steeper with more repetitions, demonstrating that participants weigh the stimuli DV information more in the P3 compared to the P1 sequence. Additionally, we observed that participants showed a bias to report diagonal orientations more often than cardinal orientations (logistic curves are displaced slightly to the left from the middle point), despite the probability of both categories being the same in the experiment. This might be explained by an implicit bias to judge every stimulus that is not perfectly vertical or horizontal as diagonal.

Finally, in Figure 5.1C, we explored whether participants' responses were biased towards or away from previous choices. To test this possibility, we divided the trials into two distributions: one with trials in which participants had responded cardinal in the previous presentation (in green) and one including trials in which participants responded diagonally in the previous presentation (in purple). Our results revealed that when participants had responded cardinal in the previous trial, their responses in the current trial were biased towards the cardinal choice, and when participants had responded diagonally, their responses were biased towards the diagonal. This pattern appeared only after the first presentation and is reflected by a shift in the logistic regression curves.

In summary, our behavioral results demonstrate that participants take advantage of the repeated information to improve their performance in a perceptual decision-making task, but this improvement might be limited by an implicit bias to repeat previous choices.

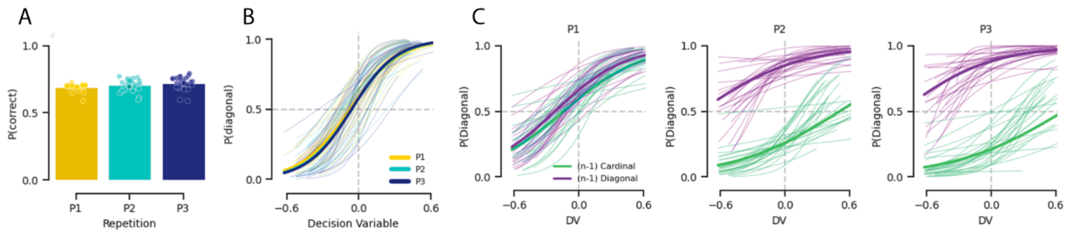


FIGURE 5.1: A) Group average proportion of correct responses as a function of the repetition number. Points represent individual scores. B) Proportion of diagonal responses as a function of what is the mean DV in each trial for each repetition. Thick and thin lines represent group average and individual logistic regression fits to the empirical data respectively. C) Proportion of diagonal responses as a function of the mean DV for trials in which participants responded cardinal or diagonal in the immediate previous trial. Same labeling convention as in B.

5.2 ERP Activity

Recall from Figure 3.2C and Figure 3.2C the ERP activity diminishes after each presented orientation. This complements the behavioural results discussed earlier.

In the next analyses, we will apply different machine learning approaches to the neural data collected during the task to understand how stimuli information encoded in neural patterns changes at different processing stages in each trial.

5.3 Orientation Decoding Results

As stated on Chapter 3, aside from the Main data set processed from the EEG signals, we also have the Stim data set, to focus on individual orientations. Here we will work mainly on this Stim data set since we are targeting sensory level decoding.

5.3.1 Stimuli Data Set

Firstly, we applied the forward encoding model to the Stim dataset, and studied the decoding results from the orientation stimuli. In this study, we present the model's outcomes when applied to the Stim dataset around the time range of 100 ms. This is shown in (Figure 5.2A) were we see the correlation between the orientation channels and the decoded orientations. Essentially the peaks of every orientation channel match accordingly to their decoded orientation. This is to assess that the forward encoding model is working correctly and for later studies, centering of these curves and adding the confidence intervals will be done in order to provide better insights of the results.

