BRAIN-COMPUTER INTERFACE BASED ON THE RELAXATION STATE

Final report

VALENTIN VIERGE, GANSHENG TAN, SHUHUI WANG
WEI MU, MIN WANG

WITH THE SUPPORT OF:

ANTOINE CHAILLET, HUGUES MOUNIER, LUCA GRECO

MAY, 2019

Table of Contents

1 Abstract

The link between relaxation and brain waves has been uncovered but the literature contains many contradictions about it. The quantification of the state of relaxation using EEG (Electroencephalographic) signals could give rise to interesting applications, notably in the learning of meditation. In this study we develop two machine learning models for classifying EEG signals into two states, one being Su-Soku meditation and the other a reference state corresponding to intellectual stimulation. The first model is a global classifier trained on 20 experiments, using 10, 30 or 60s samples; the second is an individual classifier trained specifically on each subject, using 1s samples. We develop the choice of features and their optimization : the main features are Mean Band Amplitude (MBA), Power spectral density (PSD) indicators in the Delta, Theta, Alpha and Beta frequency bands. Theta-Beta Ratio, Alpha-Theta Ratio and Phase Amplitude Coupling (PAC) between Theta and Gamma bands are also explored. Biomarkers such as: Skin temperature, Skin conductance, Heart rate variability, respiration serve as auxiliary tools to detect the state of meditation. The algorithms tested are: Support Vector Machine (SVM), Random Forest and k-Nearest Neighbors (KNN) using random subspaces of features. Their evaluation is done on our database of 20 experiments including the EEG signals on 21 electrodes and biomarker signals. In addition to the use of a filter, we developed a short algorithm for the filtration of artifacts due to swallowing and face movements. SVM individual classifier using Skin conductance (SC) and Heart rate variability (HRV) reaches more than 92% of accuracy for 75% of the tested subjects. The Random Forest individual classifier uses five of the most significant features among MBA (Mean Band Amplitude), ATR (Alpha Theta Ratio), TBR (Theta Beta Ratio) which are respectively the mean absolute value of the sum of complex Fourier series coefficient of a certain wave band, the ratio between MBA of alpha band and MBA of theta band, the ratio between MBA of theta band and MBA of beta band. It reaches 88% for 75% of the tested subjects. Besides, introducing PCA as a feature increases the accuracy of this classifier. The global classifier using 24 features reaches an accuracy of more than 91% for 30s samples.

2 Introduction

It has always been the dream of human beings to control the external objects through consciousness like in the science fiction world. The appearance of Brain Computer Interface made the assumption come true. BCI is defined as a combination of hardware and software that allows brain activities to control external devices or even computers. The Figure1below shows the working mechanism of BCI:

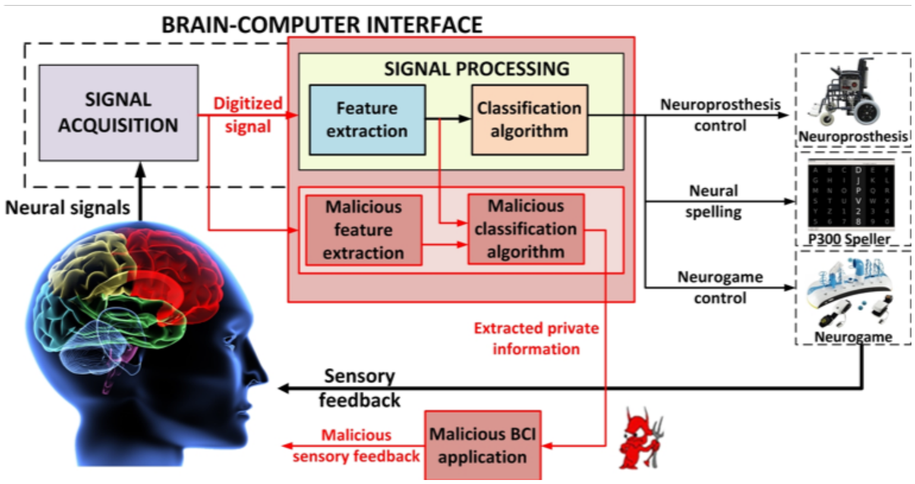


Figure 1: The working mechanism of BCI

First step is to acquire the neural signal. Brain wave activation represents the electrical activity of neurons, specifically the voltage fluctuations from ionic flow within neurons in the brain. This electrical activity is recorded via electroencephalogram (EEG), and the EEG will represent this electrical activity as waves or oscillations.

EEG spectral power and coherence estimates in the individually defined delta, theta, alpha, beta and gamma bands, which is classified by their frequencies, were used to identify and characterize brain regions involved in meditative states, in which focused internalized attention gives rise to emotionally positive experience. Table 1 below shows the frequency components of brain waves[1]:

Band	Frequency Range (Hz)	Voltage Level (μV)	Corresponding Brain Activity
Delta	0.5-4	20-200	Sleeping
Theta	4-7	< 20	Dreaming; Meditation
Alpha	8-13	30-50	Relaxation
Beta	13-30	5-10	Concentration
Gamma	above 30	5-10	Conscious perception

Table 1: Components of Brain Waves

However, it is still a challenge to change the weak brain activity into digital signals which can be processed and classified to realize the interface between the brain and the external world. Not only because of the complexity of brain activity but also the barriers lying on classification of signals.

Meditation is commonly considered to be the willful manipulation of one's state of mind and intentional self-regulation of attention. Meditation has been found in numerous studies to reduce experience of anxiety, and increase power of concentration[2].

By virtue of the positive impact of meditation on the health, recent research has begun to explore the mechanisms of meditation. One aroused great interest is the research on mechanisms of brain wave during meditation. So analysing the EEG performance during meditation is probably significant for building a useful classifier to determine the state of relaxation. In addition, the biomarkers such as: skin temperature, skin conductance, heart rate volume (HRV), respiration serve as auxiliary tools to detect the status of meditation. For example, skin temperature, it was found to increase during relaxation[3]; while there are contradictions on the skin conductance, a study by Cauthen[3] found that there is no significant change in skin conductance whereas a study by Vempati[4] showed reduction in the subjects' skin conductance.

In fact, there already exists some application based on EEG and meditation. One of American company called $MUSE^{TM}$ is an online company that acknowledged themselves as the first tool in the world that can give accurate, real-time feedback on what's happening in brain while meditating by detecting EEG. But the feedback of customers casts light on being deceptive.

Even though in the literature, there exists many varying results on the correlation between EEG and meditation. Some studies, for example, showed increased alpha or theta electroencephalographic (EEG) activity during meditation (Hebert 1977; Banquet 1973)[5] [6], whereas others show more beta or even some delta EEG present (Aftanas and Golocheikine, 2001)[7]. Some others show that using mindfulness meditation causes significant decreases in left-sided anterior alpha power in meditators compared with non-meditators(Henz and Diana,2018)[8]. The dissimilarities between studies on the brain wave of meditation may be due to the variability of meditation practices as well as methodological flaws.

Hence, our study is aimed to find out a correlation between meditation and brain waves with an adapted meditation method. Then the second step is to use machine learning to build a classifier base on EEG and biomarkers for detecting the state of meditation.

Based on many research protocols, we chose the method Su-soku (by Takahashi [8]) where the eyes are closed to avoid the influence of blink. To remedy the defect of lacking control experiment, an additional under controlled experiment (where the experiment subjects are required to solve several history questions) was taken into consideration. Subjects following strictly the protocol in the experiment, EEG data is collected on 21 electrodes which are distributed evenly on the prefrontal, frontal and temporal–central and posterior in the brain regions.

After 20 experiments, we acquired the EEG and biomarker information of our participants through the Nexus 32 Hardware set and BioT ace+ software. In our research, data pre-processing is used to remove the artifacts in EEG signals caused by swallowing or other muscles movements but because of time limit, the filtered data weren't used in final data processing.

Regarding our enormous data, we decided to use machine learning to build a classifier for the classification of the state of meditation. It consists of three parts: classification model, feature engineering and validation.

Our data is labeled into two status: meditation and no meditation, so it is a binary labels problem. The common used machine learning algorithm for classification like: Logistic Regression, Support Vector Machine (SVM), Random Forest and k-Nearest Neighbors algorithm (KNN) which are four most common supervised learning algorithm for classification. Logistic regression resembles linear regression but is used when the dependent variable is not a number, but something else (like a Yes/No response), in other words, it is used for prediction of output which is binary. So it is widely performed on the relationship between variables to get the model. Support Vector Machine (SVM) is used for both regression and classification. It is based on the conception of decision planes that define decision boundaries. A decision plane (hyperplane) is one that separates a set of objects having different class memberships. The learning of the hyperplane in SVM is done by transforming the problem using linear algebra. It performs quite well on high dimensions. Random Forest Classifier is an ensemble algorithm based on bagging bootstrap aggregation. Ensemble methods combines more than one algorithms of the same or different kind for classifying objects, for example, an ensemble of SVM, Naive Bayes or Decision Trees. The general idea is that a combination of learning models that increases the overall result is selected. Random Forest searches for the best features among a random subset of features. This results in a wide diversity that generally results in a better model. K-Nearest Neighbors algorithm (KNN) is one of the simplest classification algorithm and it is used to identify the data points that are separated into several classes to predict the classification of a new sample point. KNN is a non-parametric, lazy learning algorithm. It classifies new cases based on a similarity measure (e.g. distance functions). KNN works well with a small number of input variables p, but struggles when the number of inputs is very large [9] [10].

