Sagol School of Neuroscience
Tel Aviv University

Machine Learning Identification of Exploratory Journey During NF Trials

By

Aviv Rabi

Sharon Tirza

Conducted under the supervision of

Prof. Talma Hendler, M.D. Ph.D.

Guy Gurevich, Ph.D.

1. Background

Our project was conducted in Prof. Talma Hendler's laboratory at Sagol Brain Institute in Ichilov. This laboratory is at the forefront of developing methodologies involving audiovisual brain-computer interfaces, virtual reality, and augmented reality for applications in psychiatry, rehabilitation, and cognitive enhancement. The lab's primary focus revolves around pioneering research that integrates these technological advances to address psychopathologies such as depression, anxiety, addiction, cognitive rehabilitation, and learning (Hendler, n.d.). A central aspect of the lab's research involves the utilization of neurofeedback (NF) as a pivotal component in achieving its objectives.

Real-time functional MRI NF (rt-fMRI-NF) stands as an innovative tool for mastering self-regulation of brain activity. In rt-fMRI-NF studies, the blood-oxygen level dependent signal emanating from a designated brain region is quantified and subsequently presented to the subject through an MRI scanner.

In our dataset, data collection was conducted using an EEG scanner. The substantial cost and restricted mobility associated with fMRI present considerable challenges in terms of scalability, accessibility, and cost-effectiveness, particularly for clinical applications. Thus, the limited applicability of rt-fMRI is addressed by leveraging an EEG model equipped with improved spatial resolution to target and observe brain activity (Keynan et al., 2018).

The ultimate goal of the NF approach is to empower the subject to gain control over the activity within the chosen brain area, thereby influencing the corresponding mental state. A prominent domain of study within the realm of Neurofeedback centers around emotion regulation (Caria et al., 2019). Originating in the late 1960s, EEG biofeedback (NF) emerged as a technique for reshaping brainwave patterns through operant conditioning. Since its inception, a substantial body of research has accumulated, demonstrating the efficacy of NF in addressing a variety of conditions (Hammond, 2008).

A "mental strategy" pertains to a cognitive construct encompassing perceptual, affective, cognitive, or meta-cognitive components (Lubianiker et al. (In preparation)). NF offers individuals the ability to learn how to modulate their own brain activity (Kober et al., 2013).

The fundamental learning mechanisms underpinning NF training are associative and implicit, predominantly operating beyond the purview of conscious awareness. Nonetheless, several other aspects of training that contribute to the outcomes are accessible to conscious processing. For instance, the results of sensorimotor rhythm (SMR) up-regulation training align with the strategies articulated by participants (Autenrieth et al., 2020). Earlier research further establishes connections between NF regulation levels across targets and certain features of reported mental strategies (Lubianiker et al. (In preparation)).

Notably, it has been discovered that individuals who do not experience initial success may encounter frustration, potentially leading them to invest excessive mental effort by employing various mental strategies. This phenomenon could hinder performance and impede subsequent learning endeavors (Kober et al., 2013).

These observations prompt inquiries into whether the exploratory journey within the realm of mental strategies influences the outcomes of NF success.

Therefore, our project aimed to facilitate the laboratory's enhanced comprehension of the associations and potential correlations between subjects' exploratory journeys and their rates of regulatory success. Our primary goal is to generate a set of features that encapsulate these exploration journeys by extracting them from the data. Following this, we plan to input these features into a machine learning algorithm, aiming to predict success rates effectively. Secondly, we will conduct a more comprehensive analysis of these predictions, seeking to gain deeper insights into the key features that distinguish a successful exploration journey. Lastly, we aim to provide the laboratory with a tool that facilitates ongoing research, allowing the application of our algorithm to data from future studies encompassing diverse brain regions.