As mentioned before and on section 4.1 Figure 5.2B showcases the forward encoding model applied at every point of the data set using a cross validation procedure for all EEG channels and all participants. As seen in the figure, stimuli on this data set start appearing from time 100ms on wards and we can observe that the stimuli have a higher intensity at the beginning and at the very end of the trial. The former high intensity might indicate that the stimuli had more impact on the participants due to previous trials and the later high intensity might indicate the stimuli is

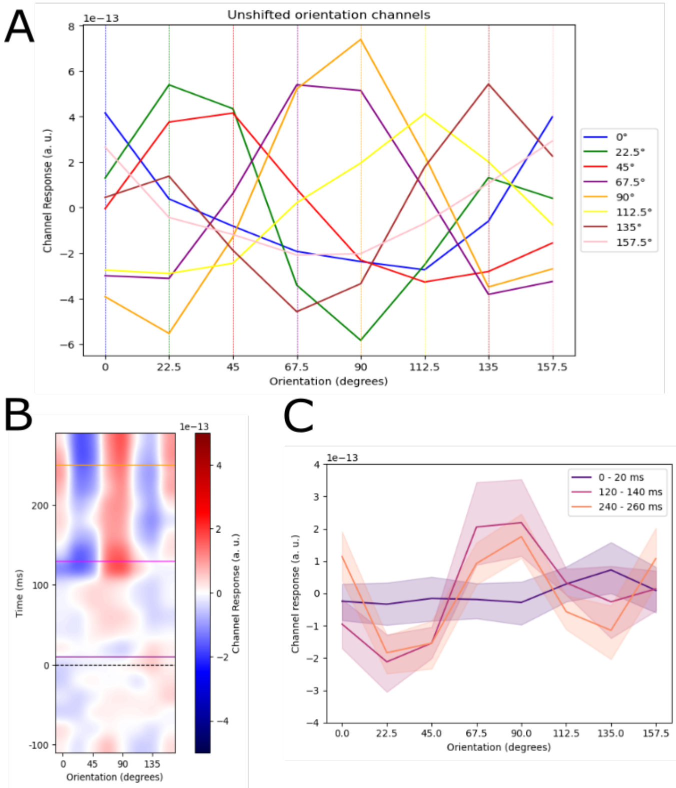


FIGURE 5.2:
A) Sample result of the forward encoding model applied for only one participant at time 100 ms. B) Results from the forward encoding model applied at the entire stim dataset, for all the participants and at every time point. The same curves from A) have been centered for visualization purposes and straight horizontal lines lines indicate the cuts that will be showcased on C. C) Transversal cuts from B for illustration of intensity values and shape of the centered stimuli curves.

presented in form of decision or response. These stimuli can be seen at Figure 5.2C that showcases the transversal cuts of Figure 5.2B where the same lines correspond to each other, essentially featuring the channel response at 0 - 20 ms, 100 - 120 ms and 240 - 260 ms, alongside the confidence interval. From Figure 5.2C we can assess the fact that the stimuli is indeed higher around 0 - 20 ms and 240 - 260 ms. At the same figure its also shown how these curves have been centered, showcasing the sinusoidal lines where the stimuli is higher.

As mentioned on the chapter 4, another study was performed on whether the changes to the orientation responses from the EEG signals were related to the previous decisions. For this study, we applied the same procedure once again, but this time, we used three different models, one for every presentation of the trial. Then we subtracted the first presentation from the second and last, resulting in Figure 5.3. However no visible correlation can be seen.

Finally, the last analysis consisted on whether the orientation responses from the EEG data were related to the decision taken on the previous trial. In this study, we also wanted to see if this decision taken on the previous trial had an effect on their orientation stimuli. This time the data set was split into two, constrained by whether