In our data processing, Logistic Regression, Support Vector Machine (SVM) and Random Forest were used for individual classifier. k-Nearest Neighbors algorithm (KNN) is applied to global classifier.

Feature engineering is the process of using domain knowledge of the data to create features that improve the performance of machine learning algorithms. Feature engineering is fundamental to the application of machine learning, and is both difficult and expensive. The performance of predictive models depends on the quality of features.

Basically we consider the four bands of EEG signals in different electrodes and the biomarkers as features for model training. In addition, new feature, phase amplitude coupling (PAC), which indicates the correlation between the two signals oscillations (signals with different frequencies) were analysed in our feature engineering. The accuracy of Logistic Regression model is less than 70% for individual, however, for global classification, its accuracy is around 80%. The result provides us with the possibility for improving the accuracy of classifier by adding PCA into the model.

After building the classification model, there is always a need to validate the stability of the machine learning model. Generally in training process, a numerical estimate of the difference in predicted and original responses is done, also called the training error. However, this only gives us an idea about how well our model does on data used to train it. Now it is possible that the model is underfitting or overfitting the data. So, the problem with this evaluation technique is that it does not give an indication of how well the learner will generalize to an independent/ unseen data set. Getting this idea about our model is known as Cross Validation[10]. The most used cross validation is called: K Fold cross validation. In K Fold cross validation the data is divided into k subsets and is repeated k times, such that each time, one of the k subsets is used as the test set/ validation set and the other k-1 subsets are put together to form a training set. The final accuracy is obtained after averaging, therefore this significantly reduces the variance of accuracy.

It is found that there exists strong correlation between SC (skin conductance) and HRV (heart rate variability which is related to heart rate) under meditation and non-meditation statues by linear regression analysis. So SVM combined with cross-validation with 10 folds by using SC and heart rate variability (HRV) is implemented. The accuracy for one subject during training step reaches 97.2 %, and the median of accuracy in testing step for 14 subjects is 100 %.

As for individual classifier−−Random Forest classifier based on EEG signals, it is found that the there does not exist a generally important feature for everybody. So for every individual classifier, the importance of features are calculated, then we take the five most important features into the model. Without considering PAC, the accuracy of Random Forest classifier using the four most important features is around 90%, for some individuals, its accuracy reaches to to 95%. After adding PAC, the accuracy can increase $1\% - 10\%$ which still needs further research because we just used PAC for two subjects' data processing.

Furthermore, a global classifier is considered. We are searching for a classifier which is adapted for every individual. By using average amplitude of alpha bands as feature, the predicating on non-meditation phase is completely correct, while the predicating accuracy is quite different from individuals. The problems can come from the artifacts in the signals. Because the accuracy of the KNN algorithm can be severely degraded by the presence of (i) noisy data, (ii) irrelevant features, and (iii) non-consistency of feature scales with their importance. Thus, for an efficient KNN based classification, (i) digital filtering has been used to reduce the noise, and (ii) the feature engineering should be done to get the features' importance which serves to reduce the computational burden on the KNN classifier [11].

Because of the time limit, we did not use the filtered data in data processing, PAC is not completely implemented for global classifiers. But it is a funny adventure to explore what is happening in our head and put our acquired knowledge into practice.

3 Experiment

3.1 Experiment Subjects

A total of 19 undergraduate students and 1 professor have been recruited to participate in this study. They consist of 12 males and 8 females. Except for one participant (52 years old), the rest of the subjects were aged between 19 and 22 years (average = 20.6 y, S.D. = 0.8 y). There were 13 French, 3 Chinese, 2 Moroccan, 1 Polish and 1 Chilean. They are all free from cardiac or pulmonary diseases that could cause autonomic nervous system dysfunctions. 2 of them were regular meditators practicing at least once a week, 1 of them was a very regular meditator practicing Mindfulness meditation once a day, and 1 of them was a very experienced meditator, practicing Taoist meditation and Buddhist meditation once or twice per day. The rest were almost new to any kind of meditation. The analysis has been performed without distinction between the subjects.

The detailed information about the subjects are in the appendices. In order to classify and process the data more conveniently, each subject has a unique code composed of 4 letters. In this report, we will also use the code to refer to the subjects.

3.2 Experiment Equipment

Based on the equipment of Lab L2S in Centralesupelec, we used Nexus 32 Hardware set and BioTrace+ software for acquiring the signal information. EEG signals were acquired through a EEG head cap with 21 electrodes. Skin temperature, skin conductance, heart rate volume (HRV), respiration were acquired by corresponding sensors. All EEG and biomarker signals are acquired and exported through $BioTrace+$ software.

3.3 Experimental Protocol

The following presents the protocol used for these 16 experiments, which is the result of several months of evolution. It has been designed with a view to obtain two different states of mind, therefore, it is trackable and easy to reach for the subject.

3.3.1 Specifications

Type of Meditation: Many different types of meditation exist, with very various procedures and objectives. Hence, the resulting state of mind – and the associated brainwaves - cannot be expected to be the same between two subjects performing different types of meditation. For our experiment, we chose Su-Soku meditation, a type of Zen meditation meant for relaxing. It consists in breathing largely and calmly, and counting every breath taken. There are two main reasons to this choice: first, Su-Soku meditation is particularly easy to learn and perform for beginners, which most of our subjects were. Then, Su-Soku meditation is trackable since the count of respirations, simultaneously done by the subject and the respiration sensor, gives a reliable information on how focused the subject was.

Baseline: For the means of comparison with meditation, a baseline "reference state" was needed. Simply letting the test subject sit and rest did not seem to be the best way to obtain a clear distinction with meditation. A way to simulate this and keep the subject in a thinking mode would be to ask them to solve some - math - problems. However, a major constraint of EEG measures is that face muscles movements generate a lot of artifacts that can obliterate the readability of the signal. Therefore, we cannot ask the subject to speak their response to the problem; and asking them to solve problems in their mind reduces the trackability of their state as well the interest and energy they might be willing to put in the process.

So, to make this more appealing, we came up with an idea: the problems are put in the form of a challenging story, which is read to the subject along all the phase. At some points in the story, a problem is presented to them. They are given a fixed amount of time to solve it, after what the story continues no matter what. They have to remember their answer and give only when the story comes to an end. In this report, this phase will be further referenced to as "Story phase".

In our story "Escape dungeon", the participant plays the role of a prince/princess who has to go through 5 rooms containing 5 enigmas in order to save a princess/prince. The story lasts 8 minutes, among which 5*45s of time given for reflection. The enigmas call for memory, spatial vision, mental computation and logic. The text is attached in appendix at the end of this report.

Feedback: A feedback questionnaire was submitted to the test subjects at the end of each phase, in order to collect their impressions and some information on their state of mind. This is an important step as relaxation is very hard to quantify and cannot be defined other than subjectively. The questionnaire can be viewed in appendix.

3.3.2 Procedure

The following table shows the exact process followed during the experiment from the beginning to the end. In order to limit the influence of the order of phases on the data collected, half of the experiments were performed beginning with the Meditation phase, and the other half beginning with the Story phase. In the following table, the Story phase comes first.

Table 2: Process of the whole experiment

Test subject	**Experimenter**
Enter the Room.	Greet, invite to seat. Explain how EEG works and what are the 2 phases.
Sit and listen.	Put on the subject helmet + sensors (heart rate, skin temperature, skin conductance, respiration rate).
	Calibration.
Listen to the instructions.	Explain the 1st phase in detail. Ask to sit as still as possible and to remain eyes closed.
Closes the eyes.	Beginning of 1st phase.
1st phase: Story phase(8 min)	
Take a break.	End of 1st phase.
Answer the questions.	Ask how many questions they solved and give them the answers. Submit story questionnaire.
Listen to the instructions.	Detail 2nd phase. Ask to sit as still as possible and to remain eyes closed.
Closes the eyes.	Beginning of 2nd phase.
2nd phase: Meditation phase (10 min)	
Take a break.	End of 1st phase.
Answer the questions.	Ask how many breaths they counted. Submit meditation questionnaire
Leave.	Remove equipment and thank the subject goodbye.

3.3.3 Story Phase

Duration: 8mins

Instructions given to test subjects: This phase is designed to be a "control phase", it serves as a baseline to be compared with meditation. Please listen carefully. I am going to read you a story. What's specific about this story is that you're going to participate. It is a sort of escape game, where you have to go through different chambers and solve enigmas to move on to the next chamber. You should solve, or try to solve, the enigmas in your head. I am going to repeat each enigma and let you some time, about 40s, to solve it. If one of them is too difficult and you can't solve it during that time, it's okay, please just focus on the continuing of the story. No matter what happens the story keeps going on, and it's important that you don't lose track of the story. I shall remind you that you need to close your eyes during the whole phase and move as little as possible.