2. **Methods** (refer to Appendix 1)

2.1. Data Collection

The dataset utilized in our study was curated by Nitzan Lubianiker and Avigail Lerner for their research titled "Quantitative Assessment of the Link Between Mental Strategies and Regulation Success". Following each NF training session, participants provided descriptions of the mental strategies employed during each cycle of practice. These verbal reports were subsequently categorized by third-party raters using a novel questionnaire known as the Mental Strategies Questionnaire for NF (MSQ-NF, detailed in appendix 2). The MSQ-NF encompassed categories related to psychological dimensions, content, and manner features.

For each strategy, a rater determined whether it included each of the MSQ features or not. This process resulted in a 51-dimensional binary vector representation for each mental strategy within a regulation cycle. The study involved a cohort of 66 participants, each of whom participated in a total of six sessions. Within each session, there were five individual runs, contributing to a grand total of 1371 vectors for analysis. Notably, certain participants exhibited instances of missing data within their submissions.

2.2. Data Preprocessing

Our data necessitated preprocessing for two primary reasons. We encountered cases of missing data, as there were instances of "NaN" values present. To delve into further detail, the binary vector, as noted earlier, served as a representation of various aspects of the mental strategy. However, some aspects posed challenges in terms of determination and consequently appeared as question marks within the raw data. In addressing this issue, we adopted a strategy of replacing all aforementioned values with the value 0.5. This decision was reached following thorough consideration, as we concluded that certain aspects couldn't be definitively classified as either present or absent in this context.

Subsequently, we proceeded to extract features that exhibited either low variance (below 0.025, in alignment with the lab's guidelines) or a high correlation with another feature (correlation exceeding 0.75, as per the direction from our lab advisor). The second challenge pertained to missing data. As previously mentioned, each subject participated in a total of six sessions, each comprising five runs. After meticulous evaluation of the dataset, we opted to retain only subjects with five sessions or more. For participants who completed all six sessions, we chose to exclude the first session — a common practice since this session typically demonstrates the lowest success rate in neurofeedback tasks. This approach was undertaken to ensure the utilization of vectors with consistent lengths, facilitating more homogeneous data analysis (graphic view is shown in Appendix 3).

2.3. Feature Extraction

The features we derived for each subject are inherently related to the inter-vector similarities. This choice was driven by our primary interest in comprehending the subject's exploratory journey, rather than focusing on the specific content of the mental strategy. The parameter we elected to use is the soft cosine similarity (refer to Appendix 4). This parameter affords a similarity metric that capitalizes on the similarity between features. In our context, this similarity was quantified through a correlation matrix constructed from the entire database.

For every subject, each of the five sessions is encapsulated by the average soft cosine value calculated across each pair of runs within that session. The subsequent ten features encompass the soft cosine value computed for each pair of sessions, where a session is defined as the mean vector across its constituent runs. In totality, each subject is characterized by 15 features: one for each session and $\binom{5}{2} = 10$ reflecting inter-session relationships (refer to Appendix 5 for visual representation and to Appendix 6 for a pairing function). Sessions composed of a solitary run were addressed by substituting the mean soft cosine value from their remaining sessions.

2.4. Prediction model

We employed a clustering algorithm, the K-means algorithm, to group the subjects into two distinct clusters, leveraging the extracted features. Building upon these cluster assignments, we subsequently applied a Random Forest algorithm to predict and assess the accuracy of the prior clustering outcomes.

We selected the Random Forest algorithm primarily due to the imbalanced classifications observed with K-means. After carefully examining various other algorithms including logistic regression, gradient boosting classifiers, and SVM, we noticed that they consistently yielded predictions limited to a single class. In contrast, the Random Forest algorithm was the only one capable of generating predictions encompassing both classes.

For the K-means classification, we divided it into a training group and a testing group, allocating a 30% portion to the test set. The accuracy score of the K-means classification was determined by calculating the percentage of correct predictions made by the Random Forest algorithm. To ascertain the robustness of our predictions, we repeated this process five times. During each iteration, we randomized the designation of samples as either the test or training set, enabling a comprehensive assessment of the model's performance across different partitionings of the data.

The performance measure utilized as “success rate” was calculated as the average success rate of the final session minus that of the first one.