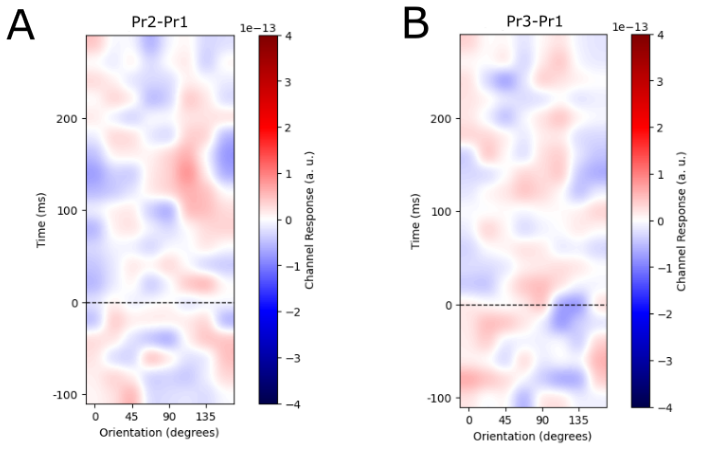
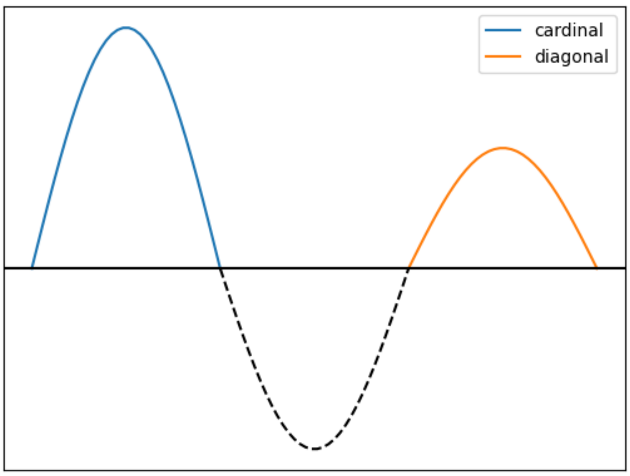


FIGURE 5.3:
Comparison between presentations. A) Difference between the results of the forward encoding model applied on presentation 2 and presentation 1. B) Difference between the results of the forward encoding model applied on presentation 3 and presentation 1.

the previous decision was cardinal or diagonal.

After applying the forward encoding model, we shifted the orientations again, moving the cardinal orientations to the left and diagonal orientations to the right in order to see if the participants were biased to the previous decision. Essentially we want to obtain a shifted transversal cut (Figure 5.2) that is similar to the following graph (case when the previous decision taken was cardinal),



Results of this procedure is portrayed in Figure 5.4. Here we observe on Figure 5.4B, that the stimuli on cardinal orientations are stronger than the diagonal orientations whenever the participants chose cardinal on the previous trial. However on Figure 5.4C, the same thing doesn't happen when the previous choice was diagonal. Nevertheless, from the former finding, it might be enough to be able to assess our hypothesis, that the participants (up to some degree) were biased in their selection.