3.3.4 Meditation Phase

Duration: 10mins

Instructions given to test subjects: This is the "core exercise" of the experiment. We are going to ask you to meditate, but not any type of meditation : you're going to perform Su-soku meditation. Su-soku meditation is based on respiration counting. Close your eyes, breath very slowly and largely. Count when you inspire (3s), then expire (4s). Try to empty your mind and focus only on breathing. Free your muscles from any tension. Count to 100. You should count every breathing, even if you think it's not deep enough. We might stop you before you reach 100 (depends on the subject: see in practice). If we have not stopped you by touching your shoulder before you reached 100, please start counting again from 1.

4 Data Acquisition and Pre-processing

4.1 Data Acquisition

All the data is collected using a 10-20 EEG cap and sensors for Skin conductance, Heart rate, Skin temperature and Respiration rate, associated to the NEXUS 32 hardware and Biotrace+ software. During the recording, data collected by NEXUS 32 is transmitted via BlueTooth connection to the computer, where it can be viewed in real time and stored in Biotrace+. After each experiment, the data is exported mainly in the form of a $.txt$ file to be processed in R and $.mat$ file to be processed in matlab.

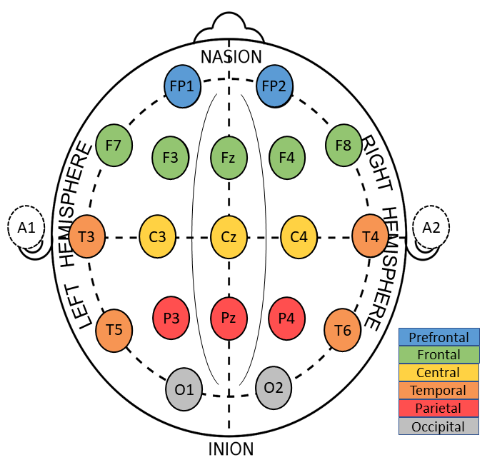


Figure 2: The 10-20 configuration with the corresponding brain areas

4.2 Digital Filtering

Because original EEG signals contain huge low frequency fluctuations and many high frequency artifacts including the 50Hz peak from the electricity grid, they are filtered before exportation. This is done thanks to a digital filter implemented in Biotrace+. Except for performing phase-amplitude coupling (PAC) analysis, this filter is a 3rd order Butterworth filter from 1Hz to 45Hz shown in Figure 3. For phase-amplitude coupling analysis more information on high frequencies was needed, so we used the same type of filter from 1Hz to 80Hz.

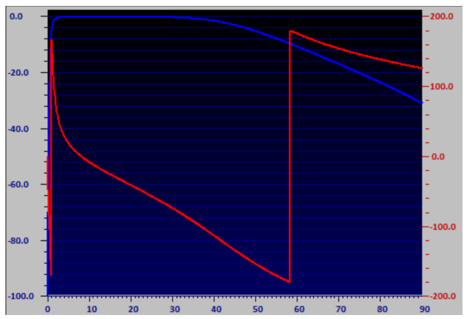


Figure 3: Frequency response of the filter (phase in red, amplitude in blue)

After the digital filter implemented in Biotrace+, the low frequency fluctuations and the 50Hz from the power grid are removed. However, the signals are still not clean. Artifacts such as swallowing and muscle movement around the eyes, which are between 1Hz and 45Hz, are still included in the data. These artifacts can affect the frequency analysis, so further cleaning is needed to remove them.

4.3 Artifacts Cleaning

As said before, swallowing and muscle movements are clearly some nuisances in investigations of EEG. These movements that occur in synchrony with the events that we want to observe introduce unwelcome artifacts into the data.

4.3.1 Artifacts Identification

In order to figure out the pattern of swallowing and muscle movement in EEG signals, we took note of the moment where the subject moved during the recording. The confrontation of these timings with those of artifacts found in the signals revealed that swallowing and muscle movement mainly causes some fluctuations in three electrode: $F7$, $F4$ and $F8$. Here is one example shown in Figure 4 that shows the EEG artifacts elicited by swallowing or muscle movement. Red frames indicate the moments where an obvious movement of the subject was observed during the experiment.

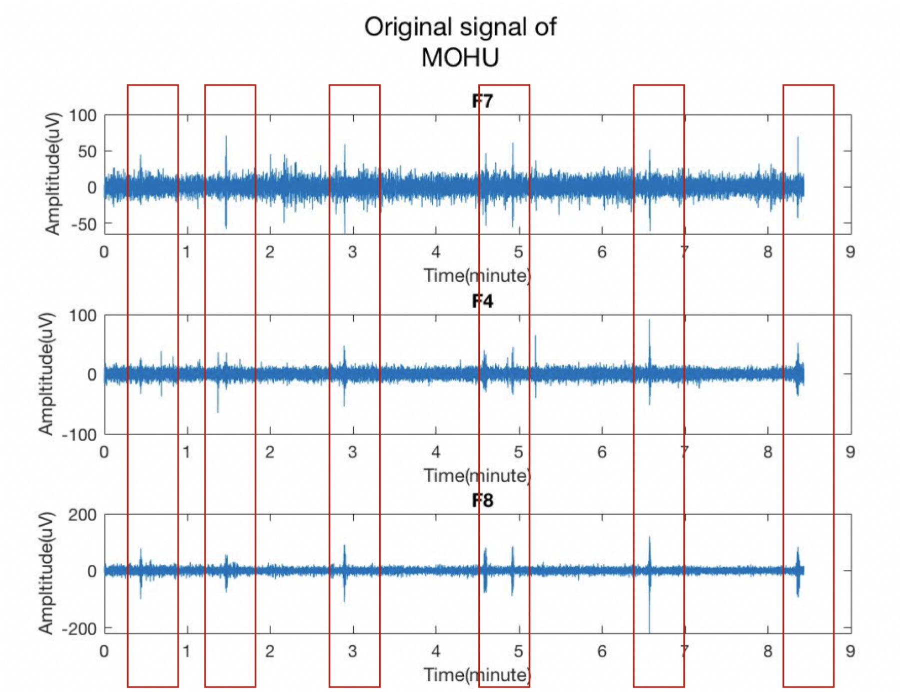


Figure 4: The original EEG signal of subject MOHU

After observing all of the subjects, two features are attained:

- While normal EEG peak-to-peak amplitude is around 40 μV, this amplitude ranges from 80 to 300 μV during artifacts;
- The swallowing process usually lasts 3-4 seconds. While for muscle movements, it depends on which kind of movement it is.

4.3.2 Artifacts Elimination

The whole process consists of 4 steps:

1. Moving Average This procedure is to find the interval where it exists the high amplitude so that it enables us to detect the appearance of artifacts. To do so, we must divide the signals in frame of length L. Overlap rate r was introduced for moving average to adjust the overlap between neighbor frames. so the moving average is attained by:

$$Avg_i = \frac{\sqrt{\sum x^2(t)}}{L} \mathbb{1}_{[(i-1)\times L\times(1-r), i\times L\times(1-r)]}$$

   In this way, we can get the moving average which resembles the figure5:

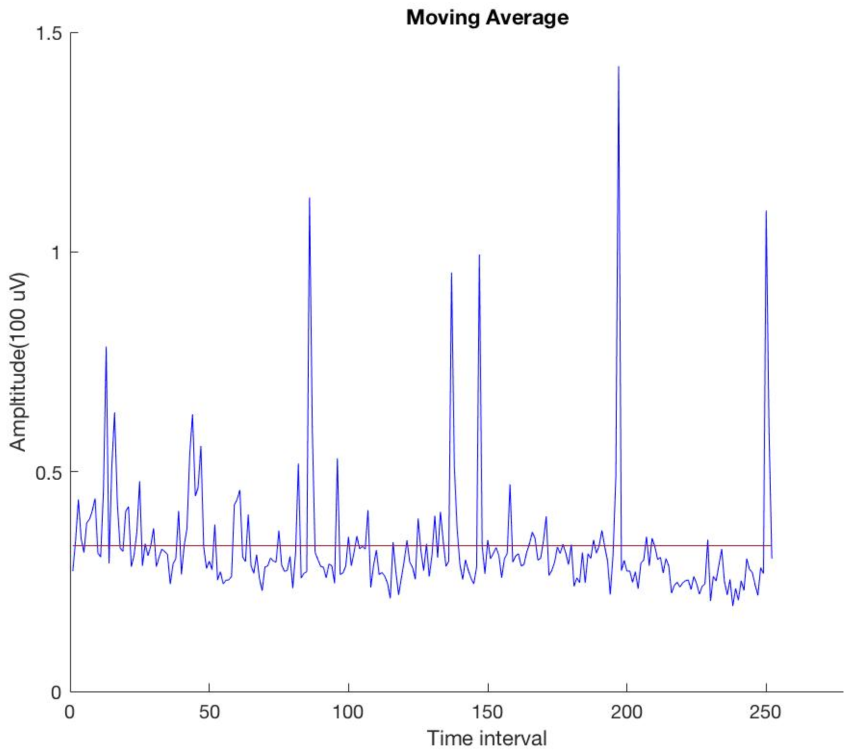


Figure 5: The amplitude moving average: L=512 samples, overlap=0

From the Figure above, we can easily find where is the artifacts through the peak.

