Tjalle Galama 6936385 - Utrecht University 2021 - supervisors prof. dr. J.L. Kenemans, dr. C.P. Janssen

Bottom-up prediction of cognitive control features in EEG signals

Linear classification of pre-stimulus stop-signal task response inhibition

May the Sun of Righteousness Enlighten Us

Abstract

The prediction of failures in response inhibition is potentially very valuable for understanding pathologies as well as for the support of critical operations carried out by humans. To realize this aim, Fisher's linear discriminant analysis with leave-one-out cross-validation has been performed per 1 Hz frequency interval on the 1000 milliseconds electroencephalogram Fourier-transformed periodograms preceding go-stimuli that were followed by a stop-signal in an auditory stop-signal task, using the power values per electrode as features. Ocular correction was applied, while no other artifacts were removed. This repeatedly class-balanced two-class classification resulted in significant accuracies for 21 of the 29 participants. It was hypothesized that both theta oscillations originating from the prefrontal cortex during anticipation as well as a balance between alpha oscillations over visual and those over auditory cortex both would promote successful stopping. The results translate to a partial validation of these theories. In future studies, the procedure can be extended and this intuitive multivariate pattern analysis machine learning method may continue to yield insights for research in artificial intelligence and human computer interaction.

Introduction

This research project is about predicting inhibitory failures from pre-stimulus electroencephalography brain activity and is continuing on theories about how certain cortical activation patterns implement cognitive control, for which frontal theta is a compelling candidate (Cavanagh & Frank, 2014). We can define or theorize cognitive control as a resource which enables an individual to inhibit responses to irrelevant stimuli and focus on relevant information. The ability to suppress unwanted or inappropriate actions and impulses ('response inhibition') is a crucial component of flexible and goal-directed behavior. Response inhibition is essential for navigating everyday life and its derailment is considered integral to numerous neurological and psychiatric disorders, and more generally, to a wide range of behavioral and health problems (Verbruggen et al., 2019).

There are a few indications that behavior which has not yet taken place can be predicted from EEG, but we do not know yet to what extent this is possible. Specific neural signatures of attentional lapses are registered in the EEG up to twenty seconds before an error (O'Connell et al., 2009). In a Go-noGo task, elevated occipital alpha and sensorimotor mu activity just prior to the presentation of the stimuli predicted an upcoming error, wherein an error resulted in increased frontal theta activity and decreased posterior alpha activity (Mazaheri et al., 2009). Elevated alpha power in parieto-occipital brain areas and elevated levels of theta power are observed in frontal brain areas when auditory attention is sustained successfully (Keller et al., 2017).

A dataset is used to try and determine the predictability of failed response inhibition. Herein, EEG is recorded when subjects perform a choice task. After a split-second interval, the visual go-stimuli are occasionally followed by an auditory signal - dictating that all ongoing actions are stopped or suppressed. This split-second interval is manipulated to induce about

50% failed stops. A pilot study using this data investigated whether the above alpha and theta dynamics also apply in relation to response inhibition (Kandiah, 2020). In the failed stop condition there is a trend towards more elevated alpha power in frontal brain areas in comparison to parieto-occipital brain areas. Moreover, pre-stimulus theta power in frontal brain areas is significantly higher in the successful stop condition compared to the failed stop condition. Consequently, it is hypothesized that strong proactive cognitive control ('attend', 'get ready to go or stop') during anticipation of the signals promotes successful stopping, wherein it is assumed that cognitive control manifests as theta oscillations originating from the prefrontal cortex. It is also hypothesized that susceptibility ('sensory readiness') to either visual or auditory signals promotes successful stopping, which originates in the balance between EEG-alpha oscillations over visual and those over auditory cortex.

A recurring question of methodological origin within the cognitive sciences is to what degree one has to test a (very) specific hypothesis to be able to foresee what action will happen, or whether a bottom-up approach (of pattern recognition) might do the trick. The above hypotheses will be tested using machine learning techniques, using the data prior to the stimulus as a starting point. Our main operational question of research that will be dealt with is: what are the predictors of performance failures from a perspective that goes beyond theta and alpha in terms of frequency and location? Most generally put we thus ask: what patterns can be found in the data? We can expect that the classifier will rely on the theta and maybe also on the alpha frequency band.

This investigation may over time also contribute to solutions for the creation of external intelligence, since computational architectures are inspired by contemporary progress in neuroscience (Haber et al., 2015; Lieto et al., 2018). Research in machine learning and brain-computer interaction is characterized by its plurality of methods and approaches. These scientific fields can foster the validation and development of links between cognitive theory and neurophysiology of human behaviour (Vahid et al., 2020), proceeding on the conjecture that every aspect of intelligence can in principle be so precisely described that a machine can simulate it and hence make predictions about it.

Method

Participants

The data of 29 participants was obtained from (Kenemans, van der Heiden, Van Bijnen & Logemann, in preparation). The subjects (age: M = 22.8, SD = 3.26) were all healthy and studying at Utrecht University, having normal hearing and normal vision. Informed consent (approved by the ethics committee of the University Medical Centre Utrecht) was signed.

EEG data acquisition

For recording, the ActiveTwo Biosemi system with 64 Ag-AgCl electrodes has been used. The electrodes were placed following the 10/10 system. EOG electrodes were positioned below and above the left eye as well as at the outer canthi of both eyes. The signals were sampled at 2048 Hz and online referenced to the Common Mode Sense/Driven Right Leg electrode and low pass filtered at DC to 400 Hz.

Procedure of the stop-signal task

After placing the EEG cap, the participants took place in a dark sound-attenuated chamber, sitting at a distance of approximately 90 centimeters from the computer screen. The participants were asked to respond as fast as possible to the visual go-stimuli and to refrain from responding when a stop-signal was presented after the go-stimulus. During the task, participants had to discriminate between “X” and “O” by pressing with either the left or right index finger. These visual go-stimuli were presented on the screen for 150 milliseconds, slightly above a central fixation cross. For preventing expectation-effects and making the onset of the visual go-stimuli unpredictable, the trial-to-trial interval changed between 1.5 to 1.8 seconds. The stop-signal task had 25 blocks in total. The first block consisted of 126 go-trials, which were used for practice and to determine a reaction-time baseline. Hereafter, two visual stop-signal conditions and one auditory stop-signal condition were presented, in a counterbalanced order across participants. Note that in our investigation only the data of the auditory stop-signal condition is used. This stopping tone of 1000hz and 72dB was presented binaurally for 150 milliseconds through headphones. Each of the three conditions started with a base-block, to estimate a go-stop interval: the stimulus-onset asynchrony. In each base-block the SOA value was fixed at 250ms. After the base-block, three more blocks would follow, each consisting of 128 trials, wherein the interval between the go - and stop stimulus was jittered over 99ms around the SOA value. After completing these four blocks, the index finger-assignment was switched before starting with another base-block and three experimental blocks. The SOA-value of subsequent blocks was based on the stop rate from previous blocks and optimized by a tracking algorithm (De Jong et al., 1995) to yield an approximate stop rate of 50%. A comparison between the mean reaction time on the go-stimuli with the go-stimuli of the initial practice block was made after each experimental block. Participants were asked to increase their speed when the average reaction time of an experimental block was higher than one and a half times the average of the practice block. They were told to slow down when having less than 40 percent successful stops. A 15 minute break was given halfway through the experiment.

EEG preprocessing

EEG data is noisy, picking up electrical activity from a variety of non-brain sources such as muscle activity, eye movements and the environment. There are many preprocessing procedures to increase the signal-to-noise ratio, such as filtering, resampling and artefact rejection. The plethora of methods does make it difficult to keep an overview (Chaumon et al, 2015). Preprocessing is by no means standardized, which is an issue for the reproducibility of neuroimaging research. To ensure the results are robust one needs to balance maximizing the SNR with minimizing preprocessing steps (Carlson et al., 2019).

BrainVision Analyzer 2.1 was used for preprocessing. Spline interpolation has been applied, but no channels were interpolated. The EXG5 channel was used for re-referencing all the other channels. A low cutoff at 0.5 Hz and a high cutoff at 28.8 Hz with zero phase shift Butterworth filters of order 2 were applied. The sampling rate was reduced to 64 Hz. There is no reason to presume that such a downsampling after including a low-pass filter at 28.8 Hz will affect the frequencies up until 20 Hz. The first segmentation had a segment position relative to go-stimuli reference markers of -1000.00 ms to 1500.00 ms, in which overlapping

segments were allowed. Eye blinks were removed with the Gratton and Coles method. No other artifact rejection has been done, since this may cause unnecessary data loss, after which the remaining data may still be containing artifacts. Also, multivariate pattern decoding (and specifically linear discriminant analysis) is assumedly more robust to artefacts than the established or traditional preprocessing methods used in univariate ERP/ERF research (Carlson et al., 2019). The second segmentation selected all the 1000 ms intervals preceding go-stimuli followed by an (auditory) stop-signal and separated these (for failed and successful inhibition). A fast fourier transformation using the half-spectrum with a resolution of 1 Hz and a periodic Hanning data window of 10% length (with variance correction) resulted in non-complex power values, for every second before the go-stimulus was presented, divided into the two classes. In the appendix the total BrainVision Analyzer history tree can be found.

Feature selection and algorithm sketch

Once we have (minimalistically) dealt with artifacts, a decision to make is what features of the EEG are going to be used. For modelling a cognitive system, our strategy depends on how much is already known about the data at hand. By taking more knowledge of the data into account, more constraints on the learning model are to be expected. There is a wide variety of linear (and nonlinear) methods to extract features from EEG signals, such as eigenvector methods, wavelet transformations and auto regressive methods (Vaid et al. 2015). In comparison with other methods of feature extraction in the frequency domain, one disadvantage of the fast fourier transform feature extraction methods is its relative weakness in analyzing nonstationary signals. Another weakness is that it suffers from large noise sensitivity. Nonetheless, this method still is an appropriate choice for narrowband signals and it has an enhanced speed of calculation over virtually all other methods in real-time applications (Al-Fahoum et al., 2014). The usage of very rigid and rough bandwidths substantially limits the capabilities when analyzing a complex signal. To improve accuracy, Buettner et al. (2020) present the idea to unfold the EEG standard bandwidths (0-100 Hz) in a more fine-graded equidistant 99-point spectrum. Therefore, the width of a frequency band (or ‘component’) is taken to be 1 Hz. With our hypotheses in mind, only the frequency range of 0-20 Hz is used. The combination of the components and (64) electrodes thus yields 1280 power values, for each trial.

We attempt to express one categorical dependent variable as a linear combination of continuous independent variables. Linear discriminant analysis is a method used to find a linear combination of features that characterizes or separates labeled data in two or more classes (see figure 1). This resulting combination can be used as a (linear) classifier as well as for dimensionality reduction before other successive classification methods. There are several comparable techniques which differ in requirements on the sample, that fall under the name *linear discriminant analysis*. We will perform “leave-one-out” Fisher’s linear discriminant analysis, for which it is not necessarily assumed that the classes are normally distributed nor to have equal covariances (the “homoscedasticity” assumption) - which would have to be assumed for general linear discriminant analysis. When these assumptions are satisfied, our outcome will be the same as the generalized linear discriminant analysis.

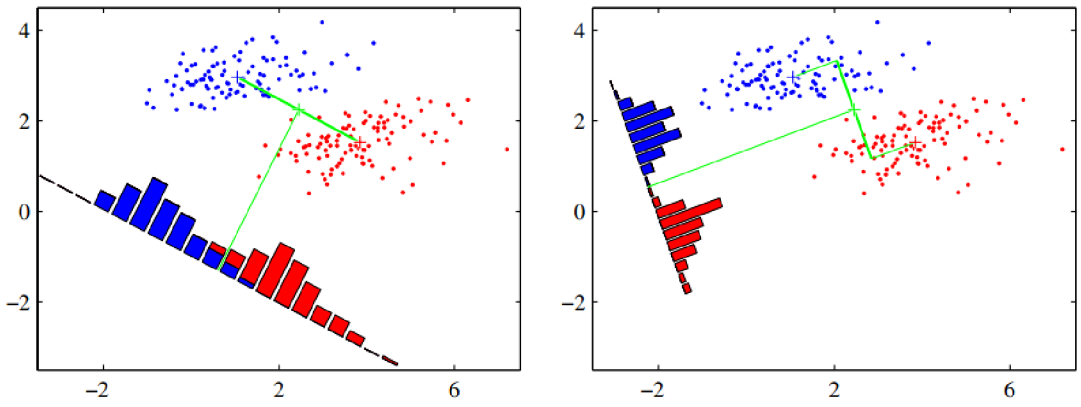


Figure 1: illustration of Fisher's linear discriminant analysis (Bishop, 2006)

The high-dimensional data will be mapped along the axis of the projection vector *w* onto the one-dimensional line orthogonal to *w*, where classification can be carried out. To find *w,* we first sum all the power values of the trials X_i per class and divide by the number of trials n_1, n_2 , which results in two (vertical) vectors of means per electrode:

$$Mf = (\frac{1}{n_1}) \sum_{X_i \in C_f} X_i \text{ and } Ms = (\frac{1}{n_2}) \sum_{X_i \in C_s} X_i \quad \text{(equation 1)}$$

Using w , these means can also be projected (by in-product) to one-dimensional values:

$$mf = w^T \cdot Mf, \quad ms = w^T \cdot Ms \quad \text{(eq. 2)}$$

This suggests that we want to find a w that maximizes the distance $(mf - ms)^2$, called the between-group variance. We can enlarge this expression arbitrarily by enlarging w , which does not yield the optimal separation. As can be seen in figure 1, the groups are best separated when each class has a minimum within-group variance while the between-group variance is maximized. The within-class variance is defined as:

$$s_k^2 = w^T \cdot (\sum_{X_i \in C_k} (Xi - Mk)(X_i - Mk)^T) \cdot w = \sum_{X_i \in C_k} (w^T \cdot X_i - mk)^2 \quad \text{(eq. 3).}$$

Fisher proposed a function that minimizes the overlap of the projected classes by having the largest separation between the projected class means while at the same time having the smallest projected within-class variances at its maximum. Thus, he defined this criterion for the separation between the two distributions as the ratio of the variance between the classes to the variance within the classes, which is formulated as:

$$J(w) = (mf - ms)^2 / (s_f^2 + s_s^2) \quad \text{(eq. 4).}$$

So, how we maximize $J(w)$? Here, the numerator can be rewritten to

$$(mf - ms)^2 = (w^T Mf - w^T Ms)^2 = w^T (Mf - Ms)(Mf - Ms)^T w = w^T S_B w \quad \text{(eq. 5).}$$

This *n* by *n* matrix S_B is called the *between-class covariance* matrix.