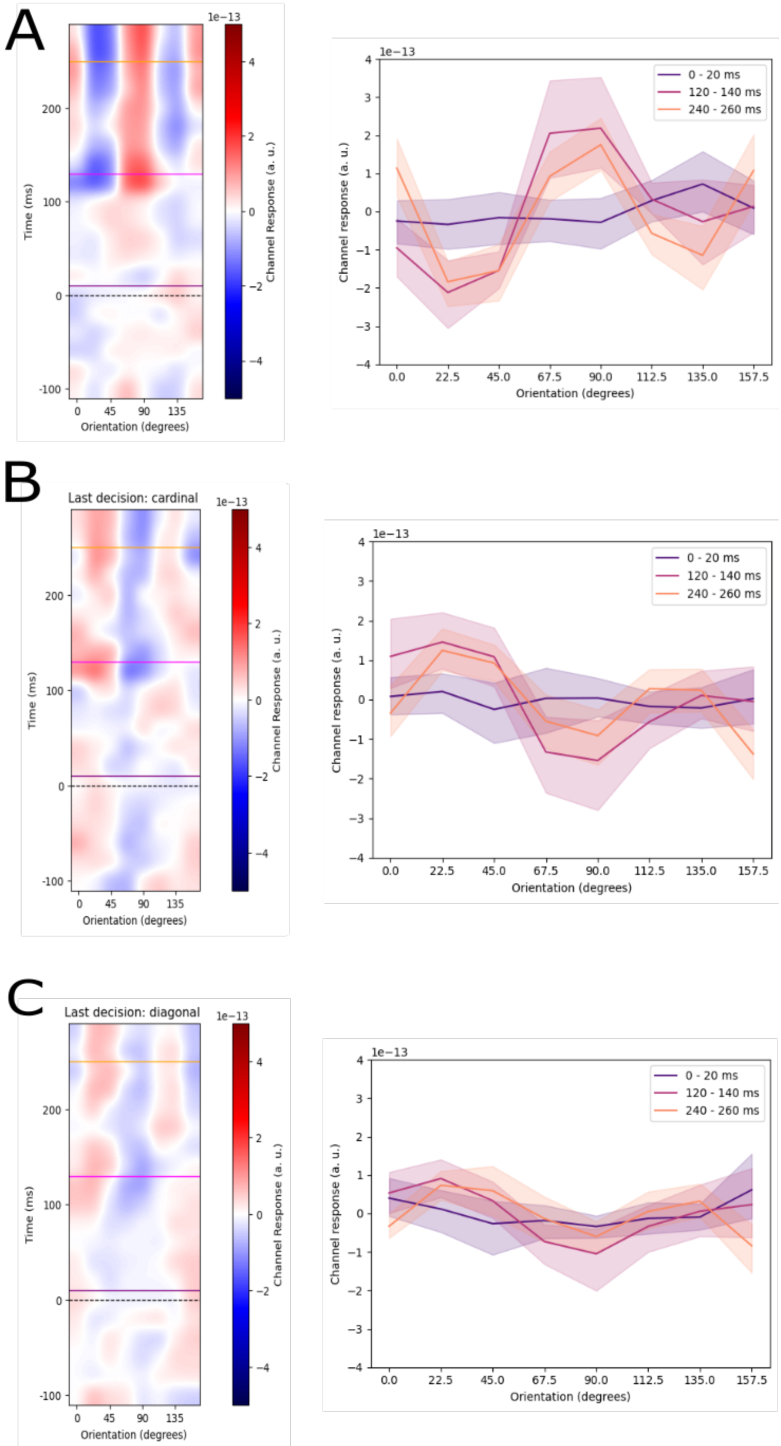


FIGURE 5.4:
Comparison between previous decisions. A) Results of the forward encoding model on the entire data set. B) Results of the forward encoding model applied on the data set where the previous decision was cardinal. C) Results of the forward encoding model applied on the data set where the previous decision was diagonal.