The reason for this decision was our greater emphasis on monitoring and enhancing performance over the course of their sessions, rather than solely focusing on the raw success rate. This is because the raw success rate can vary among subjects due to their inherent ability to naturally regulate their brain activity, without direct consideration of their capability to learn from neurofeedback training procedures.

As a point of reference, we explored alternative approaches to predicting success rates. Initially, we examined each subject's mental strategy vectors and condensed them into a single vector by calculating the feature-wise mean, effectively representing a mean mental strategy. Subsequently, we determined the success rates by computing the mean of each subject's success rates across the entire trial. The prediction of these success rates was achieved using a linear regression model. We estimated the model’s success using Pearson’s

R coefficient (refer to Appendix 7). Similarly, we repeated this process five times. After reviewing the results, a choice was made to embrace an alternative approach: an endeavor to predict the mean success rates per session. For every subject, we generated mean mental strategy vectors per session alongside corresponding mean success rates. In this instance, our emphasis shifted horizontally, directed towards predicting success rates within each session across all subjects. Similar to the previous approach, we assessed the model's efficacy using the Pearson's R coefficient.

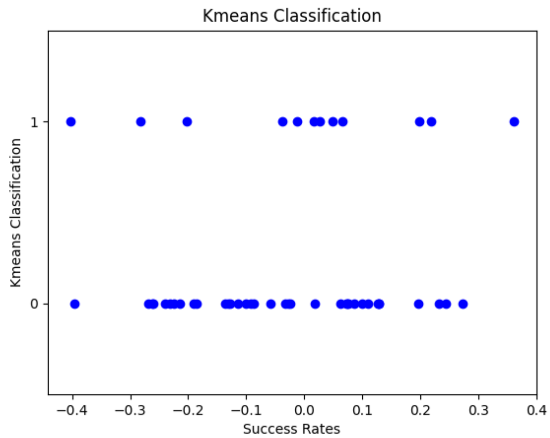
2.5. Exploratory Research

Utilizing the K-means algorithm classification, participants were divided into two distinct groups: 'successful' and 'unsuccessful,' determined by the mean success rate calculated by subtracting the mean success rate of the first session from that of the last session, as previously described.

For evaluating feature significance, we took into account two crucial parameters: the absolute disparity between the mean values of that feature within the 'Successful' and 'Unsuccessful' groups, and the general variance of values across all subjects for that particular feature, to act as a control and ensure the mean difference was not affected by extreme values.

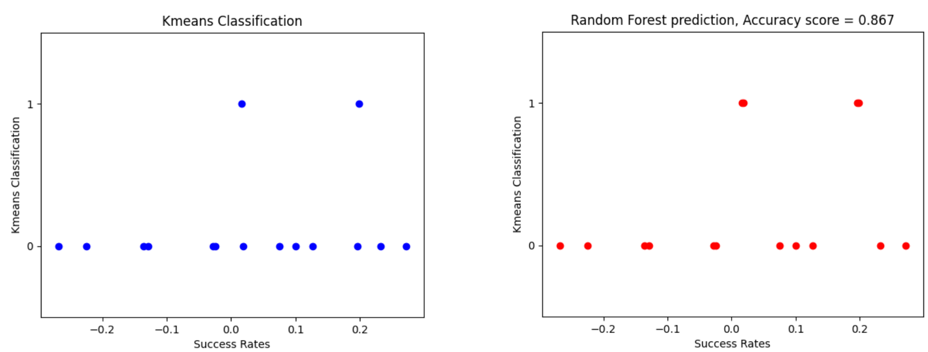
Additionally, we generated a vector to represent the mean exploration journey for each group. This vector was computed by calculating the feature-wise mean for all subjects within their respective group.