2. Then based on moving average, we use the function $findpeaks$ in matlab to return the indices at which the peaks occur. Besides, we set the $MinPeakHeight = average(moving - average) \times 750$ which means that it restricts $findpeaks$ to return only those peaks higher than 'MinPeakHeight'. $average(moving - average)$ is the average of the data that we got in moving average. In this case, we got the indices of peaks in the moving average signal, as shown in Figure6. After this, we change the indices to the indices of peaks in original signals(shown in Figure7):

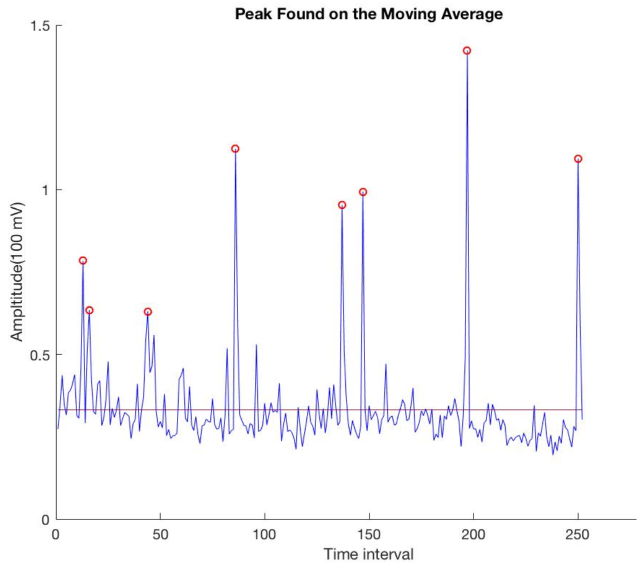


Figure 6: Peaks on amplitude moving average

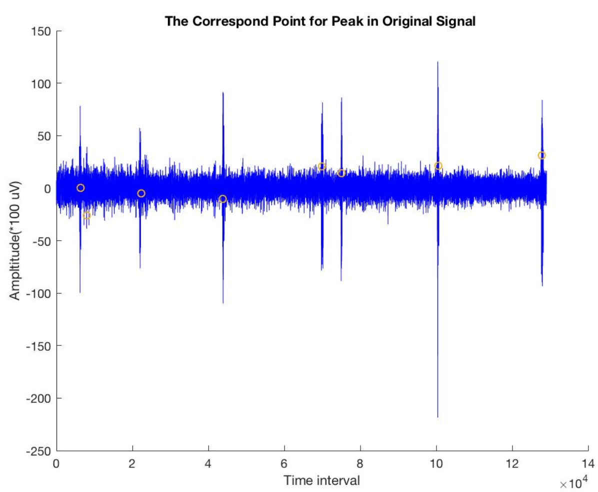


Figure 7: Peaks on original signal

Until now, we succeeded to find the indices of artifacts.

3. Now in this step, we take the three main electrodes($F7$, $F4$, $F8$) into consideration. For each electrode, the indices of peaks are obtained. Then we made an union of all indices as the final indices to remove for all electrodes, so that we can keep the synchronization of all the electrodes.

4. Finally, once the indices are fixed, for each indices, we remove the 3 seconds' intervals around the indices corresponding in the signal.Therefore we got the signals before and after the filtration((Figure8 Figure9):

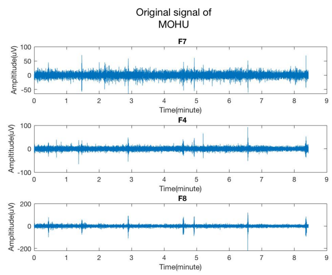


Figure 8: The original signal of MOHU

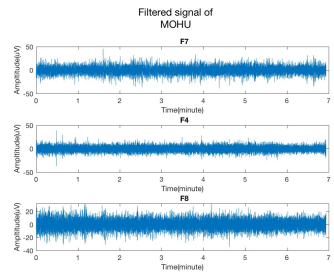


Figure 9: The filtered signal of MOHU

5 General Tendencies

The goal of this section is to present the investigation work performed in order to find interesting variables that describe the data and allow the differentiate the Meditation and the Story state. This research gives an overview of how the differences between these two states can be quantified, which is a preliminary step to the definition of classification features.

5.1 First Observations from EEG Data

5.1.1 Definition of Mean Band Power

The first idea is to compute the Power spectral density (PSD) of the whole signal and to see whether the two phases exhibit different power distributions. We call power distribution the distribution of power between the 4 frequency bands Delta, Theta, Alpha and Beta (band power is defined as the integral of PSD on that frequency band). From the following figures we can have a first look at this distribution. In this example, the subject exhibits an increase in Alpha power during Meditation compared to Story.

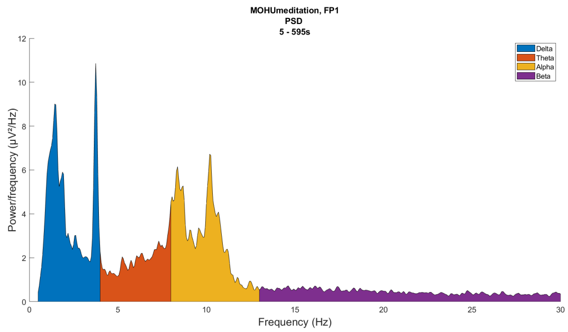


Figure 10: PSD of a full-length FP1 signal during Meditation

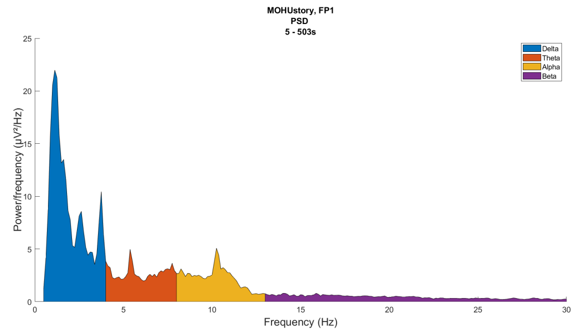


Figure 11: PSD of a full-length FP1 signal during Story

From the computation of PSD in each electrode we can define Mean Band Power (MBP) of each brain area as the average band power of all the electrodes composing that area. Hence each of the 6 areas is characterized by Delta MBP, Theta MBP, Alpha MBP and Beta MBP. Then the relative variation of MBP during Meditation compared to Story can be computed.

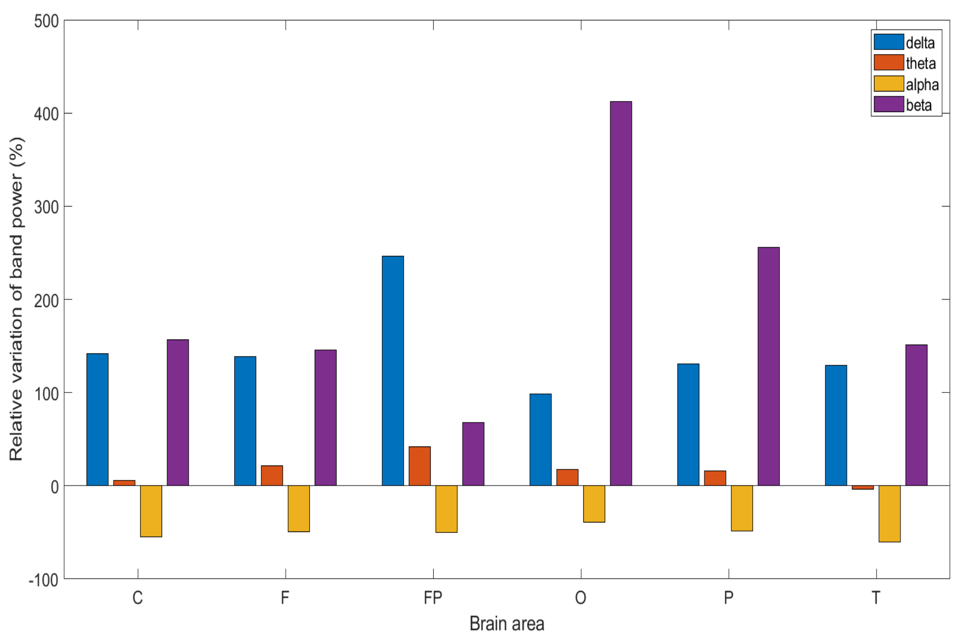


Figure 12: MBP variation in the 4 frequency bands of the 6 brain areas, for one subject

5.1.2 Statistical Results

The following boxplots show the relative variations of MBP during Meditation compared to Story for 15 subjects. To be able to differentiate the two states thanks to these variations, we expect to spot some boxes situated completely above/below zero, denoting a pair [band,area] with an increase/decrease of MBP among at least 75% of the subject, which can be considered as significant. Such boxes (framed in blue) can indeed be observed, which is very encouraging regarding the possibility to discriminate the two state with MBP.