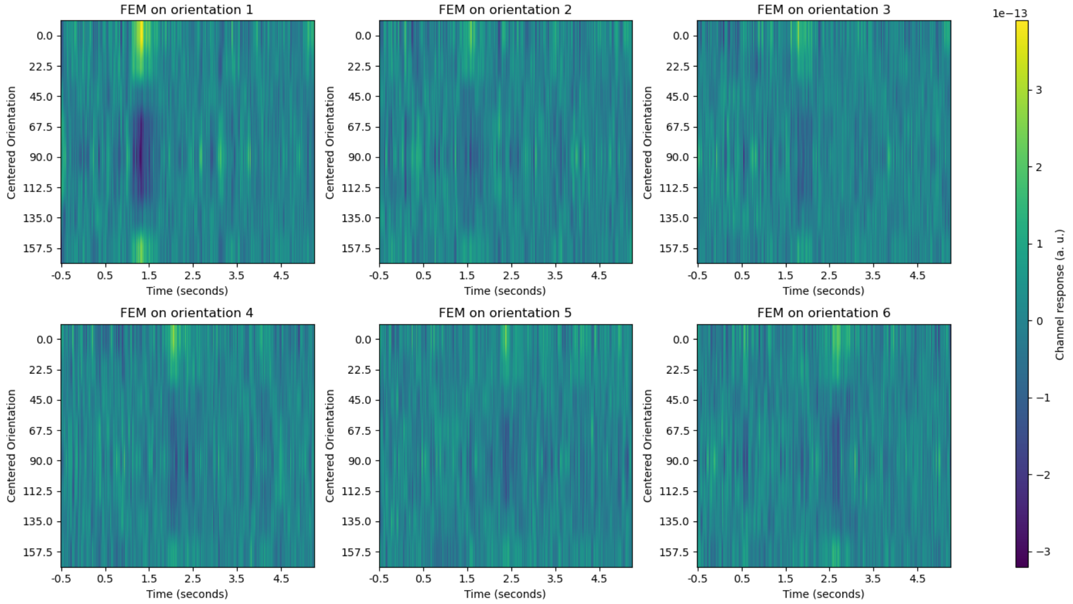


FIGURE 5.5:
Results of the forward encoding model applied to the main data set.

5.3.2 Main Data Set

Same exact procedure was done on the main data set, as seen in Figure 5.5, where we applied the forward encoding model procedure for each of the orientations. We can observe that channel responses become noticeable starting at approximately the 1-second mark and conclude around the 3-second mark. No additional analyses were done following this approach. However these results will be quite interesting for the next section.

5.4 Decision and Response Decoding Results

In this section we will only show the results from the time point classification. More in-depth details of the trivial results from the other methods were treated (although not extensively) on the corresponding notebooks in GitHub.

From the time point-wise classification method, we obtain Figures 5.6 and 5.7, targeting response values and decision values respectively. As stated previously in Chapter 4, we approached the classification problem with SVM and Random Forest Ensembles.

5.4.1 Decision Classification

The results for the decision variable are shown in Figure 5.6. For both the classifiers, no substantial information could be found, since the results from the model trained on the correct labels are within the bounds of the confidence intervals of the model trained on shuffled labels.

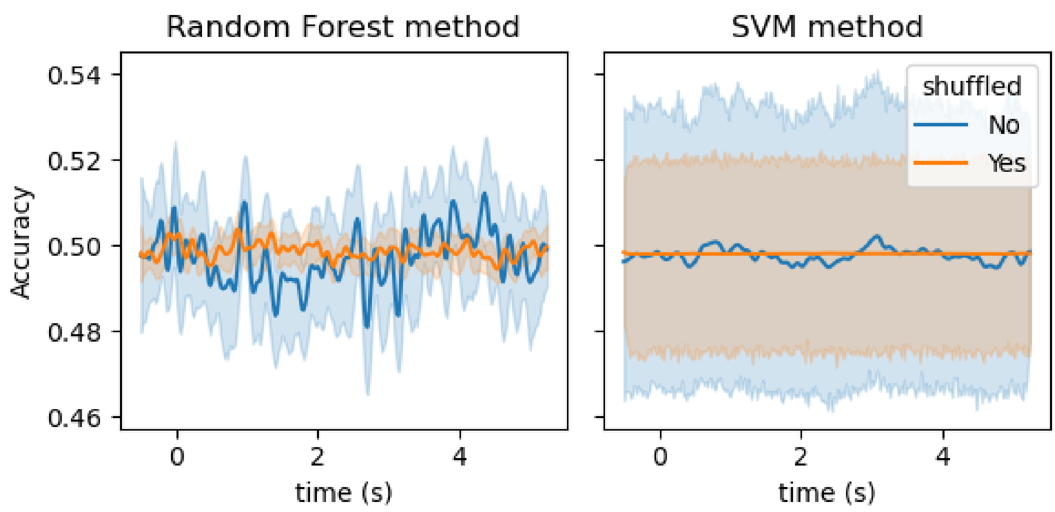


FIGURE 5.6:
Left, time point wise accuracy using Random Forest. Right, time point wise accuracy using SVM, targeting decision value.

5.4.2 Response Classification

However for the response variable, we observe that the results from the model trained on correct labels between the 3-second and 5-second time interval are higher than than the chance level results from the model trained on shuffled labels. (Figure 5.6).