3. Results



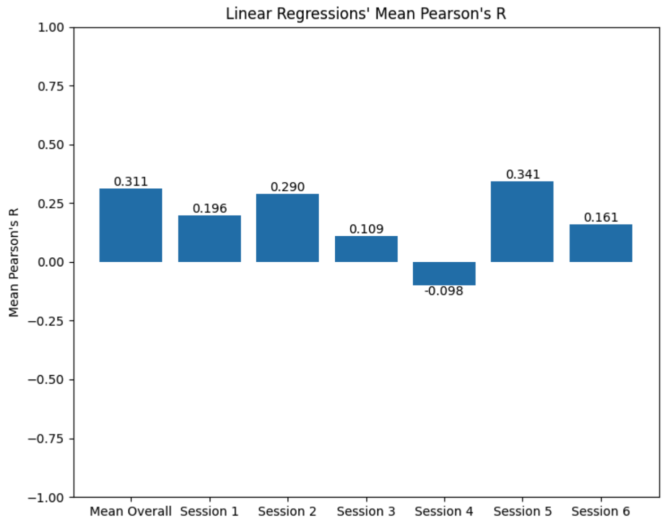
[Figure 1 - K-means classification]

The K-means algorithm divided the subjects into two groups, aligning with our focus, relying on the feature vectors. As depicted in Figure 1, this classification exhibits an imbalance, predominantly favoring class "0".



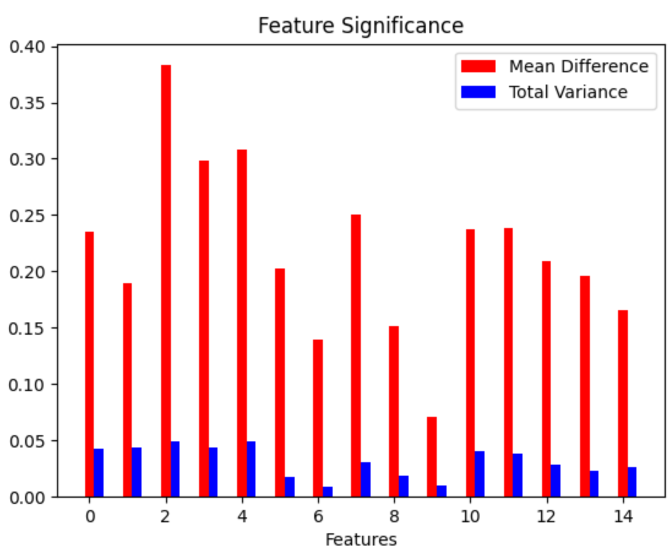
[Figures 2-3 - Random forest prediction vs. K-means classification]

The above Figures display the predictions made by the Random Forest model in comparison to the actual classifications generated by K-means clustering, which were subsequently employed to compute the accuracy score for K-means classification, as detailed in the Methods section. The highest attained accuracy score was 0.867, with an average score of 0.72 across the five iterations.



[Figure 4 - Linear Regression across all runs]

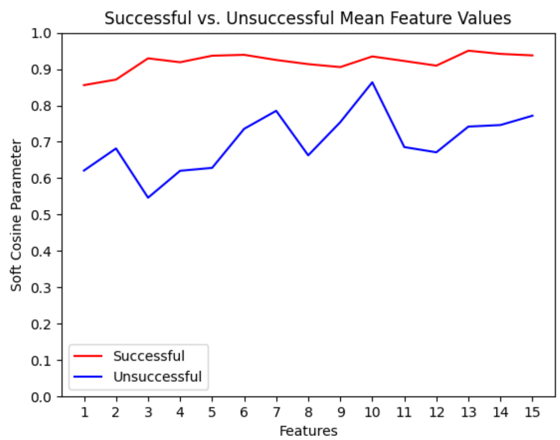
As illustrated in Figure 4, the average Pearson's R coefficient, computed over five iterations, indicates a relatively low value, implying a limited degree of success in predicting both on a per-session basis and in terms of the overall mean.



[Figure 5 - Feature significance]

As depicted in the bar graph in Figure 5, the features that exhibited the most notable disparity between subjects categorized as successful and those classified as unsuccessful by the K-means algorithm are the third, fourth, and fifth features. These features pertain to the corresponding sessions (refer to Appendix 6).