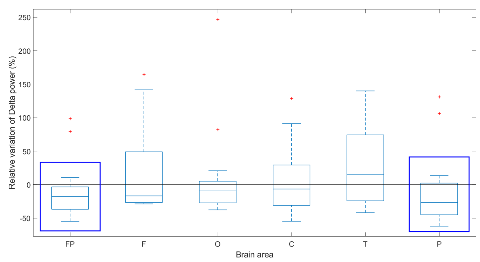


Figure 13: Boxplot of Delta MBP variation for 15 subjects

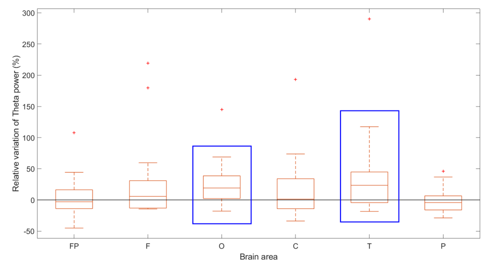


Figure 14: Boxplot of Theta MBP variation for 15 subjects

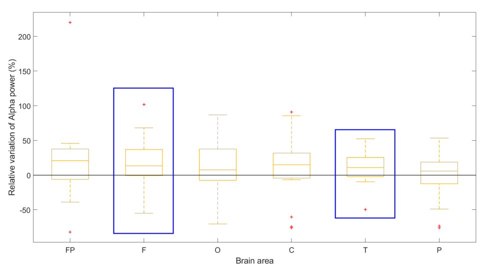


Figure 15: Boxplot of Alpha MBP variation for 15 subjects

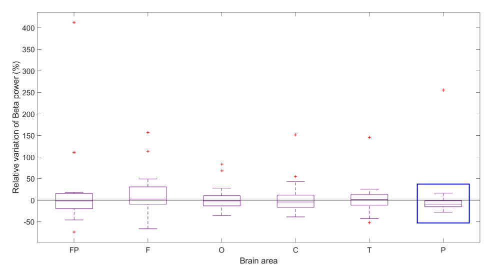


Figure 16: Boxplot of Beta MBP variation for 15 subjects

5.1.3 Time Consistency

Until now, MBP has been computed on the full-length signals. We may ask ourselves if the significant MBP variations observed are valid on a smaller time scale. This presents a high interest in the construction of a real-time classifier : indeed we would like to able to classify the state based on an acquisition of 1min or less. Figure 17 shows the FP1 mean band power computed on windows of 60s noted MBP60, with an overlap of 0.9. The lines show the MBP over the full length signal, so they are the averages of moving curves. We can see that on windows of 60s, the MPB variations fluctuate a lot around their mean values, especially in the Delta and Theta band here.

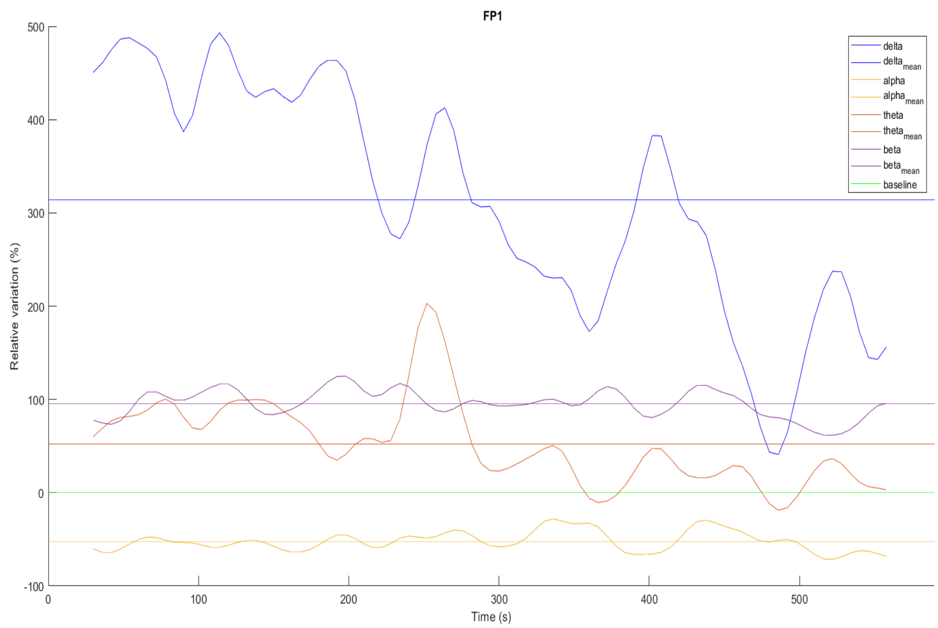


Figure 17: Moving average on 60s of mean band power (MBP60), overlap = 0.9

5.2 Correlation between EEG signals and biomarkers

In this section, we explain the process to transform the $.txt$-form recording file into analysable R-language data frame. This process is aimed at the construction of a clear input data set for classification which we will present in section 7. In addition, we also present the correlation between different signals.

5.2.1 Data Preparation

For statistical analysis, the data is directly exported from the BioTrace+ software in the form of a text file. This txt file contains data of 21 EEG channels and 4 sensors of biomarkers: Skin Temperature (ST), Skin Conductance (SC), Respiration Rate (Resp), Heart Rate (HR).

The EEG signals are filtered by a 3rd-order band pass Butterworth filter from 1 Hz to 45 Hz. The text file exported is a matrix delimited by comma. The sample rate of EEG signals is 256 SPS (sample per second) and the sample rate for biomarkers is 32 SPS. So each line of the matrix is the data acquired in that 1/256 second for the EEG signals and 1/32 second for other biomarkers. Because the simple rate are not the same, the data from biomarkers are repeated 8 times in order to have the same number of samples as EEG signals. Each column of the matrix is one electrode of EEG or one sensor of biomarkers.
Units: EEG signals are in micro-Volts (μV), ST in Celsius degrees, SC in micro-Siemens (μS), HR and Resp are in times per minute.

To do a statistical analysis, we set a time length standard for all the subjects, which is that we take 450 seconds during story phase and 580 seconds during meditation. In order to minimize the effect of physical exercises before the experiment as well as the impact of dismantling devices at the end of each phase, we take 450 seconds in the middle of story phase and cut the same length of the beginning and at the ending. We do the same on meditation.

5.2.2 Definition of Mean Band Amplitude and RMSSD

For EEG signals, we suppose that if we divide the signals in frames then they can be decomposed as sum of sinusoidals. First thing we do is to do a time-frequency analysis. For each electrode, we divide the signal in frames of length L, indexed by n, where it can be considered stationary. Afterwards, we compute the coefficients of the Fourier series of each frame, indexed by k. We denote the signal as $x(t)$, and the window function $w(t)$, the step a.

$$g[k,n] = \frac{1}{L}\int_{\mathbb{R}} x(t)w_L(t-na)exp(-i2\pi kt/L)dt$$

In practice, we take $L = 512\ samples$, which represents 2 second, $a = L/2$ meaning the step is one second. $g[k,n]$ is a complex matrix, we calculate $\sum_{k\ in\ the\ wave\ band} \frac{|g[k,n]|}{\#k}$ where k is the number of k, and denote the sum **Mean Band Amplitude** (MBA). After this processing, each electrode has four MBAs, respectively AlphaMBA, BetaMBA, ThetaMBA, DeltaMBA for each second.

Then, we want that the traditional biomarkers have the same length as MBA, that's to say, there is one value for each second. So the value per second of Heart Rate, Respiration Rate, Skin Temperature, Skin Conductance are obtained by interpolating raw data points. In practice, this is done by using $approx$ function in R. This function returns a list with a desired length which linearly interpolate given data points.

Finally, The well-known definition of HRV is the RMSSD (Root mean square successive difference) of R-R interval[12]. This is illustrated by the figure18 below. Supposing HR is constant during one second, the R-R interval can be represented by $\frac{60}{heartrate}$.

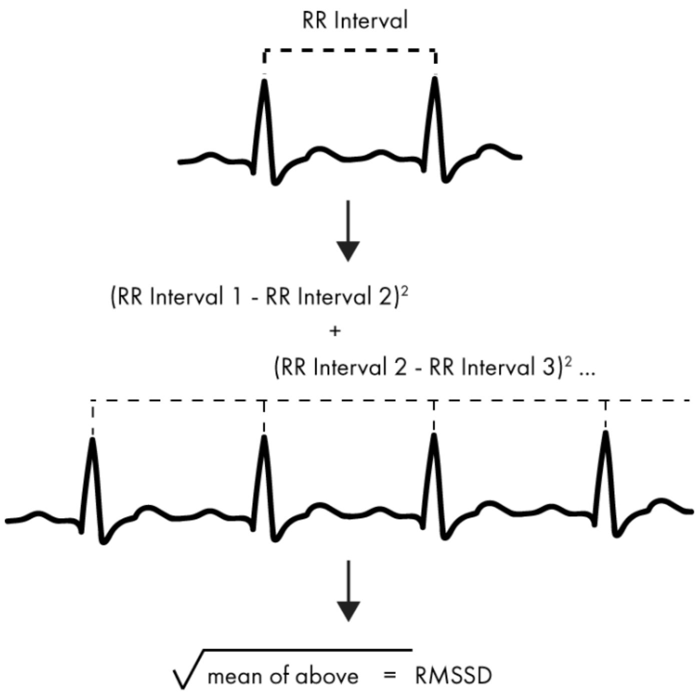


Figure 18: Calculation of RMSSD [12]

Thus, we define heart rate variability as MSSD (mean square successive difference) of heart rate. Practically, we build HRV from pre-processed heart rate (BPM). After building the matrix df, we have a heart rate value for each second, than we suppose that HRV in the first second is 0, and we calculate the HRV of resting time by using MSSD of heart rate. At the end, we add the label for each sample, so we will have a input matrix looking like this:

	FP1AlphaMBA	⋯	FP1DeltaMBA	⋯ MBA of other electrode ⋯	conventional biomarkers	label (0 for story phase and 1 for meditation)
sample1(1 sec of story phase)	350.63	⋯	338.32	⋯	⋯	0
⋮	⋮	⋮	⋮	⋮	⋮	⋮
sample451(1 sec of MEDITATION)	320.23	⋯	309.15	⋯	⋯	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 3: df matrix example

5.2.3 Statistical Results

In order to see the correlation between biomarkers more clearly, we present one example of the scatter plot (that of subject MOHU) shown in the figure 19. In this example, we can see the relation among all the traditional biomarkers: heart rate (HR), skin conductance (SC), heart rate variability (HRV), respiration rate (Resp), skin temperature (ST). As a comparison, we also include an EEG biomarker: the thetaMBA of FP1 electrode (FP1ThetaMean). In the figure, each point is a sample of one seconds. The x axe and y axe are the corresponding biomarkers.