An important insight of this result is the fact that on Figure 5.5, where we applied a forward encoding model on the Main data set, we can see that the orientation stimuli is present until the 3-second mark, whereas in the time point classification method, on Figure 5.6, the accuracy on the response values start to appear around 3.5s, slightly after the orientation stimuli ends. This indicates that the same EEG signals indeed contain insights for handling both experiments.

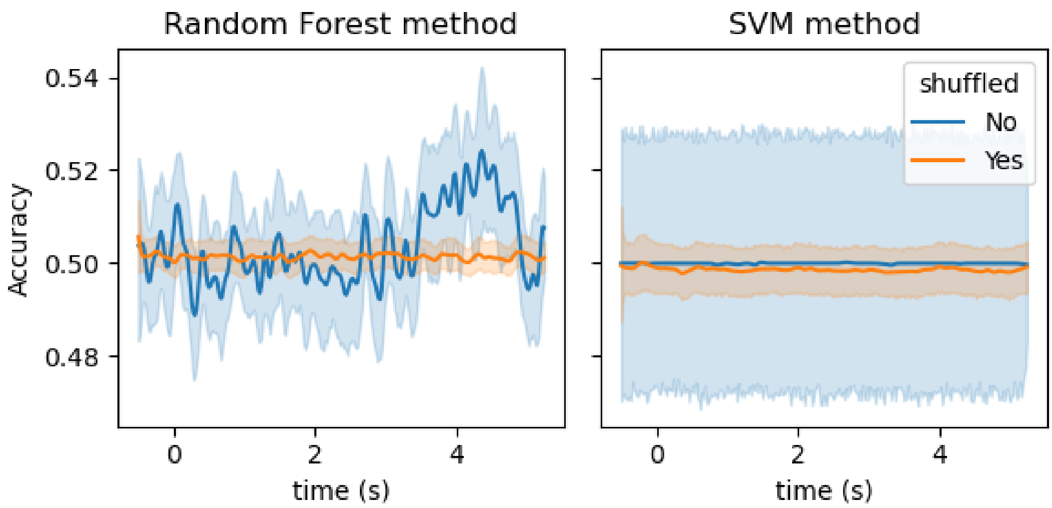


FIGURE 5.7:
Left, time point wise accuracy using Random Forest. Right, time point wise accuracy using SVM, targeting response value.

Chapter 6

Future Work

Due to the large, complex nature of the data, this work was limited by both experimental time and physical hardware constraints. Even though a down-sampling from 500Hz to 50Hz was implemented, some experiments in the previous section would often take 3 hours per experiment and they were mostly run on a Intel-Core i5-13500 CPU. Therefore, the main concern throughout the work was time and memory constrains. For this matter, specialized software, such as the MNE library, is certainly better for the analyses. However, the MNE library provides very dedicated and specific Machine Learning algorithms that might not align with our objectives. Therefore, something related to an adaptation of the MNE code might be the solution for this concern.

Regarding the forward encoding model from Section 4.1, a lot of hyper-parameter searching has been done although not documented since the results are not quantitative. Following work from this section might be to perform a temporal generalization, i.e. testing the performance of the trained model on data points or time steps that were not part of the training set but are from the same temporal domain. This helps assess whether a model has learned to generalize its predictions or classifications to unseen time points or sequences. Some work regarding the forward encoding model applied to the Main data set was done in 5.3.2 with meaningful results, however as stated, the work was discontinued. Perhaps the entire procedure that was done for the Stim data set could have been also done for the Main data set to see how the model performed on the data that contained the entire trial.

In the LSTM network, depending on how we structured the data for tensor conversion we would often run into Kernel crashing problems [1], therefore more approaches regarding this method would be batch training data or different data structuring. Moreover, a different LSTM network with other layers and hyper-parameters might have given the desired results. Also, the implementation of other Recurrent Neural Networks might also be helpful.