For the denominator we have:

$$s_f^2 + s_s^2 = w^T(\sum_{X_i \in C_f}(X_i - Mf)(X_i - Mf)^T + \sum_{X_i \in C_s}(X_i - Ms)(X_i - Ms)^T)\, w = w^T S_W\, w \quad \text{(eq. 6)},$$

where S_W is the *within-class covariance* matrix. This results in the following equation:

$$J(w) = w^T S_B w \,/\, w^T S_w w \quad \text{(eq. 7)}.$$

The ratio of quadratic forms can be maximized by differentiating and equating it to 0:

$$(2S_B w / w^T S_W w) - (w^T S_B w / (w^T S_W w)^2) 2 S_W w = 0 \quad \text{(eq. 8)}.$$

Since $w^T S_W w$ and $S_B w / (w^T S_W w)^2$ are scalars, we can write $S_W w = \lambda S_B w$ for some constant λ. Hence we are dealing with an eigenvalue problem: $w = \lambda S_W^{-1} S_B w$. Furthermore:

$$S_B w = (Mf - Ms)(Mf - Ms)^T w = (Mf - Ms) * k \quad \text{(eq. 9)}.$$

Since we are only interested in the direction of w, we leave both the scalers λ, k out to find:

$$w = S_W^{-1}(Mf - Ms) \quad \text{(eq. 10)},$$

which is the direction that maximizes the separation between the projections of the two classes of data. This proves how the optimal projection vector w is found by subtracting the vector of means (the mean per electrode of the successful trials) from the other (the mean per electrode of the failed trials) and then taking the dot product with the inverse of the within-class scatter matrix. Lastly, the decision criterion c is found by projecting the average of the two vectors of means using *w*:

$$c = w^T \cdot ((Mf - Ms) / 2) \quad \text{(eq. 11)}.$$

In summary, given the power values of the fourier-transformed periodograms of the successful (S) and failed (F) of the second preceding the go-stimulus followed by a stop-stimulus, the following steps are taken for each participant separately:

- Leave one trial out
- Compute the mean and within-class covariance matrix across the remaining trials for all features (e.g. 64 sensors), per frequency (e.g. 0-20 Hz), separately per class
- Subtract the vector of means (Ms) from the other (Mf) and take the dot product with the inverse of the within-class scatter matrix, which yields the weight vector *w*
- Now project *w* on the average of the 2 vectors of means to find *c*
- Then **classify**: if *w* projected on left-out feature vector is
  - larger than *c* → classify as successful
  - smaller than *c* → classify as failed
- Repeat for all the other trials
- Count correct classifications for all trials, per frequency
- Finally, conduct statistical evaluation

With smaller sample sizes, the chance increases that the performance level is high by chance. This is true for LDA, naive-Bayes and support vector machines alike, regardless of the type of cross-validation performed (Combrisson & Jerbi, 2015). Their solution is to evaluate machine learning results for statistical significance against sample-size specific thresholds instead of theoretical chance performance. They provided a "look-up table" revealing the minimum performance that is needed to significantly exceed chance (figure 2).

c / p / n	2-classes			
	p<0,05	*p<0,01*	*p<10-3*	*p<10-4*
20	70,0%	75,0%	85,0%	90,0%
40	62,5%	67,5%	75,0%	77,5%
60	60,0%	65,0%	70,0%	73,3%
80	58,7%	62,5%	67,5%	70,0%
100	58,0%	62,0%	65,0%	68,0%
200	56,0%	58,0%	61,0%	63,0%
300	54,7%	56,7%	59,0%	60,7%
400	54,0%	55,7%	57,7%	59,2%
500	53,6%	55,2%	57,0%	58,2%

Figure 2: minimum accuracies

Results

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	SUM
1	0,51	0,42	0,43	0,53	0,49	0,51	0,48	0,52	0,51	0,49	0,42	0,56	0,39	0,54	0,53	0,49	0,39	0,53	0,53	0,46	9,73
2	0,49	0,46	0,58	0,46	0,5	0,59	0,64	0,44	0,45	0,48	0,56	0,48	0,54	0,52	0,5	0,51	0,53	0,54	0,57	0,53	10,37
3	0,49	0,44	0,48	0,54	0,41	0,48	0,54	0,62	0,47	0,49	0,55	0,54	0,44	0,51	0,49	0,53	0,52	0,54	0,47	0,44	9,99
4	0,5	0,51	0,55	0,54	0,5	0,57	0,49	0,52	0,48	0,5	0,42	0,4	0,45	0,5	0,5	0,39	0,49	0,47	0,51	0,51	9,8
5	0,6	0,45	0,55	0,54	0,51	0,52	0,5	0,52	0,54	0,44	0,52	0,5	0,47	0,53	0,42	0,49	0,55	0,55	0,51	0,44	10,15
9	0,5	0,38	0,58	0,46	0,44	0,46	0,55	0,56	0,54	0,59	0,41	0,54	0,52	0,46	0,64	0,45	0,46	0,47	0,51	0,46	9,98
10	0,57	0,51	0,54	0,48	0,49	0,5	0,48	0,51	0,58	0,56	0,47	0,47	0,46	0,49	0,49	0,52	0,52	0,54	0,54	0,46	10,18
11	0,53	0,5	0,42	0,44	0,48	0,5	0,48	0,52	0,45	0,5	0,52	0,36	0,41	0,47	0,52	0,59	0,41	0,52	0,59	0,5	9,71
12	0,47	0,54	0,49	0,44	0,42	0,52	0,49	0,6	0,48	0,47	0,53	0,43	0,46	0,51	0,58	0,47	0,5	0,46	0,52	0,54	9,92
13	0,57	0,42	0,56	0,46	0,49	0,57	0,42	0,53	0,56	0,47	0,49	0,45	0,59	0,56	0,54	0,52	0,51	0,53	0,51	0,55	10,3
14	0,52	0,48	0,51	0,47	0,44	0,51	0,55	0,53	0,48	0,51	0,53	0,53	0,53	0,51	0,48	0,45	0,46	0,52	0,51	0,56	10,08
15	0,61	0,54	0,49	0,54	0,44	0,53	0,57	0,62	0,51	0,53	0,59	0,55	0,48	0,58	0,56	0,44	0,53	0,55	0,51	0,49	10,66
16	0,43	0,53	0,49	0,51	0,49	0,53	0,48	0,47	0,51	0,46	0,53	0,54	0,55	0,53	0,43	0,49	0,46	0,5	0,5	0,49	9,92
17	0,53	0,47	0,53	0,48	0,59	0,55	0,54	0,51	0,55	0,47	0,45	0,47	0,47	0,53	0,49	0,55	0,49	0,54	0,52	0,57	10,3
18	0,52	0,49	0,51	0,44	0,51	0,53	0,52	0,52	0,52	0,43	0,61	0,62	0,51	0,46	0,47	0,57	0,43	0,51	0,53	0,6	10,3
19	0,56	0,62	0,53	0,53	0,4	0,6	0,44	0,53	0,62	0,64	0,58	0,51	0,53	0,58	0,47	0,43	0,54	0,47	0,5	0,54	10,62
20	0,59	0,48	0,54	0,61	0,58	0,56	0,52	0,57	0,53	0,49	0,51	0,53	0,49	0,58	0,56	0,59	0,52	0,49	0,62	0,53	10,89
21	0,46	0,45	0,47	0,58	0,57	0,53	0,41	0,45	0,57	0,64	0,57	0,51	0,53	0,51	0,53	0,62	0,5	0,54	0,57	0,47	10,48
22	0,44	0,41	0,51	0,44	0,43	0,45	0,39	0,51	0,48	0,48	0,61	0,56	0,46	0,54	0,52	0,57	0,46	0,57	0,5	0,51	9,84
101	0,56	0,49	0,46	0,44	0,5	0,43	0,49	0,48	0,43	0,56	0,45	0,61	0,44	0,54	0,5	0,53	0,5	0,47	0,56	0,53	9,97
103	0,48	0,5	0,51	0,55	0,47	0,49	0,55	0,44	0,48	0,52	0,51	0,52	0,44	0,46	0,51	0,51	0,49	0,46	0,5	0,53	9,92
104	0,5	0,42	0,58	0,25	0,5	0,33	0,42	0,42	0,5	0,5	0,5	0,75	0,58	0,5	0,67	0,42	0,25	0,42	0,25	0,5	9,26
106	0,53	0,53	0,56	0,52	0,54	0,54	0,48	0,53	0,54	0,53	0,52	0,54	0,47	0,57	0,56	0,52	0,49	0,54	0,5	0,55	10,56
107	0,51	0,52	0,52	0,46	0,55	0,56	0,57	0,42	0,56	0,49	0,51	0,49	0,55	0,56	0,51	0,51	0,48	0,47	0,56	0,51	10,31
108	0,56	0,44	0,46	0,44	0,46	0,49	0,45	0,53	0,43	0,44	0,53	0,56	0,5	0,53	0,53	0,53	0,49	0,52	0,53	0,56	9,98
109	0,5	0,48	0,6	0,51	0,49	0,5	0,51	0,47	0,5	0,44	0,5	0,54	0,46	0,52	0,47	0,46	0,48	0,5	0,52	0,54	9,99
110	0,54	0,54	0,51	0,44	0,58	0,42	0,51	0,53	0,49	0,47	0,49	0,6	0,55	0,52	0,62	0,48	0,56	0,51	0,48	0,51	10,35
111	0,47	0,5	0,54	0,51	0,5	0,51	0,52	0,54	0,47	0,52	0,45	0,47	0,46	0,53	0,46	0,51	0,52	0,53	0,57	0,5	10,08
112	0,48	0,52	0,47	0,55	0,46	0,46	0,5	0,55	0,42	0,54	0,47	0,6	0,6	0,46	0,52	0,51	0,49	0,47	0,46	0,5	10,03
SUM	15,02	14,04	14,97	14,16	14,23	14,74	14,49	14,96	14,65	14,65	14,8	15,23	14,33	15,1	15,07	14,65	14,02	14,73	14,95	14,88	

Figure 3: averaged accuracies of the 5-times repeated class-balancing LOO-procedure