The two classes were categorized as 'Successful' and 'Unsuccessful' based on the subjects' success rates. The mean success rate for individuals in class '0' was -0.041, whereas for those in class '1', it was -0.0002. It is worth noting that a lower value indicates a more pronounced level of improvement in success in regulating brain activity throughout the NF trial. Consequently, class '0' was designated as 'Successful,' and class '1' as 'Unsuccessful.' Subsequently, we computed a vector comprising the mean values per feature for each class.



[Figure 6 - Successful vs Unsuccessful]

The graph in Figure 6 illustrates the relationship between success and lack of success with respect to the soft cosine metric.
In this graph, a notable dissimilarity becomes evident, notably in the third, fourth and fifth sessions. Remarkably, the soft cosine value consistently remains at a high level for the successful group. Conversely, the unsuccessful group exhibits consistently lower soft cosine values across all features.

4. Discussion

Our initial milestone involved identifying distinct patterns of behavior, specifically focusing on the strategies subjects employed when transitioning between different mental approaches. Given that each subject underwent multiple runs during which they could alter their mental strategy, our primary task was to preprocess the data and extract pertinent features. When attempting to predict success rates from the content of the mental strategies, both per subject and per session, we encountered fairly low prediction rates (peaking at 0.341). This finding, combined with previous knowledge of the effect of mental strategies on NF success rate, led us to believe there was more to the data than is revealed using this technique. Through the application of our methodologies, a comparatively high prediction rate was attained (0.867 using K-means), contrasting with said results for the control method. This disparity implies that a more comprehensive understanding can be derived from the exploratory journeys of the subjects. Building upon this achievement, we can delve into data analysis to discern the distinguishing features that underlie the projected elevated success rate. In accordance, our second milestone aimed to gain fresh insights into the potential to guide subjects in discovering the most suitable mental strategy for them based on the observations made in the first milestone. Using the obtained outcomes, we conducted an exploratory analysis to ascertain the significance of the different features we extracted, to deepen our comprehension of the components that should ideally constitute an optimal mental strategy journey. When examining the features representing the mean soft cosine parameter within sessions (See appendix 5), Our feature significance analyses revealed that the most pronounced distinction between subjects characterized as successful and unsuccessful materialized within the third, fourth and fifth features, representing corresponding sessions (See appendix 6). We desired to look further into this matter and while examining the mean values per feature across both classes ('Successful' and 'Unsuccessful') we see this discrepancy manifested through fluctuations in the chosen mental strategy.

Successful subjects consistently exhibited higher and stable soft cosine values, indicating a greater similarity in their mental strategies during and across sessions. In contrast, the unsuccessful group showed lower soft cosine values, suggesting more pronounced fluctuations in their mental strategies across runs within each session. When we shift our focus to features 6-15, we observe a second set of prominent features, specifically 7 through 12. These features represent the soft cosine values between sessions, as detailed in Appendix 6. Interestingly, there is a lack of stability in the soft cosine values for these features. This introduces another dimension of stability in the mental strategy choices of the successful class, contrasting with the variability seen in the unsuccessful class.

When examining the unsuccessful class, it's essential to underscore that the lower values of the soft cosine value within sessions (features 1st-5th), as compared to the values between sessions (features 6th-15th), suggest an absence of stability in their strategy choices within individual sessions. However, there is a higher degree of similarity or a narrower range of exploration between sessions.

We would like to put forth an interpretation for the change in the soft cosine value starting from the 6th feature. This observation leads us to suggest that unsuccessful subjects were more inclined to modify their selected mental strategy across runs within individual sessions. However, they tended to maintain relatively similar mental strategies across multiple sessions.

Simultaneously, it's worth noting that the soft cosine values for transitions between sessions consistently remain lower than those observed in the successful group. This observation leads us to infer that the successful subjects drew insights from previous sessions to identify more advantageous mental strategies, which they then adhered to. The graph illustrates how successful subjects swiftly grasped the most beneficial strategy for them (as evident in the rise from the 1st to the 2nd feature) and maintained a consistent mental strategy.