As can be seen from the example, the correlation among SC, HRV, ST is clear and obvious. The points are nearly linearly distributed. However, there are no clear correlation of HR, RESP, and FP1ThetaMean with other biomarkers.

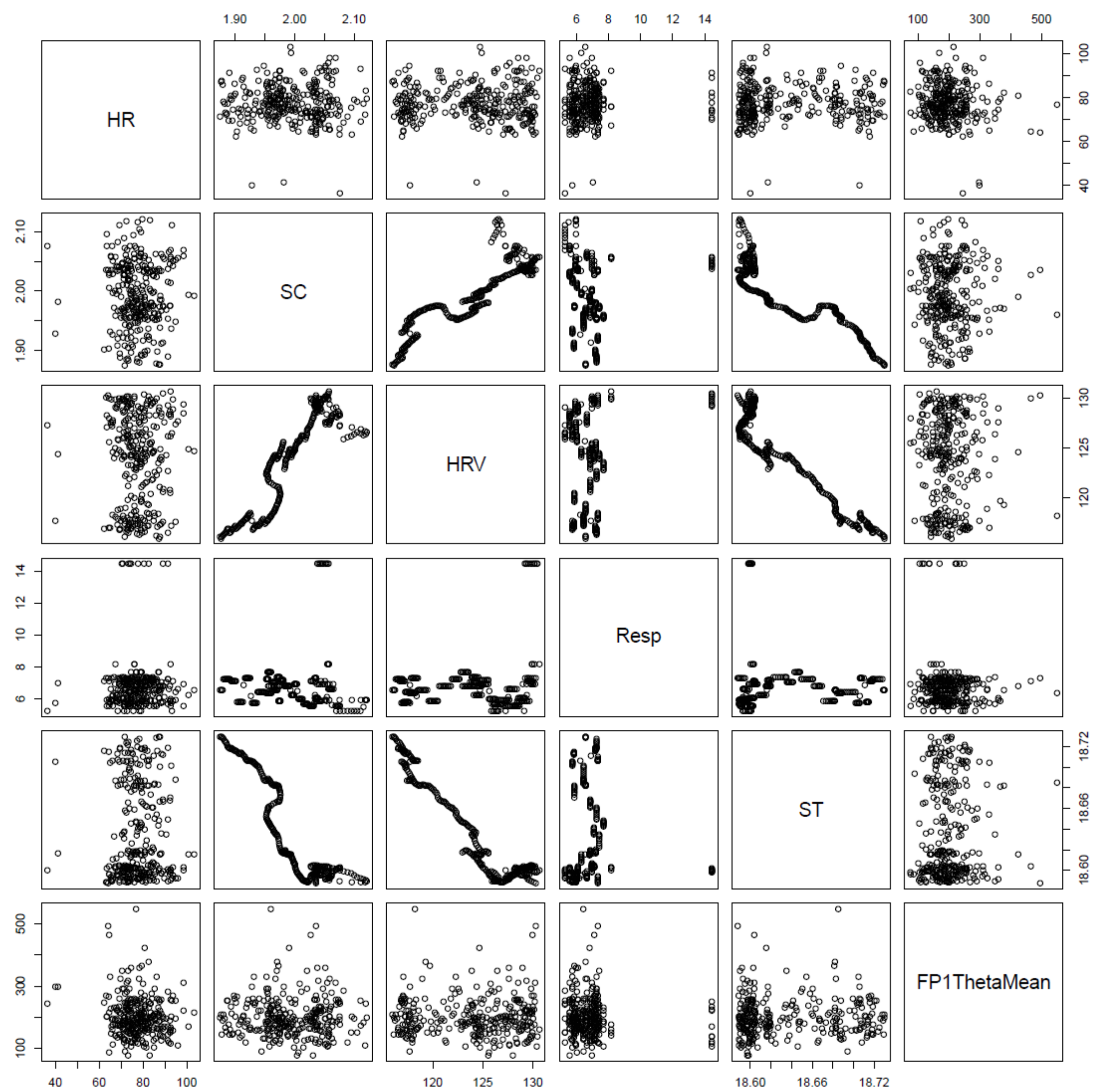


Figure 19: An Example of Scatter Plot of Biomarker (MOHU, Meditation Phase)

Because skin temperature (ST), skin conductance (SC), heart rate variability (HRV) are strongly correlated between each other, so in order to reduce the scale of calculation, when we calculate the correlation between traditional biomarkers and EEG signals, we just calculate the correlation between HRV and all the MBA of each electrode. In other words, we use HRV as a representative for traditional biomarkers. If one MBA is correlated with HRV, it is correlated with SC and ST too.

Figure20 is the scatter plot of subject *MOHU*, where we can seethe skin conductance increases rapidly with heart rate variability during meditation phase, while during story phase there is a slower increase. We will hence do the regression separately. These three line with different colors represent one regression during meditation and two regressions during story phase where one have set a threshold for HRV value. When we calculate the linear regression of biomarkers respectively during meditation and story phase, we have the p-values which are quantitative values to evaluate how strong the biomarkers are correlated with each other. The smaller the p-value is, the stronger the correlation is. We saved the p-value between HRV and EEG band power for all the test subjects and for both the story phase and the meditation phase. Then we averaged the p-values among all the subjects to see which EEG signals are more correlated with HRV. The top 5 are shown in the table below.

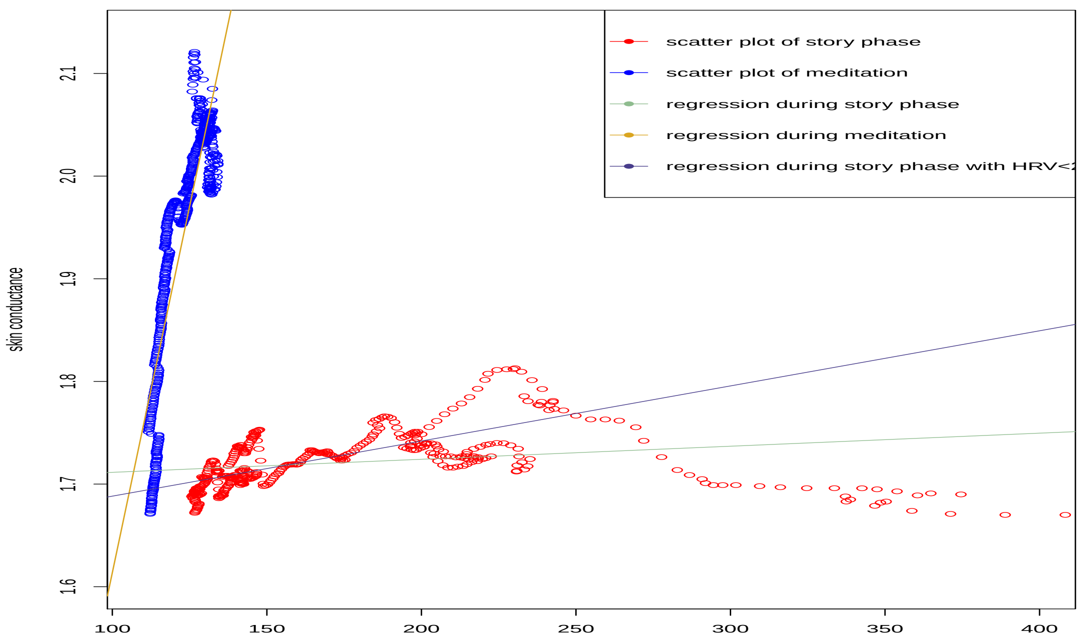


Figure 20: Regression of HRV-SC

Electrode	EEG band	Phase of Experiment	p-value
T4	Beta	Story	0.0329
O1	Delta	Meditation	0.0454
T6	Delta	Meditation	0.0540
T4	Beta	Meditation	0.0618
O1	Alpha	Meditation	0.0714

Table 4: P-values of How EEG Bands Correlate with HRV

5.3 Phase-Amplitude Coupling

Phase-amplitude coupling (PAC) is the modulation of the phase of low frequency components according to the amplitude of high frequency components.