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0,51	0,42	0,43	0,53	0,49	0,51	0,48	0,52	0,51	0,49	0,42	0,56	0,39	0,54	0,53	0,49	0,39	0,53	0,53	0,46
2	0,49	0,46	0,58	0,46	0,5	0,59	0,64	0,44	0,45	0,48	0,56	0,48	0,54	0,52	0,5	0,51	0,53	0,54	0,57	0,53
3	0,49	0,44	0,48	0,54	0,41	0,48	0,54	0,62	0,47	0,49	0,55	0,54	0,44	0,51	0,49	0,53	0,52	0,54	0,47	0,44
4	0,5	0,51	0,55	0,54	0,5	0,57	0,49	0,52	0,48	0,5	0,42	0,4	0,45	0,5	0,5	0,39	0,49	0,47	0,51	0,51
5	0,6	0,45	0,55	0,54	0,51	0,52	0,5	0,52	0,54	0,44	0,52	0,5	0,47	0,53	0,42	0,49	0,55	0,55	0,51	0,44
9	0,5	0,38	0,58	0,46	0,44	0,46	0,55	0,56	0,54	0,59	0,41	0,54	0,52	0,46	0,64	0,45	0,46	0,47	0,51	0,46
10	0,57	0,51	0,54	0,48	0,49	0,5	0,48	0,51	0,58	0,56	0,47	0,47	0,46	0,49	0,49	0,52	0,52	0,54	0,54	0,46
11	0,53	0,5	0,42	0,44	0,48	0,5	0,48	0,52	0,45	0,5	0,52	0,36	0,41	0,47	0,52	0,59	0,41	0,52	0,59	0,5
12	0,47	0,54	0,49	0,44	0,42	0,52	0,49	0,6	0,48	0,47	0,53	0,43	0,46	0,51	0,58	0,47	0,5	0,46	0,52	0,54
13	0,57	0,42	0,56	0,46	0,49	0,57	0,42	0,53	0,56	0,47	0,49	0,45	0,59	0,56	0,54	0,52	0,51	0,53	0,51	0,55
14	0,52	0,48	0,51	0,47	0,44	0,51	0,55	0,53	0,48	0,51	0,53	0,53	0,53	0,51	0,48	0,45	0,46	0,52	0,51	0,56
15	0,61	0,54	0,49	0,54	0,44	0,53	0,57	0,62	0,51	0,53	0,59	0,55	0,48	0,58	0,56	0,44	0,53	0,55	0,51	0,49
16	0,43	0,53	0,49	0,51	0,49	0,53	0,48	0,47	0,51	0,46	0,53	0,54	0,55	0,53	0,43	0,49	0,46	0,5	0,5	0,49
17	0,53	0,47	0,53	0,48	0,59	0,55	0,54	0,51	0,55	0,47	0,45	0,47	0,47	0,53	0,49	0,55	0,49	0,54	0,52	0,57
18	0,52	0,49	0,51	0,44	0,51	0,53	0,52	0,52	0,52	0,43	0,61	0,62	0,51	0,46	0,47	0,57	0,43	0,51	0,53	0,6
19	0,56	0,62	0,53	0,53	0,4	0,6	0,44	0,53	0,62	0,64	0,58	0,51	0,53	0,58	0,47	0,43	0,54	0,47	0,5	0,54
20	0,59	0,48	0,54	0,61	0,58	0,56	0,52	0,57	0,53	0,49	0,51	0,53	0,49	0,58	0,56	0,59	0,52	0,49	0,62	0,53
21	0,46	0,45	0,47	0,58	0,57	0,53	0,41	0,45	0,57	0,64	0,57	0,51	0,53	0,51	0,53	0,62	0,5	0,54	0,57	0,47
22	0,44	0,41	0,51	0,44	0,43	0,45	0,39	0,51	0,48	0,48	0,61	0,56	0,46	0,54	0,52	0,57	0,46	0,57	0,5	0,51
101	0,56	0,49	0,46	0,44	0,5	0,43	0,49	0,48	0,43	0,56	0,45	0,61	0,44	0,54	0,5	0,53	0,5	0,47	0,56	0,53
103	0,48	0,5	0,51	0,55	0,47	0,49	0,55	0,44	0,48	0,52	0,51	0,52	0,44	0,46	0,51	0,51	0,49	0,46	0,5	0,53
104	0,5	0,42	0,58	0,25	0,5	0,33	0,42	0,42	0,5	0,5	0,5	0,75	0,58	0,5	0,67	0,42	0,25	0,42	0,25	0,5
106	0,53	0,53	0,56	0,52	0,54	0,54	0,48	0,53	0,54	0,53	0,52	0,54	0,47	0,57	0,56	0,52	0,49	0,54	0,5	0,55
107	0,51	0,52	0,52	0,46	0,55	0,56	0,57	0,42	0,56	0,49	0,51	0,49	0,55	0,56	0,51	0,51	0,48	0,47	0,56	0,51
108	0,56	0,44	0,46	0,44	0,46	0,49	0,45	0,53	0,43	0,44	0,53	0,56	0,5	0,53	0,53	0,53	0,49	0,52	0,53	0,56
109	0,5	0,48	0,6	0,51	0,49	0,5	0,51	0,47	0,5	0,44	0,5	0,54	0,46	0,52	0,47	0,46	0,48	0,5	0,52	0,54
110	0,54	0,54	0,51	0,44	0,58	0,42	0,51	0,53	0,49	0,47	0,49	0,6	0,55	0,52	0,62	0,48	0,56	0,51	0,48	0,51
111	0,47	0,5	0,54	0,51	0,5	0,51	0,52	0,54	0,47	0,52	0,45	0,47	0,46	0,53	0,46	0,51	0,52	0,53	0,57	0,5
112	0,48	0,52	0,47	0,55	0,46	0,46	0,5	0,55	0,42	0,54	0,47	0,6	0,6	0,46	0,52	0,51	0,49	0,47	0,46	0,5

Figure 4: (in)significant combinations of averaged-5-times-balanced procedure, p<0,05

The described procedure is repeated five times. Each time we randomly select a subset of the majority class to balance the sets of the trials. The balanced amounts of trials are listed in the third row of table 1. Hereafter, the resulting accuracies were averaged and are depicted in figure 3. A customary statistical test for significance is added in the appendix.

participant	1	2	3	4	5	9	10	11	12	13	14	15	16	17	18	19	20	21	22	101	103	104	106	107	108	109	110	111	112
n.o. trials	192	192	192	192	192	192	192	192	192	192	192	192	192	192	191	192	192	192	192	257	256	256	257	288	256	288	256	256	289
balanced	178	108	180	152	154	162	182	64	172	170	192	176	188	148	178	156	156	184	162	228	224	12	210	280	254	214	186	254	212
majority class	s	f	f	f	f	f	f	f	s	s		s	s	s	f	s	s	f	f	f	s	f	s	f	f	s	s	f	s
percentage	54	72	53	60	60	58	53	83	55	56	50	54	51	61	53	59	59	52	58	56	56	98	59	51	50	63	64	50	63

Table 1: participant, number of trials, n.o. trials when balanced, majority class and its percentage

There are multiple participants for which significant classification has been possible, in line with the given sample-size specific thresholds (figure 4). The light-green scores are almost-significant accuracies for participants who did not have a significant classification.

The procedure can also be applied without balancing the classes (figure 5). To clarify the problems of imbalanced classes we can make use of different metrics. Let *positives* denote trials that were originally labeled as successful inhibitions and *negatives* as failed ones and write TruePositives, TrueNegatives, FalsePositives (type I errors) and FalseNegatives (type II) accordingly. The metrics *precision* TP/(TP+FP), *recall/sensitivity* TP/(TP+FN), *specificity* TN/(TN+FP), *negative predictive value* TN/(TN+FN) and *F1-score* 2TP/(2*TP+FP+FN) may all give different insights, for example revealing how the model was not capable of classifying successful inhibition because it trained almost entirely on classifying the failed inhibitions.

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
1	0,53	0,48	0,42	0,53	0,44	0,53	0,42	0,53	0,51	0,58	0,43	0,55	0,41	0,55	0,51	0,49	0,42	0,54	0,53	0,45
2	0,58	0,47	0,64	0,49	0,53	0,56	0,65	0,52	0,53	0,56	0,57	0,49	0,64	0,64	0,59	0,48	0,48	0,66	0,53	0,53
3	0,49	0,46	0,5	0,56	0,4	0,47	0,5	0,63	0,42	0,48	0,56	0,52	0,43	0,52	0,54	0,55	0,56	0,56	0,48	0,44
4	0,54	0,55	0,57	0,53	0,54	0,51	0,47	0,53	0,54	0,51	0,42	0,45	0,47	0,48	0,5	0,44	0,5	0,48	0,53	0,52
5	0,57	0,47	0,47	0,53	0,43	0,51	0,54	0,56	0,61	0,47	0,48	0,52	0,53	0,56	0,43	0,55	0,57	0,52	0,52	0,45
9	0,56	0,44	0,59	0,48	0,48	0,45	0,52	0,55	0,52	0,54	0,5	0,53	0,48	0,48	0,6	0,45	0,48	0,48	0,52	0,44
10	0,56	0,52	0,56	0,45	0,47	0,52	0,46	0,49	0,57	0,57	0,47	0,43	0,47	0,49	0,51	0,51	0,52	0,54	0,46	0,45
11	0,53	0,61	0,69	0,59	0,64	0,71	0,68	0,65	0,6	0,62	0,63	0,62	0,64	0,66	0,68	0,58	0,55	0,58	0,59	0,53
12	0,43	0,47	0,51	0,46	0,41	0,48	0,47	0,6	0,46	0,47	0,53	0,44	0,49	0,48	0,54	0,49	0,52	0,52	0,53	0,56
13	0,56	0,42	0,56	0,52	0,47	0,56	0,44	0,56	0,54	0,47	0,47	0,53	0,56	0,55	0,47	0,52	0,5	0,5	0,54	0,49
14	0,52	0,48	0,51	0,47	0,44	0,51	0,55	0,53	0,48	0,51	0,53	0,53	0,53	0,51	0,48	0,45	0,46	0,52	0,51	0,56
15	0,58	0,52	0,53	0,52	0,46	0,56	0,49	0,61	0,53	0,55	0,58	0,55	0,46	0,61	0,51	0,46	0,48	0,51	0,49	0,48
16	0,44	0,51	0,52	0,52	0,49	0,53	0,49	0,48	0,47	0,48	0,55	0,54	0,58	0,54	0,47	0,49	0,42	0,54	0,52	0,47
17	0,52	0,5	0,64	0,55	0,59	0,57	0,56	0,51	0,55	0,49	0,48	0,4	0,51	0,51	0,49	0,53	0,51	0,5	0,55	0,49
18	0,48	0,46	0,52	0,45	0,51	0,53	0,5	0,52	0,44	0,45	0,63	0,61	0,47	0,46	0,47	0,59	0,47	0,47	0,56	0,58
19	0,54	0,52	0,56	0,51	0,48	0,56	0,49	0,51	0,51	0,59	0,58	0,4	0,52	0,55	0,54	0,48	0,61	0,53	0,42	0,5
20	0,55	0,52	0,53	0,57	0,55	0,53	0,55	0,59	0,61	0,45	0,51	0,55	0,51	0,58	0,56	0,58	0,5	0,42	0,62	0,6
21	0,46	0,4	0,49	0,59	0,54	0,49	0,42	0,41	0,56	0,65	0,54	0,53	0,49	0,53	0,53	0,59	0,51	0,55	0,55	0,48
22	0,49	0,47	0,53	0,46	0,48	0,51	0,43	0,51	0,43	0,49	0,53	0,55	0,48	0,52	0,5	0,54	0,49	0,54	0,49	0,52
101	0,57	0,48	0,4	0,48	0,51	0,44	0,5	0,49	0,41	0,54	0,47	0,56	0,48	0,54	0,49	0,49	0,49	0,47	0,54	0,53
103	0,52	0,52	0,54	0,51	0,52	0,5	0,54	0,48	0,48	0,55	0,51	0,46	0,43	0,43	0,54	0,53	0,5	0,52	0,5	0,53
104	0,86	0,88	0,81	0,79	0,8	0,82	0,91	0,86	0,77	0,8	0,79	0,77	0,86	0,87	0,89	0,84	0,89	0,86	0,86	0,8
106	0,54	0,47	0,56	0,47	0,54	0,59	0,55	0,59	0,49	0,56	0,49	0,53	0,42	0,59	0,52	0,56	0,45	0,53	0,55	0,53
107	0,53	0,5	0,51	0,48	0,55	0,55	0,53	0,44	0,54	0,51	0,52	0,47	0,55	0,54	0,52	0,51	0,47	0,46	0,54	0,52
108	0,56	0,45	0,46	0,45	0,46	0,47	0,45	0,54	0,42	0,45	0,54	0,55	0,5	0,54	0,51	0,53	0,5	0,51	0,52	0,55
109	0,56	0,49	0,58	0,52	0,58	0,54	0,44	0,53	0,5	0,39	0,45	0,56	0,53	0,5	0,4	0,45	0,54	0,52	0,49	0,58
110	0,45	0,55	0,46	0,49	0,52	0,45	0,49	0,48	0,52	0,52	0,54	0,54	0,54	0,46	0,52	0,48	0,52	0,57	0,48	0,45
111	0,47	0,5	0,54	0,5	0,5	0,5	0,51	0,54	0,48	0,51	0,45	0,47	0,47	0,52	0,48	0,52	0,52	0,53	0,57	0,49
112	0,51	0,55	0,5	0,48	0,48	0,56	0,46	0,57	0,55	0,53	0,58	0,55	0,61	0,46	0,51	0,51	0,51	0,52	0,51	0,52

Figure 5: accuracies for the Leave-One-Out procedure (no classes are balanced)

Thus, class imbalance may compromise the process of learning when the model will learn to correctly classify the prevalent class while ignoring the less occurring events. In other words, a model can be accurate while still not being valuable at all. Moreover, *Matthews Correlation Coefficient* (TP*TN-FP*FP) / sqrt((TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)) is an acclaimed metric for when positive and negative cases are of equal importance, to summarize the performance in a single value (Chicco & Jurman, 2020). It depends on the applicative goal which of the metrics deserves preference. For the class-balanced procedure, the outcomes of all the mentioned metrics for the 4-12 Hz range are listed in the appendix.

Our solution for dealing with this skewness was thus to equalize the number of trials used. After the repeated class-balancing procedure, we can determine which electrodes were most discriminative by having received the highest weights. Including irrelevant predictors can worsen performance on new data, because one would end up fitting to noise. To find more specific explanations and to improve accuracy scores at the same time, the feature space can be limited while trying to optimize for a subset of electrodes. The resulting weight-vectors are averaged, for every participant and frequency separately. Note that an alternative strategy would be to start with taking the absolute values of the resulting weights vectors *before* averaging them, which has not been done. Next, when taking the absolute values of the averaged weight-vectors, the weights in these vectors are ordered. The corresponding electrodes can be counted to find an ordering of occurence. For a first indication of importance, we sum the top-15 most important electrodes for the theta and alpha band and all participants together (15*8*29 electrodes) and take the top-20 of that ordering, showing how the parietal area is of strong influence in our calculations (figure 6).