About the time point classification, since we could extract relevant information using the Main data set (recall the results on Figure 5.5 and Figure 5.6 from the previous chapter), future work regarding this procedure might also consist of a better hyper-parameter searching or implementation of other classifier, like XGBoost or CatBoost (however, this approach would exponentially increase the time computation), although quantitative results might not be needed for relevancy, since we only aim to decode any information regarding decision and response within the EEG data. Another approach, given the findings on the time point classification experiment, would be to perform something similar to a sliding window. Although we implemented a simple version of it (we did not manage to classify above chance level), a more extensive approach could be done.

[1] The error was always memory problems, even though for the LSTM network we used a NVIDIA RTX4070ti GPU, with CUDA and a Linux setup.

Finally, for each study, a statistical test might have been necessary to assess the statistical significance of the study's results. Nevertheless, the current results can be evaluated as evidence of information decoding beyond chance by comparing them to surrogate data and assessing their conformity with expected results.

Chapter 7

Conclusions

In this project studied the concept of perceptual decision making in humans. Our behavioral results showed that participants' performance improved with the number of repetitions, demonstrating that human observers integrate information over time. In addition, we observed that participants were biased to repeat previous choices. Next, we looked at neural activity in order to explore how the number of repetitions and previous choice affected the representation of decision information. Using a forward encoding model, we successfully decoded stimuli orientations within each sequence.

Importantly, we found that although neural responses decreased with repetition (Figure 3.2), the information that could be decoded remained constant across repeated presentations. However, partial results from the previous decision variable was present on the neural data.

Finally, the later experiment from the decision and response decoding section, shows us that there is indeed information in neural activity regarding the response of the participants but we were not able to decode the decision variable with the current methods.

Overall, the uncertainty of the exact pipeline for a given study and the fact that data sets we encountered in this study proved to be considerably more complex and distinct from the conventional two-dimensional data sets we encountered during our academic journey, highlight indeed the intricate nature of the human brain.

Appendix A

Electrodes

The following table contains the corresponding labels of the electrodes that we used throughout the experiments.

Electrode Number	Electrode Label
1	Fp1
2	F3
3	F7
4	FT9
5	FC5
6	FC1
7	C3
8	T7
9	LM
10	CP5
11	CP1
12	Pz
13	P3
14	P7
15	O1
16	Oz
17	O2
18	P4
19	P8
20	RM
21	CP6
22	CP2
23	Cz
24	C4
25	T8
26	Eye
27	FC6
28	FC2
29	F4
30	F8
31	Fp2
32	Fz

TABLE A.1: List of electrodes' labels

Appendix B

Metadata columns

Columns of the Metadata data set corresponding to the neural data of the Main data set.

index
subj
nblock
ntrial
nrep
trial_type
cond-1
cond
rDV
DV
resp
deci-2
deci-1
deci
corr-1
r_map
correct
confi
RT
d1
conf_lvl
correct-1
d2
d3
d4
d5
d6
o1
o2
o3
o4
o5
o6
confi-1
conf_lvl-1

TABLE B.1: Metadata columns

Appendix C

LSTM network

Layers of the LSTM network used in 4.2.4.

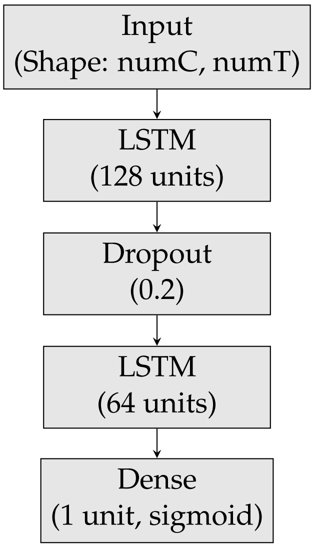


FIGURE C.1: LSTM-based Sequential Model

Bibliography

Bhimraj, K. and R. J. Haddad (2017). "Autonomous Noise Removal from EEG Signals Using Independent Component Analysis". In: *SoutheastCon 2017*. DOI: 10.1109/secon.2017.7925330.

BrainVision (2022). *BrainVision Products*. https://brainvision.com/.

Elsworth, Steven and Stefan Güttel (2020). *Time Series Forecasting Using LSTM Networks: A Symbolic Approach*. arXiv: 2003.05672 [cs.LG].