Conversely, the unsuccessful group appears not to have reached similar conclusions. Notably, the soft cosine values demonstrate stability and relatively high values within the later features (13th-15th), highlighting the similarity between sessions. This pattern suggests a preference among unsuccessful subjects for maintaining consistent or closely related mental strategies to the more beneficial ones they had identified, particularly as they progressed toward the end of the trial.

In summary, our findings lead us to propose an optimal approach for achieving success in neurofeedback sessions. It involves maintaining consistency in the choice of a mental strategy throughout runs within a session. Simultaneously, subjects should, once found, stay with a mental strategy proven effective, when transitioning between sessions. This trend of high similarity both within and between sessions enables the subjects to practice their chosen mental strategy and better their regulating skills across the entire trial. Thus utilizing the long process, as opposed to the unsuccessful subjects who seem to not infer from previous sessions.

As the dataset expands, its capacity to precisely determine the most fitting journey for each brain region, from which the data originates, is enhanced.

To facilitate the repeated application of the algorithm within the lab's framework, we introduced a Python-based GUI application (see Appendix 8). This application permits convenient execution of the algorithm whenever necessary, offering substantial utility to the lab. The GUI serves two primary purposes: firstly, when there's an update to an ongoing trial involving new subjects, necessitating an augmentation of the algorithm's dataset. Secondly, it could be utilized when new research emerges that the lab intends to explore using our algorithm. The entire project we engaged in has been written as a code package (see Appendix 9), intended for future use by the laboratory. The GUI illustrates the desired input (as depicted in Appendix 8), and accordingly, the user obtains the plots that we have introduced in the results section (Figures 1-3, 5-6).

Throughout the course of our work, we encountered several challenges. Firstly, the dataset's scale was relatively modest, encompassing a total of 66 subjects. Moreover, the dataset exhibited some sparsity, manifesting as instances where subjects were absent for entire sessions or, at times, multiple runs within sessions. Additionally, the raw data contained instances of missing information represented by question marks, as expounded in the Data Preprocessing section of the Methods. Successfully navigating these challenges necessitated numerous subjective judgments on our part, which likely impacted the scale of our results and conceivably even led to false positives.

Furthermore, intriguingly, the correlation between mental strategies and success rates does not exhibit a progressive increase with each subsequent session. While one might intuitively anticipate such an increase and correspondingly anticipate more straightforward model predictions, which in turn facilitates a heightened Pearson's R coefficient. Our findings diverge from this presumption and no steady rise in the correlation coefficient is shown as the sessions progress. This prompts us to propose the pursuit of additional research to delve into the reasons underlying this observation, aiming to determine the factors contributing to the absence of a consistent rise in the correlation between success rates and mental strategies across sessions.

Additionally, we recommend a potential avenue for further research involving informed subjects. In this proposed study, participants would be apprised of the insights drawn from our data analysis. The aim would be to assess whether the application of these conclusions leads to increased success rates during neurofeedback sessions. By revealing our findings and their implications, we encourage the investigation of whether participants who actively incorporate this knowledge can enhance their performance and attain more favorable outcomes. Such further research could provide valuable insights into the practical applicability of our observations in real-world scenarios.

5. Bibliography

Autenrieth, M., Kober, S. E., Neuper, C., & Wood, G. (2020, May 14). *How much do strategy reports tell about the outcomes of neurofeedback training? A study on the voluntary up-regulation of the Sensorimotor Rhythm*. Frontiers. https://www.frontiersin.org/articles/10.3389/fnhum.2020.00218/full

Caria, A., Kadosh, K. C., Cyders, M. A., Gucht, V. D., Emmert, K., Greer, S. M., Groenewold, N. A., Hamilton, J. P., Hellrung, L., Herwig, U., Johnston, S. J., Kalisch, R., Kohn, N., Koush, Y., Lawrence, E. J., MacDuffie, K. E., Mataix-Cols, D., Misaki, M., Morawetz, C., … Gröne, M. (2019, March 9). *FMRI neurofeedback in emotion regulation: A literature review*. NeuroImage. https://www.sciencedirect.com/science/article/pii/S1053811919301788