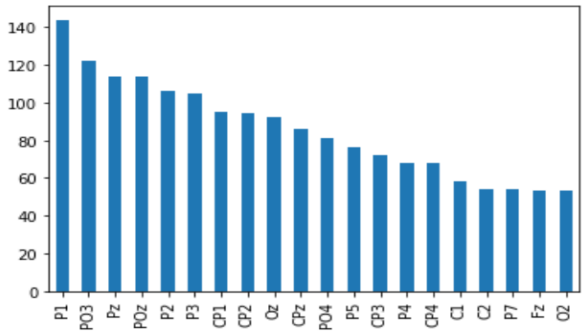


Figure 6: cross-subject top-20 of the top-15 discriminating electrodes in theta + alpha band

Using a specific subset of electrodes for every 1 Hz frequency interval and every participant (29*20 = 580 subsets) would result in the highest accuracy scores. Since there are 2^64 possible subsets, a more (computationally) feasible strategy is to find one subset of electrodes per frequency, across participants. This has been done by putting the top-32 most informative electrodes of the *averaged 5-times repeated class-balancing LOO-procedure* for every participant (for one frequency at a time) in a list. Then, the occurrences were counted and ordered based on this summation. Next, we execute the same procedure with an increasing number of electrodes from that ordered list, starting with the single most occuring electrode and finishing with all. After each time of running the procedure with one set of electrodes, the accuracies are summed. The subset of electrodes corresponding to the highest cross-participant sum is returned. The 8 resulting subsets can be found in the appendix. This optimization yields the following accuracy-scores for the 4-12 Hz range:

	4	5	6	7	8	9	10	11
001	0,57	0,47	0,56	0,54	0,51	0,52	0,52	0,55
002	0,48	0,57	0,48	0,46	0,45	0,52	0,48	0,53
003	0,52	0,53	0,52	0,61	0,47	0,53	0,53	0,54
004	0,57	0,59	0,58	0,54	0,48	0,55	0,5	0,51
005	0,49	0,55	0,51	0,52	0,54	0,49	0,56	0,55
009	0,49	0,59	0,54	0,56	0,54	0,56	0,52	0,59
010	0,5	0,45	0,53	0,52	0,58	0,54	0,6	0,53
011	0,56	0,58	0,53	0,84	0,55	0,47	0,58	0,42
012	0,52	0,46	0,52	0,62	0,48	0,52	0,49	0,52
013	0,54	0,46	0,54	0,53	0,56	0,5	0,45	0,45
014	0,5	0,43	0,55	0,54	0,48	0,55	0,47	0,48
015	0,44	0,56	0,46	0,56	0,51	0,48	0,56	0,6
016	0,56	0,48	0,51	0,46	0,51	0,51	0,57	0,53
017	0,62	0,53	0,49	0,49	0,55	0,54	0,5	0,56
018	0,49	0,55	0,48	0,49	0,52	0,49	0,55	0,57
019	0,44	0,55	0,56	0,54	0,62	0,52	0,49	0,53
020	0,49	0,51	0,58	0,56	0,53	0,56	0,54	0,62
021	0,51	0,46	0,49	0,47	0,57	0,49	0,58	0,55
022	0,48	0,51	0,51	0,53	0,48	0,51	0,57	0,62
101	0,53	0,5	0,52	0,5	0,43	0,51	0,49	0,59
103	0,54	0,54	0,53	0,47	0,48	0,5	0,5	0,49
104	0,75	0,58	0,58	0,67	0,5	0,58	0,58	0,83
106	0,55	0,63	0,49	0,56	0,54	0,47	0,52	0,5
107	0,49	0,49	0,47	0,41	0,56	0,54	0,51	0,5
108	0,52	0,56	0,53	0,53	0,43	0,48	0,5	0,54
109	0,5	0,51	0,57	0,48	0,5	0,54	0,54	0,5
110	0,55	0,53	0,59	0,56	0,49	0,49	0,56	0,52
111	0,45	0,52	0,53	0,54	0,47	0,51	0,49	0,49
112	0,55	0,52	0,55	0,53	0,42	0,52	0,56	0,54
SUM	15,2	15,2	15,3	15,6	14,8	15	15,3	15,8

Figure 7: 4-12 Hz accuracies, using semi-optimized subsets of electrodes per frequency

For testing the hypothesis of frontal theta yielding predictability, the *averaged 5-times repeated class-balancing LOO-procedure* was executed with the (fixed) subset of electrodes [Fp1, AF7, AF3, F1, F3, F5, F7, FT7, FC5, FC3, FC1, Fpz, Fp2, AF8, AF4, AFz, Fz, F2, F4, F6, F8, FCz, FC2, FC4, FC6, FT8], which resulted in the following accuracies:

	4	5	6	7
001	0,43	0,49	0,49	0,49
002	0,45	0,44	0,48	0,43
003	0,43	0,44	0,56	0,56
004	0,54	0,5	0,52	0,47
005	0,44	0,56	0,57	0,52
009	0,54	0,41	0,57	0,6
010	0,52	0,51	0,48	0,48
011	0,53	0,5	0,48	0,47
012	0,44	0,51	0,57	0,5
013	0,41	0,56	0,48	0,54
014	0,53	0,51	0,6	0,53
015	0,48	0,49	0,49	0,56
016	0,48	0,49	0,5	0,5
017	0,65	0,49	0,46	0,58
018	0,51	0,47	0,49	0,46
019	0,49	0,56	0,51	0,48
020	0,56	0,58	0,58	0,51
021	0,56	0,57	0,42	0,51
022	0,48	0,49	0,49	0,45
101	0,53	0,52	0,47	0,44
103	0,45	0,5	0,52	0,47
104	0,17	0,42	0,42	0,42
106	0,48	0,49	0,49	0,56
107	0,5	0,55	0,53	0,46
108	0,51	0,53	0,46	0,53
109	0,54	0,54	0,51	0,47
110	0,49	0,39	0,53	0,54
111	0,48	0,44	0,53	0,46
112	0,5	0,49	0,51	0,56
SUM	14,1	14,4	14,7	14,6

Figure 8: frontal electrodes on 4-8 Hz

Discussion

The question ‘what are the predictors of performance failures from a perspective that goes beyond theta and alpha in terms of frequency and location?’ can now be dealt with. Based on the cross-participants summation of the 5 times repeated class-balanced procedure using all electrodes, the 11 to 12 Hz frequency interval turns out to yield the most predictability (sum = 15,23), followed by 13-14 Hz (sum = 15,1) and 14-15 Hz (sum = 15,07). With this procedure, significant accuracies are found for 21 of the 29 participants. Also, with the use of a semi-optimized unique subset of the electrodes per frequency interval (on the 4-12 Hz range), the same frequency interval (11 to 12 Hz) has resulted in the highest predictability. Here, the decision to start from the top-32 most informative electrodes is not well-founded. Since the ordering of these tops already matter, there are other subsets of electrodes that are likely to give more legitimate representations of importance. Better explanations can be found when starting with the top-1 of cross-participant most-informative electrodes and executing the procedure with an increasing number of electrodes from that top - which is a very intensive calculation. Upon this, one more possibility is to optimize for finding a subset of electrodes per participant instead of for a cross-participant 1 Hz frequency interval.

Another aspect to consider is that one of the main prevailing challenges is the low signal-to-noise ratio. Our ratio is presumably imperfect because of technical glitches and other artifacts. Blind-source-separation techniques such as Independent - and Principal Component Analysis could have been deployed to decompose the signal, to improve the signal to noise ratio by removing certain artifactual components. Because all rejection would require some level of supervision, Chaumon et al. (2015) recommend a semi-automatic approach, in which (informed) decisions can be made about which components to reject, based on a number of statistical measures. Urigüen and Garcia-Zapirain (2015) argue that the best method is to be found in combining more than one algorithm and removing one type of artifact at a time. Also, Pion-Tonachini et al., (2019) made available an automated independent component classifier method. Jiang et. al (2019) reviewed (a selection of) contemporary practices and concluded that dealing with artifacts continues to be an open problem. In addition to this view, Grootswagers et al. (2017) suggest that even the sole removal of trials with eye-blink artefacts is a preprocessing step that potentially could be left out from their decoding analysis preprocessing pipeline, because classifiers (such as ours) possibly have the capacity to learn to suppress noise during training (when neither over- nor underfitting). Still, these artefacts can be potential source of discrimination when confounding with a condition. If this were to happen, it may be hard to determine whether the classifier was actually decoding the experimental condition or just the correlating difference in artefacts.

Of course, a good predictor is not necessarily a true cause. To go beyond correlation, intervention (for example with transcranial magnetic stimulation) would have to take place, since no causality has yet been shown. By investigating how much change in one variable affects the outcome given all other feature values remain fixed, causal inference is possible as well. For future research it would be also interesting to use different time windows on the dataset. Although both alpha and theta oscillations are associated with the allocation of attention to sensory stimuli, these oscillations may operate through very different mechanisms (Keller et al. 2017). Ouyang et al. (2015) reason that there are three types of events occurring in a typical psychological EEG experiment: events that are strictly time-locked to the stimulus, events that are strictly time-locked to the response, and events that occur at variable latencies between the two others. Therefore, implementing the model as a continuous application would be of help discovering more fine-grained interplay. The traditional stop-signal task itself (and the used dataset) may still be highly valuable to provide insights in both ‘higher’ (e.g., proactive top-down) and lower-level (reactive, more bottom-up driven) mechanisms of inhibition and motor inhibition, and be a useful tool for understanding pathologies as well as inspiring new treatments (Kenemans, 2015).

Decisions made at both preprocessing (e.g. dimensionality reduction, subsampling, trial averaging) and decoding (e.g. classifier selection, cross-validation design) stages of the analysis can significantly affect the results (Grootswagers et al., 2017). For future research, the current procedure can be expanded by further developing and optimizing the subset-selection of the most informative electrodes. Still, there is the chance of missing out on important information when we leave out certain features. Clearly, it is to be assumed that there are many more patterns left to unravel in this data. Finding rare events is a type of prediction in itself and the scarcest ones are the hardest to find (Haixiang et al., 2017).

Although in EEG analysis linear methods are often used, the usage of nonlinear approaches has increased in their presence because they reveal aspects that cannot be measured with linear approaches (Rodriguez-Bermudez & García-Laencina, 2015). Beyond its usage as a demonstrably powerful classifier, fisher's linear discriminant analysis method can also be used as a preprocessing tool for dimensionality reduction, to first determine which features are most important and thereafter feed those to more refined algorithms. Within this wide variety of (deep) neural network architectures, classification has already been applied to many EEG tasks (Lotte et al., 2018); Craik et al., 2019; Roy et al., 2019). Herein, further research to compare how these (ideally hybrid) networks interpret raw versus denoised EEG is encouraged as well. There are methods to use more state-of-the-art (deep learning) algorithms while at the same time warranting interpretability (Ribeiro et al., 2016; Guidotti et al., 2018). Nonetheless, such classifiers may still be considered to be less transparent. Models should also be interpretable, even when the actual task of interpretation appears to be underspecified (Lipton, 2018). Since the definition of what constitutes a viable explanation is unclear, even strong regulations such as 'right to explanation' can be undermined with less-than-satisfactory explanations (Rudin, 2019). Because prediction is not the main goal of decoding in neuroscience and the choice of classifier favours simplicity and interpretability over optimizing prediction accuracies, linear classifiers are generally preferred for brain decoding studies (Grootswagers et al., 2017).

Conclusion

The extent to which changes in brain activity can foreshadow human error remains uncertain yet has important theoretical and practical implications (O'Connell et al, 2009). The study of response inhibition has the advantage of dealing with a relatively straightforward process, namely the overriding of a planned or already initiated action (Bari & Robbins, 2013). In many circumstances, alternative courses of action and thoughts have to be inhibited to allow the emergence of goal-directed behavior. Among the many ways in which our brain controls its own activity, inhibitory processes are important both in everyday life as well as during emergency situations. Ursin (2005) noted how inhibition is important for our abilities to make choices, as well as for our freedom of choice.

It is imaginable how interventions can be devised for critical operations. New technologies might focus on lowering errors or increasing successful response inhibition, but it is necessary to question how the individual might benefit from such monitoring towards increased goal-directed action. Active monitoring may pose risks for the autonomy of the involved actors through the making of increasingly strong generalizations about ongoing (and upcoming) cognitive processes and states. There is a substantive divergence of interpretation and implementation of ethical AI principles (Jobin et al., 2019), because of which the importance of guideline development efforts persists.

Response inhibition can be predicted before it takes place. Cross-participant resemblances of predictors have been found. Our group of participants is nevertheless nowhere near a true societal representation (Henrich et al., 2010). The strength of classification can be increased with more sophisticated learning algorithms and with the use of Fisher's method as a preprocessor for dimensionality reduction.

References

Al-Fahoum, A. S., & Al-Fraihat, A. A. (2014). Methods of EEG signal features extraction using linear analysis in frequency and time-frequency domains. International Scholarly Research Notices, 2014.

Aron, A. R., Robbins, T. W., & Poldrack, R. A. (2014). Inhibition and the right inferior frontal cortex: one decade on. Trends in cognitive sciences, 18(4), 177-185.

Bari, A., & Robbins, T. W. (2013). Inhibition and impulsivity: behavioral and neural basis of response control. Progress in neurobiology, 108, 44-79.

Bekkar, M., Djemaa, H. K., & Alitouche, T. A. (2013). Evaluation measures for models assessment over imbalanced data sets. J Inf Eng Appl, 3(10).

Bishop, C. M. (2006). Pattern recognition and machine learning. springer.