In 2013, Lisman [13] pointed out that the exchange of information between dierent brain regions is actually completed by the orderly neural coding mechanism between the two signals oscillations. Overall, there are three ways to explore the relationships between different frequency signals. Amplitude-amplitude coupling, phase-phase coupling, and phase-amplitude coupling. them, phase-amplitude coupling, especially between low frequency and high frequency, may be related to the process of long-time memory and the integration of information exchange between neurons.

The higher frequency of EEG is in charge of the rapid information processing of local brain region, while low frequency is not only related to external sensory input and motion events, but also to internal learning and memory cognitive drive. The phase-amplitude coupling of theta and gamma may be related to the communication and change of information between neurons in learning and other tasks.

In a word, phase-amplitude coupling plays an important role in revealing the mechanism of information transmission of the cerebral cortex in different work tasks.

5.3.1 Method to Calculate PAC

In fact, what we will do is find difference between meditation state and story state, in other words, new features for our classifier, either our global classifier or the individual one.

Three methods have been introduced in the paper[14], including the envelope-to-signal measure (ESC), the modulation index method (MI) and the cross-frequency coherence method (CFC).

The method we choose is the modulation index method (MI).Next we will briefly introduce the principle of this method[14].

$$Z_{fph,famp} = A_{famp}(t).e^{i\beta_{fph}(t)}$$

A_{famp} means the high frequency amplitude envelope values.And β_{fph} means the low frequency signal's instantaneous phase. Z is a complex valued. All variables are obtained from the composite signal. The MI value,that is, PAC value without normalization, is calculated as the absolute average value of Z.

$$MI_{fph,famp} = |average(Z_{fph,famp}(t))|$$

We research basically the relation between theta (4hz-8hz) and gamma (32hz-80hz) ,with data filtered by 80hz.

5.3.2 Statistical Results

In theories, overall, the PAC value of meditation state will be smaller. For example: when thinking, our brain is more active, the modulation effect should be more obvious. However for now, when analyzing our data, the result is not ideal.

However, when the results were processed, we found that the difference between people is very large, not only the difference between the values, but also the difference in the active level of the brain.

For some people, the PAC value of meditation state is higher. Overall, there are three types, higher PAC value in meditation state (see Figure21), higher in story state (see Figure22), and similar between two states (see Figure23).

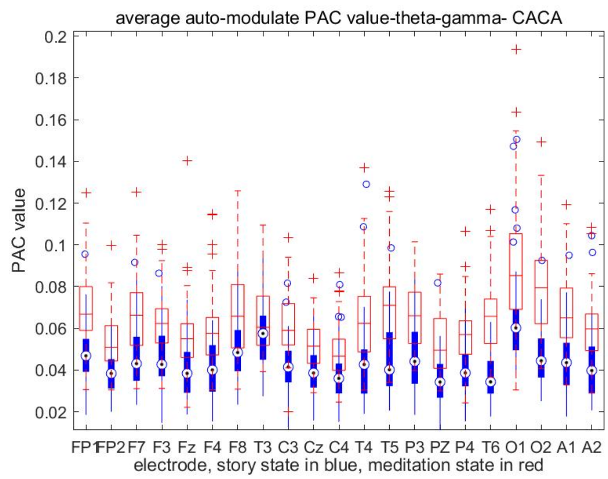


Figure 21: one type of PAC result - higher PAC value in meditation state

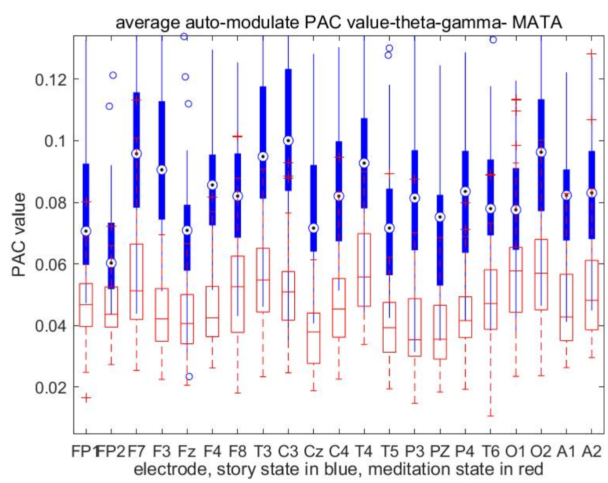


Figure 22: one type of PAC result - lower PAC value in meditation state

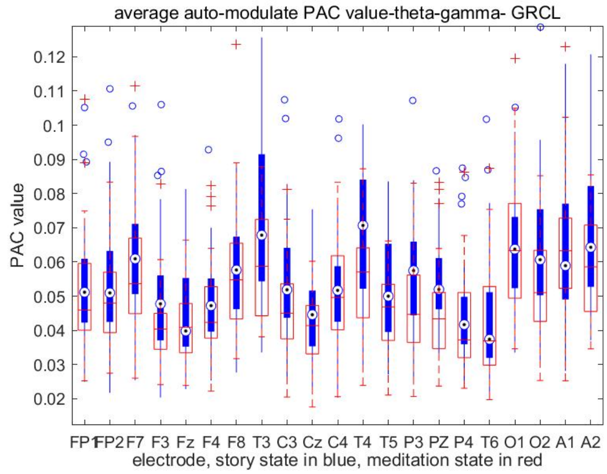


Figure 23: one type of PAC result - similar PAC value in two states

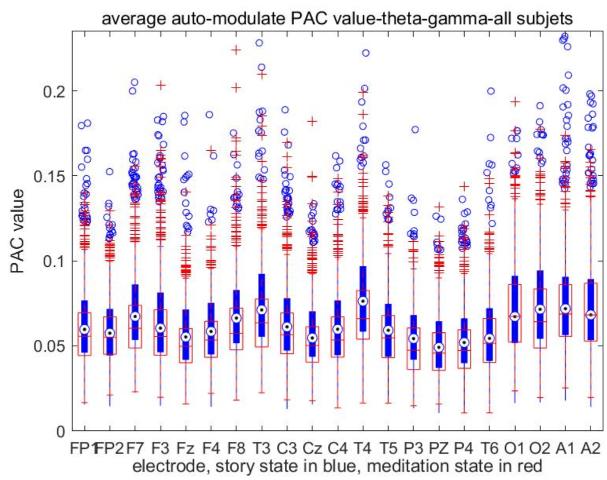


Figure 24: ALL subjects' result

There is almost no obvious tendency in the overall processing. It is verified by logistic regression algorithm that only the PAC value is used as a reference, and the accuracy of the global classifier is about 60%. In this way, it seems better to put PAC value as feature of individual classifier.

5.3.3 Logistic Regression Method

Logistic regression algorithm has been used to verify the quality of this new feature. It is simple, efficient and easy to interpret. It has fast calculation speed and easy parallelism. It is very suitable for large-scale data. It is mainly suitable for solving linear separable problems, especially the classification problem of two states. However, it is easy to underfit. In most cases, manual engineering is required.

Here we will introduce briefly the theory of logistic regression algorithm.

As shown in the figure below, it gives a line in order to divide the whole into two parts. It is an example of linear regression. The difference between linear regression and logistic regression is an activation function, for example, sigmoid.

$$Sig = \frac{1}{1+e^{-h}}$$

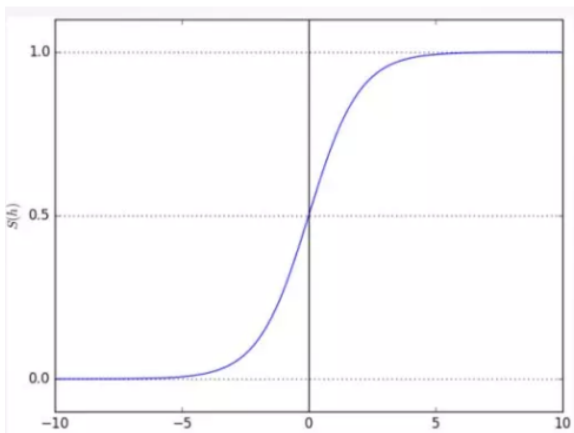


Figure 25: sigmoid function

The most important thing is to minimize the loss function, where m is the number of training sample,$h_\theta(x)$ is the value predicted by parameters θ and x, y is the value in original training sample, i means ith sample. Python is a very convenient language for machine learning. It has a number of file libraries. The logistic regression model exists in the 'sklearn' library. Therefore, it is feasible to choose logistic regression to verify the feasibility of this feature quantity.

$$J(\theta) = -\frac{1}{m}.[\sum_{i=1}^{m} y^{(i)} \log h_\theta(x^{(i)}) + (1-y^{(i)}). \log(1-h_\theta(x^{(i)}))]$$

What's more, three definitions would be introduced, which would be used in order to quantify the quality of the classification.

AUC (Area Under the Curve), is defined as the area under the ROC curve, and it is obvious that the value of this area will not be greater than 1. The value of AUC is generally between 0.5 and 1. As a numerical value, the classifier corresponding to the larger AUC works better.