Hammond, D. C. (2008, September 7). *What is neurofeedback?: Journal of neurotherapy: Vol 10, no 4*. What Is Neurofeedback? https://www.tandfonline.com/doi/abs/10.1300/J184v10n04_04

Hendler, T. (n.d.). *Immersive Brain-Computer Interfaces - Sagol Brain Institute*. brain-research. https://www.cbf-tlv.com/immersive-brain-computer-interfaces

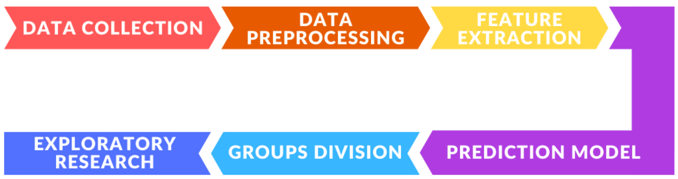
Kober, S. E., Witte, M., Ninaus, M., Neuper, C., & Wood, G. (2013, October 2). *Learning to modulate one's own brain activity: The effect of spontaneous mental strategies*. Frontiers. https://www.frontiersin.org/articles/10.3389/fnhum.2013.00695/full

Keynan, J. N., Cohen, A., Jackont, G., Green, N., Goldway, N., Davidov, A., Meir-Hasson, Y., Raz, G., Intrator, N., Fruchter, E., Ginat, K., Laska, E., Cavazza, M., & Hendler, T. (2018, December 10). *Electrical fingerprint of the amygdala guides neurofeedback training for stress resilience*. Nature News. https://www.nature.com/articles/s41562-018-0484-3

Lubianiker et al. (In preparation)

6. Appendices

Appendix 1 - Methods Pipeline



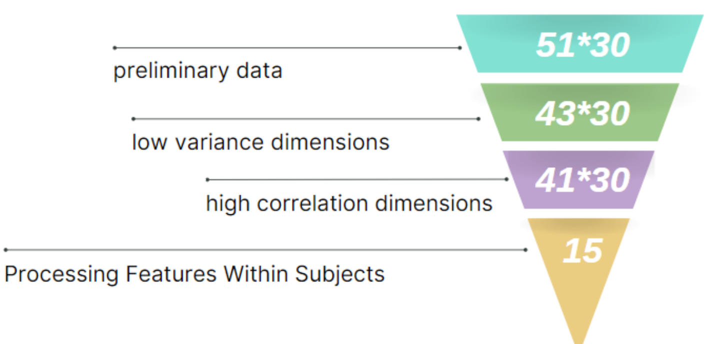
Appendix 2 - Mental Strategies Questionnaire for Neurofeedback

Content																					Manner			Psychological dimensions		
Current sensations											Episodic/Semantic															
Sensory exteroception					Somatic sensations				Affect		Imagery exteroception															
Vision	Auditory	Smell	Taste	Tactile	Pulse	Visceral sensations	Breathing	Muscle sensations	Emotion	motivation	Vision Imagery	Auditory Imagery	Smell Imagery	Taste Imagery	Tactile Imagery	Memory or imagination	Motor imagery	Lingual	Conceptual/arithmetic	Interaction with other people	Engaged or detached from interface	Involving one or many strategies	Rhythmic	Arousal	Valence	Agency

No	No	No	No	No	?	?	Yes	No	Yes	Yes	No	No	No	No	Yes	Yes	Yes	?	No

" I was recalling a concert I attended to, remembered how I danced and felt excited and happy"