Buettner, R., Rieg, T., & Frick, J. (2020). Machine Learning based Diagnosis of Diseases Using the Unfolded EEG Spectra: Towards an Intelligent Software Sensor. In Information Systems and Neuroscience (pp. 165-172). Springer, Cham.

Carlson, T. A., Grootswagers, T., & Robinson, A. K. (2019). An introduction to time-resolved decoding analysis for M/EEG. arXiv preprint arXiv:1905.04820.

Cavanagh, J. F., & Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. Trends in cognitive sciences, 18(8), 414-421.

Chaumon, M., Bishop, D. V., & Busch, N. A. (2015). A practical guide to the selection of independent components of the electroencephalogram for artifact correction. *Journal of neuroscience methods*, *250*, 47-63.

Chaumon, M., Crouzet, S. M., & Busch, N. A. (2015). Cutting-edge methods for EEG research on cognition. *Journal of neuroscience methods*, *250*, 1-2.

Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC genomics, 21(1), 1-13.

Combrisson, E., & Jerbi, K. (2015). Exceeding chance level by chance: The caveat of theoretical chance levels in brain signal classification and statistical assessment of decoding accuracy. Journal of neuroscience methods, 250, 126-136.

Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: a review. Journal of neural engineering, 16(3), 031001.

De Jong, R., Coles, M. G., & Logan, G. D. (1995). Strategies and mechanisms in nonselective and selective inhibitory motor control. Journal of experimental psychology: Human perception and performance, 21(3), 498.

Esch, L., Dinh, C., Larson, E., Engemann, D., Jas, M., Khan, S., ... & Hämäläinen, M. S. (2019). MNE: Software for Acquiring, Processing, and Visualizing MEG/EEG Data. Magnetoencephalography: From Signals to Dynamic Cortical Networks, 355-371.

Grootswagers, T., Wardle, S. G., & Carlson, T. A. (2017). Decoding dynamic brain patterns from evoked responses: A tutorial on multivariate pattern analysis applied to time series neuroimaging data. Journal of cognitive neuroscience, 29(4), 677-697.

Guidotti, R., Monreale, A., Ruggieri, S., Turini, F., Giannotti, F., & Pedreschi, D. (2018). A survey of methods for explaining black box models. ACM computing surveys (CSUR), 51(5), 1-42.

Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., & Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. Expert Systems with Applications, 73, 220-239.

Haber, R. E., Juanes, C., del Toro, R., & Beruvides, G. (2015). Artificial cognitive control with self-x capabilities: A case study of a micro-manufacturing process. Computers in Industry, 74, 135-150.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. Nature, 466(7302), 29-29.

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. Nature Machine Intelligence, 1(9), 389-399.

Kandiah, S. (2020). Theta- and alpha-power dynamics related to response inhibition

Kane, N., Acharya, J., Beniczky, S., Caboclo, L., Finnigan, S., Kaplan, P. W., ... & van Putten, M. J. (2017). A revised glossary of terms most commonly used by clinical electroencephalographers and updated proposal for the report format of the EEG findings. Revision 2017. Clinical neurophysiology practice, 2, 170.

Keller, A. S., Payne, L., & Sekuler, R. (2017). Characterizing the roles of alpha and theta oscillations in multisensory attention. Neuropsychologia, 99, 48-63.

Kenemans, J. L. (2015). Specific proactive and generic reactive inhibition. Neuroscience & Biobehavioral Reviews, 56, 115-126.

Lieto, A., Bhatt, M., Oltramari, A., & Vernon, D. (2018). The role of cognitive architectures in general artificial intelligence.

Linear discriminant analysis - Wikipedia. (2021). Retrieved 26 June 2021, from https://en.wikipedia.org/wiki/Linear_discriminant_analysis

Lipton, Z. C. (2018). The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. Queue, 16(3), 31-57.

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., & Yger, F. (2018). A review of classification algorithms for EEG-based brain–computer interfaces: a 10 year update. Journal of neural engineering, 15(3), 031005.

Mazaheri, A., Nieuwenhuis, I. L., Van Dijk, H., & Jensen, O. (2009). Prestimulus alpha and mu activity predicts failure to inhibit motor responses. Human brain mapping, 30(6), 1791-1800.

O'Connell, R. G., Dockree, P. M., Robertson, I. H., Bellgrove, M. A., Foxe, J. J., & Kelly, S. P. (2009). Uncovering the neural signature of lapsing attention: electrophysiological signals predict errors up to 20 s before they occur. Journal of Neuroscience, 29(26), 8604-8611.

Ouyang, G., Sommer, W., & Zhou, C. (2015). A toolbox for residue iteration decomposition (RIDE)—A method for the decomposition, reconstruction, and single trial analysis of event related potentials. Journal of neuroscience methods, 250, 7-21.

Pion-Tonachini, L., Kreutz-Delgado, K., & Makeig, S. (2019). ICLabel: An automated electroencephalographic independent component classifier, dataset, and website. NeuroImage, 198, 181-197.

Ribeiro, M. T., Singh, S., & Guestrin, C. (2016, August). " Why should I trust you?" Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining (pp. 1135-1144).

Rodriguez-Bermudez, G., & Garcia-Laencina, P. J. (2015). Analysis of EEG signals using nonlinear dynamics and chaos: a review. Applied mathematics & information sciences, 9(5), 2309.

Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. Journal of neural engineering, 16(5), 051001.

Rudin, C. (2019). Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. Nature Machine Intelligence, 1(5), 206-215.

Urigüen, J. A., & Garcia-Zapirain, B. (2015). EEG artifact removal—state-of-the-art and guidelines. Journal of neural engineering, 12(3), 031001.

Ursin, H. (2005). Press stop to start: the role of inhibition for choice and health. Psychoneuroendocrinology, 30(10), 1059-1065.

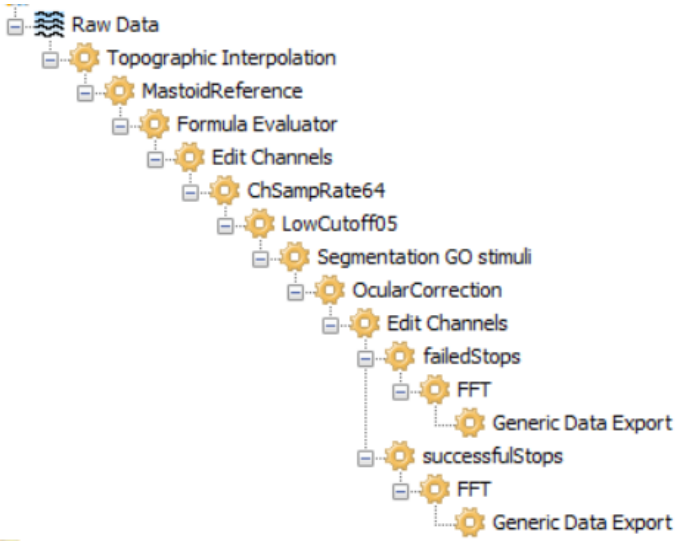
Vaid, S., Singh, P., & Kaur, C. (2015, February). EEG signal analysis for BCI interface: A review. In 2015 fifth international conference on advanced computing & communication technologies (pp. 143-147). IEEE.

Verbruggen, F., Aron, A. R., Band, G. P., Beste, C., Bissett, P. G., Brockett, A. T., ... & Boehler, C. N. (2019). A consensus guide to capturing the ability to inhibit actions and impulsive behaviors in the stop-signal task. elife, 8, e46323.

Xue, J. H., & Titterington, D. M. (2008). Do unbalanced data have a negative effect on LDA?. Pattern Recognition, 41(5), 1558-1571.

Appendix

1. BrainVision Analyzer complete pre-processing history tree




```
*** New Reference ***
Selected channels to include into the new reference:
EXG5
The implicit reference is included in the calculation of the new reference.

Channels to which the new reference applies to:
AF3 AF4 AF7 AF8 AFz C1 C2 C3 C4 C5
C6 CP1 CP2 CP3 CP4 CP5 CP6 CPz Cz F1
F2 F3 F4 F5 F6 F7 F8 FC1 FC2 FC3
FC4 FC5 FC6 FCz Fp1 Fp2 Fpz FT7 FT8 Fz
Iz O1 O2 Oz P1 P10 P2 P3 P4 P5
P6 P7 P8 P9 PO3 PO4 PO7 PO8 POz Pz
Status T7 T8 TP7 TP8
Remaining (non rereferenced) channels are kept.

Name of the new reference channel: MASTOID
```

```
*** Formula Evaluator ***

The following formulas were calculated:
VEOG = EXG4 - EXG3    Unit: µV
HEOG = EXG1 - EXG2    Unit: µV

The remaining channels were kept.
The new channels are on top.
```

```
*** Edit Channels ***

The following channels have been disabled:
EXG1    EXG2    EXG3    EXG4
EXG5    EXG6    EXG7    EXG8
Status
```

```
*** Change Sampling Rate / Resolution ***
Conversion is based on spline interpolation.
New Sampling Rate [Hz]: 64
New Sampling Interval [µS]: 15625
Data was filtered before downsampling with 28.8Hz, 24dB/oct.
```

```
*** IIR Filters ***
Zero phase shift Butterworth filters.

Global filter settings:
Low cutoff: 0.5 Hz, time constant 0.3183099, order 2
High cutoff: 28.8 Hz, order 2
Notch filter: ---

*** Segmentation ***
Segmentation relative to reference marker positions
Reference markers:
    Stimulus    S 22
    Stimulus    S 21
Advanced Boolean Expression:
    ---
Segment size and position relative to reference markers:
Start: -1000.00 ms, End: 1500.00 ms, Length: 2500.00 ms
Allow overlapped segments? Yes
Skip bad intervals?        Yes

Data was not stored but will be calculated on demand.

*** Data node specific information ***

Number of segments: 257

*** Ocular Correction (Gratton & Coles) ***
Name of HEOG channel: HEOG
Common reference
Name of VEOG channel: VEOG
Common reference

The following channels have been selected for correction:
VEOG    HEOG    Fp1     AF7
AF3     F1      F3      F5
F7      FT7     FC5     FC3
FC1     C1      C3      C5
T7      TP7     CP5     CP3
CP1     P1      P3      P5
P7      P9      PO7     PO3
O1      Iz      Oz      POz
Pz      CPz     Fpz     Fp2
AF8     AF4     AFz     Fz
F2      F4      F6      F8
FT8     FC6     FC4     FC2
FCz     Cz      C2      C4
C6      T8      TP8     CP6
CP4     CP2     P2      P4
P6      P8      P10     PO8
PO4     O2

Blinkdetection by algorithm

Number of trial types: 2
S 21: 129 Segments
S 22: 128 Segments

Number of blinks: 339

*** Edit Channels ***

The following channels have been disabled:
VEOG    HEOG
```

```
*** Segmentation ***
Segmentation relative to reference marker positions
Reference markers:
    Stimulus    S 22
    Stimulus    S 21
Advanced Boolean Expression:
    S1(0,1000) or S2(0,1000)
Segment size and position relative to reference markers:
Start: -1000.00 ms, End: 0.00 ms, Length: 1000.00 ms
Allow overlapped segments? Yes
Skip bad intervals?        Yes

*** Fast Fourier Transformation (FFT) ***
Maximum Resolution
Power
Non-Complex Output
Half Spectrum used

Data Window:
Hanning Window
Length = 10 %
Variance Correction used
Periodic


*** Data node specific information ***
Resolution: 1 Hz

*** Generic Data Export ***
File name parameter: failed101
File extension: .dat
Write header file: yes
Write marker file: no
Format: ASCII
Orientation: VECTORIZED
Line Delimiter: CRLF (PC style)
Add channel names: yes
Overwrite default decimal symbol: no
Export all channels: yes

*** Segmentation ***
Segmentation relative to reference marker positions
Reference markers:
    Stimulus    S 22
    Stimulus    S 21
Advanced Boolean Expression:
    not S1(0,1000) and not S2(0,1000)
Segment size and position relative to reference markers:
Start: -1000.00 ms, End: 0.00 ms, Length: 1000.00 ms
Allow overlapped segments? Yes
Skip bad intervals?        Yes

Data was not stored but will be calculated on demand.

*** Data node specific information ***

Number of segments: 114
```

```
*** Fast Fourier Transformation (FFT) ***
Maximum Resolution
Power
Non-Complex Output
Half Spectrum used

Data Window:
Hanning Window
Length = 10 %
Variance Correction used
Periodic


*** Data node specific information ***
Resolution: 1 Hz

*** Generic Data Export ***
File name parameter: successfull01
File extension: .dat
Write header file: yes
Write marker file: no
Format: ASCII
Orientation: VECTORIZED
Line Delimiter: CRLF (PC style)
Add channel names: yes
Overwrite default decimal symbol: no
Export all channels: yes
```