Garcia, Javier O, Ramesh Srinivasan, and John T Serences (2013). "Near-real-time feature-selective modulations in human cortex". In: *Current Biology* 23.6, pp. 515–522. DOI: 10.1016/j.cub.2013.02.013.

Gazzaniga, Michael S. (2022). *Cognitive Neuroscience: The Biology of the Mind*. Langara College.

Gramfort, Alexandre et al. (2013). "MEG and EEG Data Analysis with MNE-Python". In: *Frontiers in Neuroscience* 7.267, pp. 1–13. DOI: 10.3389/fnins.2013.00267.

Herrera, R. E. et al. (1997). "Vector autoregressive model selection in multichannel EEG". In: *Proceedings of the 19th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. DOI: 10.1109/iembs.1997.756580.

Kamitani, Yukiyasu and Frank Tong (2005). "Decoding the visual and subjective contents of the human brain". In: *Nature Neuroscience* 8.5, pp. 679–685. DOI: 10.1038/nn1444.

Kay, Kendrick N et al. (2008). "Identifying natural images from human brain activity". In: *Nature* 452.7185, pp. 352–355. DOI: 10.1038/nature06713.

Ksibi, Ameni et al. (2023). "Electroencephalography-based depression detection using multiple machine learning techniques". In: *Diagnostics* 13.10, p. 1779. DOI: 10.3390/diagnostics13101779.

Martínez González, Carla L et al. (2022). "Attention deficit and hyperactivity disorder classification with EEG and Machine Learning". In: *Biosignal Processing and Classification Using Computational Learning and Intelligence*, pp. 447–469. DOI: 10.1016/b978-0-12-820125-1.00034-8.

Musk, Elon (2019). "An Integrated Brain-Machine Interface Platform with Thousands of Channels". In: *Preprint*. DOI: doi:10.1101/703801.

Oostenveld, Robert and Peter Praamstra (2001). "The five percent electrode system for high-resolution EEG and ERP measurements". In: *Clinical Neurophysiology* 112.4, pp. 713–719. DOI: 10.1016/s1388-2457(00)00527-7.

Oostenveld, Robert et al. (2011). "FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data". In: *Computational Intelligence and Neuroscience* 2011, pp. 1–9. DOI: 10.1155/2011/156869.

Saeidi, M. et al. (2021). "Neural decoding of EEG signals with Machine Learning: A Systematic Review". In: *Brain Sciences* 11.11. DOI: 10.3390/brainsci11111525.

Salinas, E. and A. Izquierdo (2019). "Decision Letter: Lateral Orbitofrontal Cortex Promotes Trial-by-Trial Learning of Risky, but Not Spatial, Biases". In: *eLife* 8, e49744.015. DOI: 10.7554/elife.49744.015.

St. John-Saaltink, E. et al. (2015). "Expectation suppression in early visual cortex depends on task set". In: *PLOS ONE* 10.6. DOI: 10.1371/journal.pone.0131172.

Sun, Lisha, Ying Liu, and P. J. Beadle (2005). "Independent Component Analysis of EEG Signals". In: *Proceedings of 2005 IEEE International Workshop on VLSI Design and Video Technology*. DOI: 10.1109/iwvdvt.2005.1504590.

Talluri, B. C., A. Braun, and T. H. Donner (2021). "Decision making: How the past guides the future in frontal cortex". In: *Current Biology* 31.6. DOI: 10.1016/j.cub.2021.01.020.

Urai, Anne E et al. (2019). "Author Response: Choice History Biases Subsequent Evidence Accumulation". In: *eLife* 8, e46331.033. DOI: 10.7554/elife.46331.033.

Wyart, Valentin, Anna C. Nobre, and Christopher Summerfield (2012). "Dissociable Prior Influences of Signal Probability and Relevance on Visual Contrast Sensitivity". In: *SSRN Electronic Journal*. DOI: 10.2139/ssrn.4063701.