Appendix 3 - Data Preprocessing and Features Extraction per Subject

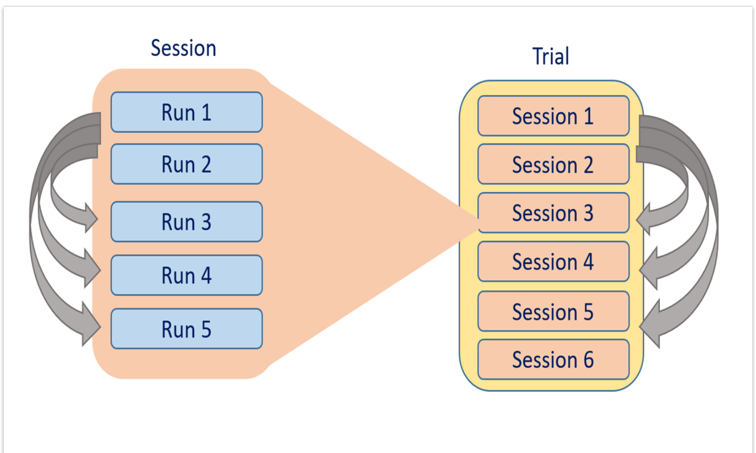


Appendix 4 - Soft Cosine Calculation

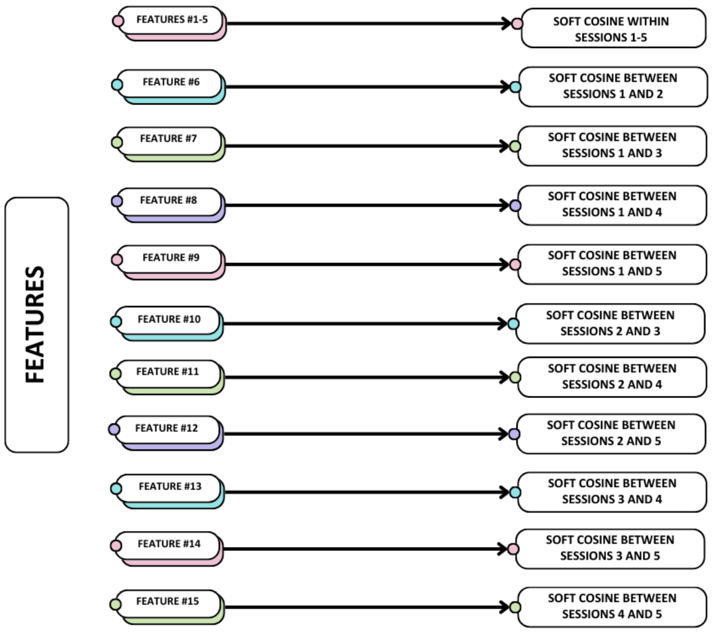
$$\text{soft\_cosine}_1(a, b) = \frac{\sum_{i,j}^{N} s_{ij} a_i b_j}{\sqrt{\sum_{i,j}^{N} s_{ij} a_i a_j} \sqrt{\sum_{i,j}^{N} s_{ij} b_i b_j}},$$

where s_{ij} = similarity(feature$_i$, feature$_j$).

Appendix 5 - Feature Extraction Representation



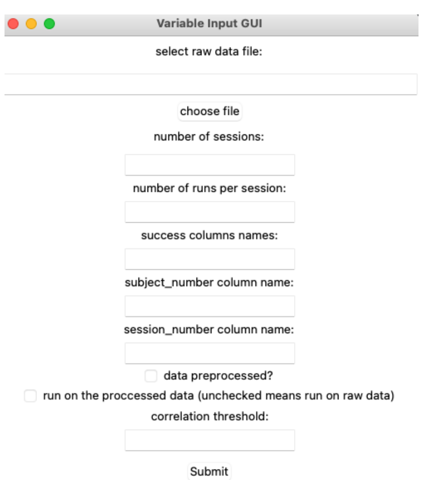
Appendix 6 - Pairing Function from Features to Their Meaning



Appendix 7 - Pearson's R Coefficient

$$r = \frac{\sum (x - m_x)(y - m_y)}{\sqrt{\sum (x - m_x)^2 \sum (y - m_y)^2}}$$

Appendix 8 - Graphic User Interface



Appendix 9 - Link to Git Repository

https://github.com/AvivR94/CS-Brain-Research-Workshop