2. Statistical evaluation of the *averaged 5-times repeated class-balancing LOO-procedure*

VARIABLE		MEAN	VARIANCE	STADEV	F-VALUE	P <	
1	1	.0179	.0021	.0462	4.3749	.0457	#
2	2	-.0159	.0026	.0509	2.8178	.1043	
3	3	.0162	.0021	.0456	3.6594	.0660	
4	4	-.0117	.0044	.0666	.8986	.3513	
5	5	-.0093	.0026	.0510	.9670	.3338	
6	6	.0083	.0032	.0562	.6290	.4344	
7	7	-.0003	.0030	.0544	.0012	.9730	
8	8	.0159	.0027	.0523	2.6705	.1134	
9	9	.0052	.0024	.0486	.3285	.5711	
10	10	.0052	.0029	.0535	.2711	.6067	
11	11	.0103	.0029	.0541	1.0585	.3124	
12	12	.0252	.0055	.0743	3.3297	.0787	
13	13	-.0059	.0029	.0536	.3466	.5608	
14	14	.0207	.0013	.0364	9.3472	.0049	##
15	15	.0197	.0032	.0570	3.4511	.0738	
16	16	.0052	.0029	.0540	.2658	.6102	
17	17	-.0166	.0035	.0594	2.2549	.1444	
18	18	.0079	.0013	.0363	1.3851	.2491	
19	19	.0155	.0039	.0624	1.7951	.1911	
20	20	.0131	.0016	.0397	3.1672	.0860	

3. Metrics 4-12 Hz for the *averaged 5-times repeated class-balancing LOO-procedure*

freq	acc	prec	rec	spec	neg	F1	MathCC	ptp
4	0,49	0,49	0,49	0,49	0,49	0,49	-0,011	001
5	0,51	0,51	0,49	0,53	0,51	0,5	0,023	001
6	0,48	0,48	0,44	0,53	0,48	0,46	-0,033	001
7	0,52	0,52	0,53	0,52	0,52	0,53	0,046	001
8	0,51	0,51	0,51	0,52	0,51	0,51	0,023	001
9	0,49	0,49	0,44	0,54	0,49	0,46	-0,022	001
10	0,42	0,43	0,45	0,39	0,42	0,44	-0,146	001
11	0,56	0,56	0,53	0,58	0,55	0,54	0,119	001
4	0,5	0,5	0,48	0,52	0,5	0,49	0	002
5	0,59	0,59	0,61	0,57	0,6	0,6	0,205	002
6	0,64	0,63	0,69	0,59	0,65	0,65	0,328	002
7	0,44	0,43	0,41	0,46	0,44	0,42	-0,122	002
8	0,45	0,46	0,48	0,43	0,45	0,47	-0,089	002
9	0,48	0,48	0,43	0,54	0,48	0,45	-0,037	002
10	0,56	0,56	0,59	0,54	0,57	0,58	0,139	002
11	0,48	0,48	0,46	0,5	0,48	0,47	-0,036	002
4	0,41	0,41	0,4	0,42	0,41	0,4	-0,164	003
5	0,48	0,48	0,43	0,52	0,48	0,45	-0,044	003
6	0,54	0,54	0,53	0,54	0,54	0,54	0,081	003
7	0,62	0,62	0,62	0,61	0,62	0,62	0,267	003
8	0,47	0,46	0,43	0,5	0,47	0,45	-0,065	003
9	0,49	0,49	0,46	0,52	0,49	0,47	-0,022	003
10	0,55	0,55	0,53	0,57	0,55	0,54	0,105	003
11	0,54	0,54	0,58	0,51	0,55	0,56	0,093	003
4	0,5	0,5	0,5	0,5	0,5	0,5	0	004
5	0,57	0,57	0,54	0,59	0,56	0,55	0,141	004
6	0,49	0,49	0,53	0,46	0,49	0,51	-0,013	004
7	0,52	0,52	0,47	0,57	0,52	0,5	0,04	004
8	0,48	0,48	0,47	0,49	0,48	0,48	-0,039	004
9	0,5	0,5	0,5	0,5	0,5	0,5	0	004
10	0,42	0,42	0,43	0,41	0,42	0,43	-0,147	004
11	0,4	0,41	0,45	0,36	0,39	0,43	-0,181	004
4	0,51	0,51	0,48	0,55	0,51	0,5	0,026	005
5	0,52	0,52	0,49	0,55	0,52	0,51	0,04	005
6	0,5	0,5	0,52	0,48	0,5	0,51	0	005
7	0,52	0,52	0,49	0,55	0,52	0,51	0,04	005
8	0,54	0,54	0,56	0,52	0,54	0,55	0,081	005
9	0,44	0,43	0,42	0,45	0,44	0,42	-0,122	005
10	0,52	0,52	0,52	0,52	0,52	0,52	0,04	005
11	0,5	0,5	0,53	0,47	0,5	0,52	0	005
4	0,44	0,44	0,42	0,47	0,45	0,43	-0,106	009
5	0,46	0,46	0,48	0,43	0,45	0,47	-0,083	009
6	0,55	0,55	0,52	0,58	0,55	0,54	0,104	009
7	0,56	0,56	0,54	0,58	0,56	0,55	0,132	009
8	0,54	0,55	0,52	0,57	0,54	0,53	0,091	009
9	0,59	0,59	0,56	0,62	0,58	0,57	0,19	009
10	0,41	0,41	0,41	0,42	0,41	0,41	-0,16	009
11	0,54	0,55	0,52	0,57	0,54	0,53	0,091	009
4	0,49	0,49	0,49	0,49	0,49	0,49	-0,022	010
5	0,5	0,5	0,51	0,49	0,5	0,51	0	010
6	0,48	0,48	0,49	0,48	0,48	0,49	-0,032	010
7	0,51	0,51	0,5	0,52	0,51	0,51	0,022	010
8	0,58	0,57	0,64	0,51	0,59	0,6	0,167	010
9	0,56	0,56	0,57	0,55	0,56	0,56	0,127	010
10	0,47	0,47	0,43	0,51	0,47	0,45	-0,053	010
11	0,47	0,47	0,45	0,49	0,47	0,46	-0,063	010
4	0,48	0,48	0,47	0,5	0,48	0,48	-0,031	011
5	0,5	0,5	0,59	0,41	0,5	0,54	0	011
6	0,48	0,48	0,34	0,62	0,49	0,4	-0,032	011
7	0,52	0,52	0,44	0,59	0,51	0,47	0,032	011
8	0,45	0,45	0,41	0,5	0,46	0,43	-0,09	011
9	0,5	0,5	0,66	0,34	0,5	0,57	0	011
10	0,52	0,51	0,56	0,47	0,52	0,54	0,032	011
11	0,36	0,36	0,38	0,34	0,35	0,37	-0,249	011
4	0,42	0,41	0,4	0,44	0,42	0,4	-0,151	012
5	0,52	0,52	0,56	0,48	0,52	0,54	0,036	012
6	0,49	0,49	0,44	0,53	0,49	0,46	-0,023	012
7	0,6	0,6	0,58	0,62	0,6	0,59	0,221	012
8	0,48	0,48	0,45	0,5	0,48	0,46	-0,046	012
9	0,47	0,47	0,48	0,45	0,46	0,47	-0,067	012
10	0,53	0,53	0,56	0,5	0,53	0,54	0,06	012
11	0,43	0,43	0,43	0,43	0,43	0,43	-0,131	012
4	0,49	0,49	0,48	0,51	0,49	0,49	-0,012	013
5	0,57	0,57	0,58	0,56	0,57	0,57	0,152	013
6	0,42	0,42	0,38	0,47	0,43	0,4	-0,143	013
7	0,53	0,53	0,51	0,55	0,53	0,52	0,061	013
8	0,56	0,56	0,54	0,58	0,56	0,55	0,125	013
9	0,47	0,47	0,47	0,47	0,47	0,47	-0,057	013
10	0,49	0,49	0,53	0,45	0,49	0,51	-0,023	013
11	0,45	0,45	0,46	0,44	0,45	0,45	-0,101	013

freq	acc	prec	rec	spec	neg	F1	MathCC	ptp
4	0,44	0,45	0,51	0,36	0,43	0,48	-0,119	014
5	0,51	0,51	0,54	0,48	0,51	0,53	0,021	014
6	0,55	0,55	0,61	0,49	0,56	0,58	0,111	014
7	0,53	0,53	0,51	0,54	0,53	0,52	0,054	014
8	0,48	0,48	0,46	0,5	0,48	0,47	-0,041	014
9	0,51	0,51	0,53	0,49	0,51	0,52	0,021	014
10	0,53	0,52	0,55	0,5	0,53	0,54	0,054	014
11	0,53	0,53	0,52	0,54	0,53	0,53	0,065	014
4	0,44	0,45	0,48	0,41	0,44	0,46	-0,108	015
5	0,53	0,53	0,56	0,51	0,54	0,54	0,071	015
6	0,57	0,57	0,53	0,6	0,56	0,55	0,147	015
7	0,62	0,61	0,65	0,59	0,63	0,63	0,274	015
8	0,51	0,51	0,53	0,48	0,51	0,52	0,011	015
9	0,53	0,53	0,49	0,57	0,53	0,51	0,059	015
10	0,59	0,58	0,65	0,52	0,6	0,61	0,189	015
11	0,55	0,55	0,57	0,53	0,55	0,56	0,108	015
4	0,49	0,49	0,49	0,49	0,49	0,49	-0,021	016
5	0,53	0,53	0,51	0,54	0,53	0,52	0,055	016
6	0,48	0,48	0,48	0,49	0,48	0,48	-0,031	016
7	0,47	0,47	0,45	0,5	0,47	0,46	-0,052	016
8	0,51	0,51	0,46	0,55	0,5	0,48	0,011	016
9	0,46	0,46	0,47	0,45	0,46	0,46	-0,082	016
10	0,53	0,54	0,48	0,59	0,53	0,51	0,066	016
11	0,54	0,55	0,49	0,6	0,54	0,52	0,089	016
4	0,59	0,59	0,65	0,54	0,61	0,62	0,211	017
5	0,55	0,55	0,55	0,54	0,55	0,55	0,099	017
6	0,54	0,54	0,55	0,53	0,54	0,55	0,085	017
7	0,51	0,51	0,51	0,5	0,51	0,51	0,014	017
8	0,55	0,55	0,57	0,54	0,56	0,56	0,115	017
9	0,47	0,47	0,46	0,47	0,47	0,46	-0,065	017
10	0,45	0,46	0,5	0,41	0,45	0,48	-0,091	017
11	0,47	0,48	0,53	0,42	0,47	0,5	-0,053	017
4	0,51	0,51	0,52	0,49	0,51	0,51	0,011	018
5	0,53	0,53	0,49	0,56	0,53	0,51	0,058	018
6	0,52	0,52	0,49	0,55	0,52	0,51	0,046	018
7	0,52	0,52	0,54	0,49	0,52	0,53	0,034	018
8	0,52	0,52	0,48	0,55	0,52	0,5	0,034	018
9	0,43	0,43	0,42	0,44	0,43	0,42	-0,136	018
10	0,61	0,6	0,7	0,53	0,64	0,64	0,259	018
11	0,62	0,63	0,56	0,67	0,61	0,6	0,272	018
4	0,4	0,39	0,37	0,42	0,4	0,38	-0,187	019
5	0,6	0,59	0,67	0,54	0,62	0,63	0,232	019
6	0,44	0,44	0,41	0,47	0,45	0,42	-0,109	019
7	0,53	0,53	0,54	0,51	0,53	0,53	0,053	019
8	0,62	0,62	0,63	0,62	0,62	0,62	0,28	019
9	0,64	0,66	0,58	0,71	0,62	0,62	0,336	019
10	0,58	0,59	0,55	0,62	0,58	0,57	0,183	019
11	0,51	0,51	0,47	0,54	0,51	0,49	0,013	019
4	0,58	0,58	0,56	0,59	0,57	0,57	0,167	020
5	0,56	0,56	0,56	0,55	0,56	0,56	0,123	020
6	0,52	0,52	0,53	0,51	0,52	0,52	0,039	020
7	0,57	0,58	0,53	0,62	0,56	0,55	0,153	020
8	0,53	0,53	0,58	0,49	0,54	0,55	0,067	020
9	0,49	0,49	0,5	0,47	0,49	0,49	-0,025	020
10	0,51	0,51	0,59	0,44	0,52	0,55	0,026	020
11	0,53	0,53	0,51	0,54	0,53	0,52	0,053	020
4	0,57	0,57	0,57	0,58	0,57	0,57	0,152	021
5	0,53	0,53	0,51	0,55	0,53	0,52	0,068	021
6	0,41	0,41	0,41	0,41	0,41	0,41	-0,161	021
7	0,45	0,45	0,43	0,47	0,45	0,44	-0,093	021
8	0,57	0,57	0,54	0,59	0,56	0,56	0,14	021
9	0,64	0,65	0,62	0,66	0,64	0,63	0,334	021
10	0,57	0,57	0,57	0,57	0,57	0,57	0,14	021
11	0,51	0,51	0,47	0,55	0,51	0,49	0,022	021
4	0,43	0,43	0,47	0,38	0,42	0,45	-0,139	022
5	0,45	0,45	0,44	0,46	0,45	0,45	-0,094	022
6	0,39	0,38	0,35	0,43	0,4	0,36	-0,202	022
7	0,51	0,51	0,48	0,53	0,51	0,49	0,012	022
8	0,48	0,47	0,47	0,48	0,48	0,47	-0,048	022
9	0,48	0,47	0,44	0,51	0,48	0,46	-0,048	022
10	0,61	0,6	0,65	0,57	0,62	0,63	0,253	022
11	0,56	0,56	0,58	0,54	0,56	0,57	0,132	022
4	0,5	0,5	0,54	0,47	0,5	0,52	0,009	101
5	0,43	0,43	0,46	0,39	0,42	0,45	-0,132	101
6	0,49	0,49	0,48	0,49	0,49	0,48	-0,026	101
7	0,48	0,48	0,5	0,46	0,48	0,49	-0,043	101
8	0,43	0,43	0,44	0,42	0,43	0,43	-0,131	101
9	0,56	0,56	0,55	0,56	0,56	0,56	0,121	101
10	0,45	0,46	0,49	0,41	0,45	0,47	-0,092	101
11	0,61	0,62	0,56	0,66	0,6	0,59	0,249	101

freq	acc	prec	rec	spec	neg	F1	MathCC	ptp
4	0,47	0,47	0,5	0,45	0,47	0,49	-0,052	103
5	0,49	0,49	0,51	0,46	0,49	0,5	-0,026	103
6	0,55	0,55	0,57	0,54	0,56	0,56	0,113	103
7	0,44	0,44	0,46	0,42	0,44	0,45	-0,11	103
8	0,48	0,48	0,54	0,42	0,47	0,51	-0,044	103
9	0,52	0,52	0,52	0,52	0,52	0,52	0,036	103
10	0,51	0,51	0,53	0,5	0,51	0,52	0,027	103
11	0,52	0,52	0,52	0,53	0,52	0,52	0,046	103
4	0,5	0,5	0,5	0,5	0,5	0,5	0	104
5	0,33	0	0	0,67	0,4	0	-0,387	104
6	0,42	0,33	0,17	0,67	0,44	0,22	-0,178	104
7	0,42	0,43	0,5	0,33	0,4	0,46	-0,156	104
8	0,5	0,5	0,33	0,67	0,5	0,4	0	104
9	0,5	0,5	0,67	0,33	0,5	0,57	0	104
10	0,5	0,5	0,33	0,67	0,5	0,4	0	104
11	0,75	0,71	0,83	0,67	0,8	0,77	0,717	104
4	0,54	0,54	0,53	0,55	0,54	0,54	0,09	106
5	0,54	0,55	0,5	0,58	0,54	0,52	0,09	106
6	0,48	0,48	0,46	0,5	0,48	0,47	-0,037	106
7	0,53	0,53	0,45	0,61	0,52	0,49	0,06	106
8	0,54	0,54	0,49	0,59	0,53	0,51	0,08	106
9	0,53	0,54	0,49	0,58	0,53	0,51	0,069	106
10	0,52	0,53	0,48	0,57	0,52	0,5	0,049	106
11	0,54	0,54	0,52	0,55	0,54	0,53	0,079	106
4	0,55	0,55	0,58	0,52	0,55	0,56	0,106	107
5	0,56	0,56	0,56	0,56	0,56	0,56	0,13	107
6	0,57	0,58	0,54	0,6	0,57	0,56	0,155	107
7	0,42	0,43	0,44	0,41	0,42	0,44	-0,14	107
8	0,56	0,56	0,54	0,59	0,56	0,55	0,13	107
9	0,49	0,49	0,54	0,44	0,49	0,51	-0,021	107
10	0,51	0,51	0,58	0,44	0,51	0,54	0,015	107
11	0,49	0,49	0,51	0,47	0,49	0,5	-0,021	107
4	0,46	0,45	0,43	0,48	0,46	0,44	-0,083	108
5	0,49	0,49	0,49	0,49	0,49	0,49	-0,023	108
6	0,45	0,44	0,4	0,5	0,45	0,42	-0,098	108
7	0,53	0,53	0,5	0,56	0,53	0,51	0,057	108
8	0,43	0,42	0,42	0,43	0,43	0,42	-0,14	108
9	0,44	0,45	0,47	0,42	0,44	0,46	-0,105	108
10	0,53	0,53	0,52	0,54	0,53	0,53	0,065	108
11	0,56	0,56	0,55	0,56	0,55	0,55	0,117	108
4	0,49	0,49	0,48	0,5	0,49	0,48	-0,028	109
5	0,5	0,5	0,52	0,49	0,5	0,51	0,009	109
6	0,51	0,51	0,5	0,53	0,51	0,5	0,028	109
7	0,47	0,47	0,49	0,46	0,47	0,48	-0,055	109
8	0,5	0,5	0,56	0,45	0,51	0,53	0,009	109
9	0,44	0,44	0,44	0,45	0,44	0,44	-0,106	109
10	0,5	0,5	0,47	0,52	0,5	0,48	-0,009	109
11	0,54	0,53	0,58	0,5	0,54	0,56	0,078	109
4	0,58	0,57	0,6	0,55	0,58	0,59	0,164	110
5	0,42	0,41	0,38	0,46	0,43	0,39	-0,15	110
6	0,51	0,51	0,53	0,49	0,51	0,52	0,022	110
7	0,53	0,53	0,55	0,52	0,53	0,54	0,067	110
8	0,49	0,49	0,52	0,47	0,49	0,51	-0,011	110
9	0,47	0,47	0,46	0,47	0,47	0,46	-0,063	110
10	0,49	0,49	0,52	0,47	0,49	0,51	-0,011	110
11	0,6	0,59	0,62	0,57	0,6	0,61	0,216	110
4	0,5	0,5	0,5	0,51	0,5	0,5	0,008	111
5	0,51	0,51	0,52	0,5	0,51	0,52	0,024	111
6	0,52	0,52	0,51	0,54	0,52	0,52	0,048	111
7	0,54	0,54	0,56	0,53	0,54	0,55	0,091	111
8	0,47	0,47	0,46	0,47	0,47	0,47	-0,061	111
9	0,52	0,52	0,54	0,5	0,52	0,53	0,032	111
10	0,45	0,46	0,51	0,39	0,45	0,48	-0,091	111
11	0,47	0,47	0,51	0,43	0,47	0,49	-0,061	111
4	0,46	0,46	0,48	0,44	0,46	0,47	-0,073	112
5	0,46	0,47	0,52	0,41	0,46	0,49	-0,073	112
6	0,5	0,5	0,5	0,49	0,5	0,5	-0,009	112
7	0,55	0,55	0,54	0,56	0,55	0,54	0,099	112
8	0,42	0,42	0,42	0,42	0,42	0,42	-0,141	112
9	0,54	0,54	0,53	0,56	0,54	0,54	0,089	112
10	0,47	0,47	0,48	0,45	0,47	0,47	-0,064	112
11	0,6	0,61	0,58	0,62	0,6	0,6	0,233	112

4. Experimentally optimized subsets of used electrodes

ALL	4	5	6	7	8	9	10	11
Fp1				Fp1				
AF7				AF7				
AF3				AF3				AF3
F1				F1				F1
F3				F3				
F5								
F7				F7				
FT7				FT7				
FC5				FC5				
FC3				FC3				FC3
FC1							FC1	FC1
C1	C1			C1			C1	C1
C3	C3			C3				
C5				C5				
T7				T7				
TP7				TP7				
CP5				CP5				CP5
CP3				CP3	CP3		CP3	CP3
CP1	CP1	CP1		CP1	CP1		CP1	CP1
P1	P1	P1		P1	P1		P1	P1
P3	P3	P3		P3	P3		P3	P3
P5				P5	P5			P5
P7		P7		P7				P7
P9				P9				
PO7				PO7				
PO3	PO3	PO3		PO3	PO3			PO3
O1				O1				
Iz								
Oz		Oz		Oz				Oz
POz	POz	POz	POz	POz	POz		POz	POz
Pz	Pz	Pz		Pz	Pz		Pz	Pz
CPz	CPz			CPz	CPz		CPz	CPz
Fpz								
Fp2				Fp2				
AF8				AF8				
AF4								
AFz				AFz				
Fz				Fz	Fz		Fz	Fz
F2	F2	F2		F2				F2
F4				F4				F4
F6				F6				
F8				F8				
FT8				FT8				
FC6				FC6				FC6
FC4				FC4				FC4
FC2				FC2				FC2
FCz				FCz				FCz
Cz				Cz				Cz
C2	C2			C2	C2		C2	C2
C4				C4				C4
C6				C6				
T8				T8				
TP8				TP8				
CP6				CP6				
CP4	CP4	CP4		CP4			CP4	CP4
CP2	CP2	CP2		CP2			CP2	CP2
P2	P2			P2		P2	P2	P2
P4				P4			P4	P4
P6				P6				P6
P8				P8			P8	P8
P10				P10				
PO8				PO8				
PO4	PO4	PO4	PO4	PO4			PO4	PO4
O2		O2		O2	O2			

5. Python code

```
""" Utrecht University 2021 - Tjalle Galama 6936385 : This tool is designed for
finding the most informative linear discriminants in pre-processed EEG data """

import os
import math
import random
import numpy as np
import pandas as pd
import time

# ( Practical sidenote: ' ; ' was used as a delimiter in both BVA and Excel )

# set this string to where the folder /script is located:
wd = "C:/Users/tjall/Onedrive/Documents/studie/UU/thesis/script"

# changeable parameters:
used_freqc = list(range(0,20)) # or for example [4,5,6,7], or just [11]
run_FLDA_LOO = False # by default not balancing classes
run_FLDA_LOO_k_times = True ; k = 5 # this one balances

find_optimal_subset = False # for given used_freqc. warning: takes a long time

# ------------------------------------------------------------------------- #
# The folders where csv's are located and where results will be stored
data_import = wd + "/import" ; results = wd + "/results/"

freq = 32 # the number of 1 Hz frequency intervals, as in the used csv files
used_electrodes = ['Fp1', 'AF7', 'AF3', 'F1', 'F3', 'F5', 'F7', 'FT7', 'FC5',
      'FC3', 'FC1','C1', 'C3', 'C5', 'T7', 'TP7', 'CP5', 'CP3', 'CP1', 'P1',
      'P3', 'P5','P7', 'P9', 'PO7', 'PO3', 'O1', 'Iz', 'Oz', 'POz', 'Pz',
      'CPz', 'Fpz', 'Fp2', 'AF8', 'AF4', 'AFz', 'Fz', 'F2', 'F4', 'F6', 'F8',
      'FT8', 'FC6','FC4', 'FC2', 'FCz', 'Cz', 'C2', 'C4', 'C6', 'T8', 'TP8',
      'CP6', 'CP4','CP2', 'P2', 'P4', 'P6', 'P8', 'P10', 'PO8', 'PO4', 'O2']

# a function to list the individual csv file locations:
def list_file_locations(folder):

    loc_failed = [] ; loc_success = []

    with os.scandir(folder) as it:
        for entry in it:
            if entry.name.endswith(".csv") and entry.is_file():
                if entry.name.startswith('failed'):
                    loc_failed.append(entry.path)

    with os.scandir(folder) as it:
        for entry in it:
            if entry.name.endswith(".csv") and entry.is_file():
                if entry.name.startswith('successful'):
                    loc_success.append(entry.path)

    participants_list = []
    for string_failed in loc_failed:
        ptcp = string_failed[-7:-4]
        participants_list.append(ptcp)

    # the function has the following two variables as output:
    return loc_failed, loc_success, participants_list

# this function is now executed (as follows):
loc_failed, loc_success, participants_list = list_file_locations(data_import)

# The second function imports the data of one participant into a dataframe:
```

```
def import_one(string_failed,string_success,used_electrodes):

    # import spectra of the seconds before failed responses
    f = pd.read_csv(string_failed,sep=";",header=None,index_col=0,decimal=',')
    # transpose the dataframe and an index is added
    f = pd.DataFrame.transpose(f)
    # only keep the columns of electrodes that we have initialized
    f = f[f.columns.intersection(used_electrodes)]
    # add a column with the class-label
    f['class'] = 'f'
    # add a column for corresponding frequencies (0 means 0-1 Hz, etcetera)
    f['freqc'] = np.tile(np.array(list(range(0,freq))),int(len(f)/freq))
    # add a column for the trial-numbers
    f['trial'] = np.repeat(np.arange(0,len(f)/freq),freq)

    # likewise, import successful
    s = pd.read_csv(string_success,sep=";",header=None,index_col=0,decimal=',')
    s = pd.DataFrame.transpose(s)
    s = s[s.columns.intersection(used_electrodes)]
    s['class'] = 's'
    s['freqc'] = np.tile(np.array(list(range(0,freq))),int(len(s)/freq))
    s['trial'] = np.repeat(np.arange(0,len(s)/freq),freq)

    # Now that we have the data, we can balance the classes.
    # Whether this will be done, depends on the parameter 'balance_classes':

    kept = pd.DataFrame([]) # storing which trials are kept
    ptcp = string_failed[-7:-4] # participant number (3-digit before '.csv')

    if balance_classes == True:
        # first define how many trials we want per class (taking the minimum):
        trials_per_class = int(min(len(s),len(f))/freq)
        if len(s)>len(f):
            # transforming an integer-float to int + set the number of s trials
            trials = int(len(f)/freq)
            # sample unique indices
            sample_indices = random.sample(range(0,trials),trials_per_class)
            # select only the trials with the corresponding sampled indices
            s = s[s['trial'].isin(sample_indices)]
            # For reproducibility, save which trials are kept
            kept['{0}, s'.format(ptcp)] = sorted(sample_indices)
            print('For ptcp {0}, {1} s-trials are kept.'.format(ptcp,trials))
        elif len(s)<len(f):
            trials = int(len(s)/freq)
            sample_indices = random.sample(range(0,trials),trials_per_class)
            f = f[f['trial'].isin(sample_indices)]
            kept['{0}, f'.format(ptcp)] = sorted(sample_indices)
            print('For ptcp {0}, {1} f-trials are kept.'.format(ptcp,trials))
        elif len(s)==len(f):
            pass

    # paste the failed and successful dataframes under each other
    dataset = pd.concat([f,s], axis=0, ignore_index=True)
    # add a column for the participant number
    dataset['ptp'] = ptcp
    # we only use 0-20 Hz, but can also give another subset to used_freqc
    data = dataset[dataset['freqc'].isin(used_freqc)]

    return data, kept

# A function to import the data of all the participants:
def import_all(used_electrodes):

    full_kept = pd.DataFrame([]) # herein, save the used trials
    full_data = pd.DataFrame([]) # herein we store all the data
```

```
    for h in range(0,len(loc_failed)):

        data, kept = import_one(loc_failed[h],loc_success[h],used_electrodes)

        full_kept = pd.concat([full_kept,kept], axis=1) # add column wise
        full_data = pd.concat([full_data,data], axis=0) # add row wise

    return full_kept, full_data

# This function performs Fisher's LDA with LeaveOneOut cross-validation:
def find_weights_LOO(dataset,used_electrodes):

    dataset = dataset.drop(['ptp'], axis=1)

    weights = [] ; out = []

    e = len(used_electrodes)

    amount_of_trials = int(len(dataset)/len(used_freqc))

    # Add a column with a new enumeration for the total amount of trials, which
    # is needed because some (or many) trial numbers are doubled (f & s)
    dataset['tot'] = np.repeat(np.arange(0,amount_of_trials),len(used_freqc))

    for h in range(0,amount_of_trials): # start looping through the trials

        left_out = dataset.loc[dataset['tot'] == h] # leave a trial out
        label = left_out['class'].values[0] # & save its class ('label')
        trial = left_out['trial'].values[0] # & save the trial

        # drop frequency, class and trial columns for the upcoming calculations
        left_out = left_out.drop(['class','trial','tot'], axis=1)

        # reset index of left-out trial to preserve frequency enumeration
        left_out = left_out.reset_index(drop=True)

        # continue with all the other trials
        rest = dataset.loc[dataset['tot'] != h]

        for i in used_freqc: # start looping through the frequencies

            # only take the row with the specific frequency
            left_i = left_out.loc[left_out['freqc'] == i]
            left_i = left_i.drop(['freqc'], axis = 1)

            trials_of_freq_i = rest.loc[rest['freqc'] == i]

            f_freq_i = trials_of_freq_i.loc[trials_of_freq_i['class'] == 'f']
            s_freq_i = trials_of_freq_i.loc[trials_of_freq_i['class'] == 's']

            f_freq_i = f_freq_i.drop(['freqc','class','trial','tot'], axis = 1)
            s_freq_i = s_freq_i.drop(['freqc','class','trial','tot'], axis = 1)

            # take means over electrodes per class
            f_i_mean = f_freq_i.mean().values[0:e]
            s_i_mean = s_freq_i.mean().values[0:e]

            # initialize the within-class scatter matrix (as an empty matrix)
            S_W = np.zeros((e,e))

            # scatter matrix of the first class
            class_sc_mat_f = np.zeros((e,e))
            for row in f_freq_i.values:

                row = row.reshape(e,1)
                mv = f_i_mean.reshape(e,1)
```

```
            class_sc_mat_f += (row - mv).dot((row - mv).T)

        # scatter matrix for the next class
        class_sc_mat_s = np.zeros((e,e))
        for row in s_freq_i.values:

            row = row.reshape(e,1)
            mv = s_i_mean.reshape(e,1)

            class_sc_mat_s += (row - mv).dot((row - mv).T)

        # sum class scatter matrices
        S_W = class_sc_mat_f + class_sc_mat_s

        # calculate the inverse of the within-class scatter matrix
        # (using pseudo-inv. instead of inv to deal with singular matrices
        # and as a solution for participants that have a lower sample size
        # for one class than the number of electrodes)
        within_inversed = np.linalg.pinv(S_W)

        w = np.dot(within_inversed,(f_i_mean - s_i_mean))

        c = np.dot(w,(f_i_mean+s_i_mean)) / 2

        if np.dot(left_i,w) > c:
            prediction = 'f' # classify as failed
        elif np.dot(left_i,w) < c:
            prediction = 's' # classify as successful

        out.append([i,trial,label,prediction])
        w = w.round(5)
        weights.append(w)

out = pd.DataFrame(out,columns=['freqc','trial','class','pred'])

weights = pd.DataFrame(weights,columns=used_electrodes)

# count correct classifications across all trials (per frequency, per ptcp)
acc_per_freq = []

for j in used_freqc:

    freq_j = out.loc[out['freqc'] == j]

    P = freq_j[freq_j['class'] == 's']
    N = freq_j[freq_j['class'] == 'f']

    TP = len(P[P['pred'] == 's'])  # TruePositive
    TN = len(N[N['pred'] == 'f'])  # TrueNegative
    FP = len(N[N['pred'] == 's'])  # FalsePositive
    FN = len(P[P['pred'] == 'f'])  # FalseNegative

    prec = round((TP / (TP + FP)), 2) # precision

    recall = round((TP / (TP + FN)), 2) # sensitivity

    spec =  round((TN / (TN + FP)), 2) # specificity

    neg = round((TN / (TN + FN)), 2) # negative predictive value

    F1 = round((2*TP / (2*TP + FP + FN)), 2) # F1-score

    CC = (math.sqrt((TP+FP)*(FP+FN)*(TN+FP)*(TN+FN) ))
    if CC == 0: CC = 0.0001 # because we cannot divide by 0
    MathCC = (TP*TN - FP*FN) / CC
```

```
        MathCC = round(MathCC, 3)

        acc = round( ( (TP+TN) / (TP+TN+FP+FN) ) , 2)

        acc_per_freq.append([j,acc,prec,recall,spec,neg,F1,MathCC])

    columns = ['freq','acc','prec','rec','spec','neg','F1','MathCC']
    accuracy_per_freq = pd.DataFrame(acc_per_freq,columns=columns)

    # concat the weights and corresponding ['freqc','trial','class','pred']
    output = pd.concat([weights, out], axis=1)

    return accuracy_per_freq, output

# A function to run the FLDA LOO procedure for all participants at once:
def find_weights_LOO_all(full_data,used_elec):

    # here the accuracies and weights of all participants are stored
    accuracies = pd.DataFrame()
    total_output = pd.DataFrame()

    start = time.time() # initiate a timer

    for ptcp in participants_list:

        print('Processing participant {0}'.format(ptcp))
        # select the data of the ptcp
        participant = full_data[full_data['ptp'] == ptcp]
        # run the function on it
        accuracy_per_freq, output = find_weights_LOO(participant,used_elec)
        # give the received outputs the participant number as a column
        accuracy_per_freq['ptp'] = ptcp ; output['ptp'] = ptcp
        # add the retrieved accuracies to the dataframe for all participants
        accuracies = pd.concat([accuracies,accuracy_per_freq], axis=0)
        # add the retrieved weights to the dataframe for all participants
        total_output = pd.concat([total_output,output], axis=0)

    end = time.time() # end the timer

    print("Time taken for this single LOO workflow: {0}".format(end - start))
    print() # print a white line

    return accuracies, total_output

if run_FLDA_LOO == True:

    balance_classes = False # by default, because the next function balances

    full_kept, full_data = import_all() # load the data of all the participants
    # when running this, full_kept is empty because we do not do any balancing

    accuracies, total_output = find_weights_LOO_all(full_data,used_electrodes)

    # save the resulting accuracies and all the unique weights vectors
    accuracies.to_excel(results + "/accuracies.xlsx")
    total_output.to_excel(results + "/total_output.xlsx")

# When balancing classes, bias may occur. A possible solution is to run the
# algorithm k-times and to average the weights and accuracies afterwards.
def find_weights_LOO_k_times(k,used_electrodes):

    averaged_accuracies = pd.DataFrame()
    averaged_output = pd.DataFrame()

    i = 0
```

```
    while i < k:

        print('Running LOO round {0} to unbias class-imbalance.'.format(i+1))
        full_kept, full_data = import_all(used_electrodes) # 'full-kept' unused

        # getting the results for one run of the procedure:
        accuracies, output = find_weights_LOO_all(full_data, used_electrodes)

        # all the accuracies are put in one dataframe
        averaged_accuracies = pd.concat([averaged_accuracies,accuracies])

        # all the weights vectors are stored together as well
        averaged_output = pd.concat([averaged_output,output])

        i += 1

    # taking the means of the accuracies per frequency+participant combination:
    averaged_accuracies = averaged_accuracies.groupby(['freq','ptp']).mean()

    # dropping the columns 'class','pred','trial', since we don't average these
    averaged_output = averaged_output.drop(['class','pred','trial'],axis = 1)

    # take the means of the weights per frequency+participant combination
    averaged_output = averaged_output.groupby(['freqc','ptp']).mean()
    # Here, for optimal subset selection purposes, an alternative would be to
    # first take the absolute values of weight vectors before averaging them.

    return averaged_accuracies, averaged_output

if run_FLDA_LOO_k_times == True:

    balance_classes = True

    used_e = used_electrodes

    averaged_accuracies, averaged_output = find_weights_LOO_k_times(k, used_e)

    # again, store the results
    averaged_accuracies.to_excel(results + "/averaged_accuracies.xlsx")
    averaged_output.to_excel(results + "/averaged_output.xlsx")

# We can create some insight by counting which electrodes were used most with
# a function to sort the electrodes based on their absolute distance to 0:
def sort_electrodes_by_weight(output):

    # beginning with some empty storage:
    most_informative = pd.DataFrame()

    i = 0

    while i < len(output):

        # first take the absolute values of the i-th row
        to_be_ordered = abs(output.iloc[[i]])

        # take the index of this row. index[0] is of the form (freq,'ptcp')
        index = to_be_ordered.index

        # now order the row
        ordered = to_be_ordered.sort_values(by=index[0],axis=1,ascending=False)

        # a new row with the same index values + electrodes names ascending
        # (based on their respective weight values, which are now omitted)
        ordered = pd.Series(list(ordered.index[0])+list(ordered.columns))

        # adding this row to the storage
```

```
        most_informative = most_informative.append(ordered, ignore_index=True)

        i += 1

    # after transforming all weight vectors to their respective ascending
    # names of electrodes, we rename the first two columns.
    most_informative = most_informative.rename(columns={0:'freq', 1:'ptcp'})

    return most_informative # (open the variable to see what it looks like)

# given the sorted electrodes, we can try to find the optimal subset for the
# complete set of frequencies that was entered into the procedure at first.
def subset_finding(ordered_electrodes, all_electrodes):

    print('Starting the subset_finding procedure')

    m_i = pd.DataFrame()

    for ptp in participants_list:

        # select the ordered electrodes of one participant
        e_participant = most_informative[most_informative['ptcp'] == ptp]

        # remove the 'freq' and 'ptcp' column
        e_participant = e_participant.drop(['freq','ptcp'],axis=1)

        # only taking the first 32 most informative electrodes per participant
        # This is an uninformed choice (!)
        e_participant = e_participant.iloc[:,:32]

        # paste all the ordered rows of most informative frequencies under each
        # other + do add those of the other participants to this dataframe too
        for row in range(0,len(e_participant)):

            m_i = pd.concat([m_i,e_participant.iloc[row,]], axis=0)

    m_i = m_i.reset_index(drop=True) # reset the index

    # count the occurences of  29 * 32 * len(used_freqc) electrodes in total
    m_i = m_i.iloc[:,0].value_counts()

    # the initial number of used electrodes (from the ordening m_i)
    i =  1

    # total accuracy score of the original set, for comparisons
    leading_accuracy_sum = sum(averaged_accuracies['acc'])

    best_e = []

    opt_av_accuracies = []
    opt_av_out = []

    while i < len(all_electrodes):

        print("The number of used electrodes is {0}".format(i))

        # the top-i of the top-32, across participants (for all of used_freqc)
        used_e = np.array(m_i.iloc[:i].index)

        # sort the new list of electrodes according to the original sorting
        # (this is needed for assigning column names in the FLDA function)
        used_e = list(filter(lambda x:x in used_e, all_electrodes))

        # run the repeated balancing function with the new subset of electrodes
        av_accuracies, av_output = find_weights_LOO_k_times(k, used_e)
```

```
        # the sum of accuracies of all used frequencies, of all participants:
        freq_summed = sum(av_accuracies['acc'])

        if freq_summed > leading_accuracy_sum:

            leading_accuracy_sum = freq_summed

            best_e = used_e

            opt_av_accuracies = av_accuracies

            # resulting weight vectors of the best solution are saved
            opt_av_out = av_output

        i += 1

    # give the result the same original ordering as the initial used_electrodes
    best_e = list(filter(lambda x:x in best_e, all_electrodes))

    opt_av_accuracies.to_excel(results + "/opt_averaged_accuracies.xlsx")
    opt_av_out.to_excel(results + "/opt_averaged_output.xlsx")

    # return the best set, the resulting accuracies as well as the counting of
    # the occurrences of electrodes in the top 32 of all participants, for all
    # the used_freqc together:

    return best_e, opt_av_accuracies, opt_av_out, m_i

if find_optimal_subset == True and run_FLDA_LOO_k_times == True:

    # apply the defined function sort_output
    most_informative = sort_electrodes_by_weight(averaged_output)

    all_e = used_electrodes # needed for naming renaming columns to electrodes

    best_e,opt_av_acc,opt_av_out,m_i= subset_finding(most_informative, all_e)
```