Recall, in other word, true positive rate (TPR), is defined:

$$Recall = TPR = \frac{TP}{TP + FN}$$

where TP is 'True Positive', FN is 'False Negative'. Accuracy is our most common metric, and it's easy to understand, that is, the number of samples being paired divided by the number of all samples. Generally speaking, the higher the correct rate, the better the classifier.However, it is far from scientific evaluate an algorithm model simply by accuracy.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

5.3.4 Evaluation

It is worth noting that this part of the assessment is only to evaluate whether the PAC can be used as the eigenvalue. The final thing we have to do is apply the PAC as a feature value to our classifier.

As shown below, For most subjects, the accuracy is around 80%. That means PAC value is quite good feature for most subjects.

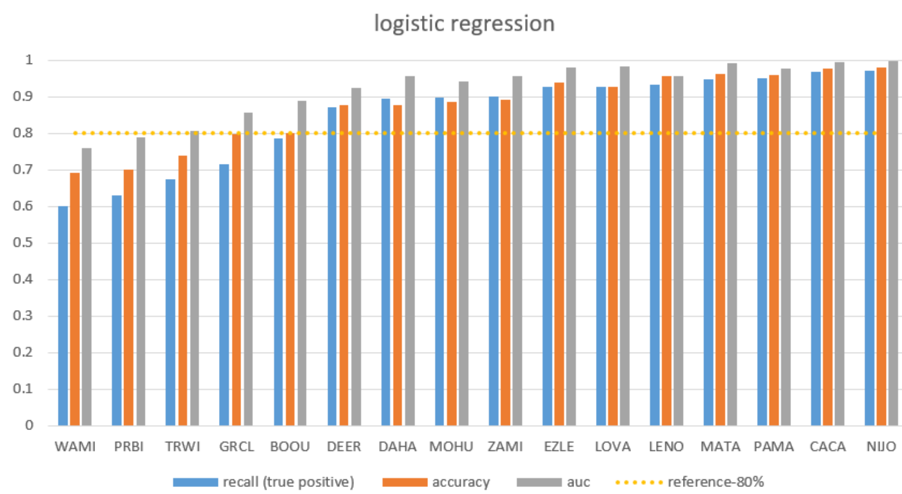


Figure 26: result of each subject

At the same time, if we use all the data, even after calibration the recall rate is only 50%. It is inappropriate to use PAC value as global feature. We will verify the result in our global classifier and individual classifier.

There is also a point to modify. In individual classifier, the sampling is one second. At the same time, PAC feature is not always valid and useful as a new feature for our all the subjects.

What's more, we successfully evaluate Phase-Amplitude Coupling (PAC) feature in individual classifier. In section 6, we integrate PAC in random forest model and see its importance.

6 Real-time Individual Classifier using EEG Signals or biomarkers

In realist application, we suppose that one meditation beginner firstly does a meditation under experienced meditators' instruction. We will use this data as training data set and we customise the BCI (Brain computer interface) product for the beginner. Then the beginner can practice meditation by himself or herself with the help of the feedback from our BCI product which acquire the EEG signal of one second with a small delay and give out a feedback. In this section, we will present the motivation of building an individual classifier using EEG signals, our proposed approach to obtain such model, our training steps as well as the results.

6.1 Motivation and Objective

Since the correlation between SC and HRV is strong. We take one subject **MOHU** as an example to explain how the classification based on SC an HRV works. In Figure27, blue points stand for samples of story phase and red points stand for meditation samples. The left one contains all data points from df matrix of this subject, the right one is a zoomed-in version of the left figure where the limit of y axis ranges from five percent HRV value to nighty-five percent HRV value. From these two figures of subject **MOHU** we could see the correlation between HRV and SC is strong. Since the labels (1 for Meditation, 0 for story phase) are separable on this 2-dimension plane, we could use a C-SVM classification type linear kernel SVM(support vector machine) classifier to see the accuracy of using conventional biomarkers. For this type of SVM, training involves the minimization of the error function:

$$\frac{1}{2}\omega^T\omega + C\sum_{i=1}^{N}\xi_i$$

subject to the constraints:

$$y_i(\omega^T\Phi(x_i) + b) \geq 1 - \xi_i \;\; and \; \xi_i \geq 0, \; i = 1, ..., N$$

C is the capacity constant, the larger the C, the more error is penalized, here we choose $C = 1$, w is the vector of coefficients, b is a constant, and ξ represent parameters for handling non-separable data (input). The index i labels the N training cases. y represents the class labels and x_i represents the independent variables. The kernel Φ is a feature map such that the kernel function $K(X_i, X_j) = \Phi(X_i)\Phi(X_j) = X_i \cdot X_j$

Firstly, we used all samples to train in order to see how wide the gap between two categories (meditation and story phase) is. The figure28 represents the separation of this hyper-plan where the black circles are support vectors (101 out of 1030 samples) and the green line is the separation line of this 2-dimensional hyper-plan, these two grey lines help visualise the maximum marge.

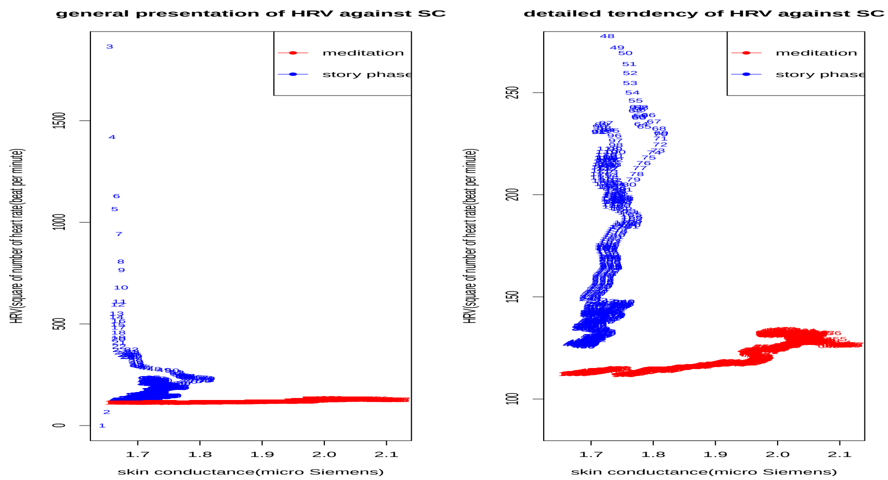


Figure 27: Scatter Plot of Skin Conductance and Heart Rate Variability

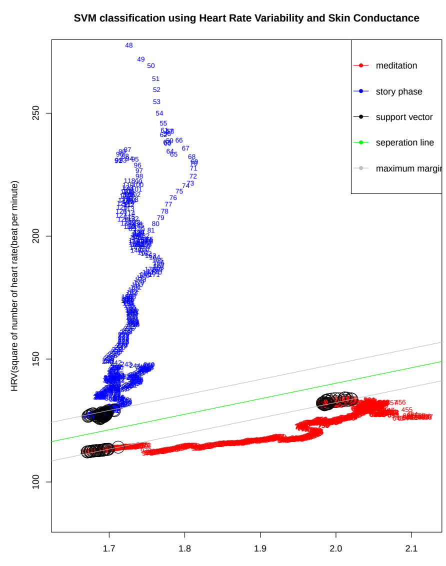


Figure 28: Preliminary Classification Using Skin Conductance and Heart Rate Variability

Furthermore, we will separate the samples into training set and test set with a $percentage\ of\ trainingset = 0.7$:

total sample	training set	test set
1030	723	307

Table 5: SVM sample summary

There are two classes (0 stands for story phase, 1 stands for meditation) in the training set. Within the training set, We used the cross-validation with 10 fold (the sample size is whether 651 or 650), repeated for 3 time. The result of Support Vector Machines with Linear Kernel in training set is:

accuracy	kappa
0.9772	0.9533

Table 6: Accuracy of training set with SVM classifier
using Skin Conductance and Heart Rate Variability

The accuracy is the ratio of correct prediction and the kappa statistic is a metric that compares the observed accuracy with the expected accuracy, The definition of kappa[15] is:

$$kappa \equiv 1 - \frac{1 - p_o}{1 - p_e}$$

where p_o is the observed accuracy, and p_e is the expected accuracy. The observed accuracy is simply the number of instances that were classified correctly throughout the entire confusion matrix. The expected accuracy is defined as the accuracy that any random classifier would be expected to achieve based on the confusion matrix. Based on the confusion matrix defined below8, we calculate the expected accuracy in this way:

$$\frac{\frac{(True\ Positive + False\ Positive) \times (True\ Positive + False\ Negative)}{number\ of\ instances} + \frac{(False\ Positive + True\ Negative) \times (True\ Negative + False\ Negative)}{number\ of\ instances}}{number\ of\ instances}$$

The kappa statistic measures how much better the classifier is comparing with guessing with the target distribution.[16]

Here we present the confusion matrix and concerning statistics in test set (see table7).

		Reference	
		1	0
Prediction	1	174	2
	0	5	264

Table 7: Confusion matrix of SVM on test set

We now look into details of this classifier in Figure9, the ”No Information Rate” is the accuracy when we pick the majority class for test set. The confidence interval of accuracy is calculated by $interval = z * \sqrt{\frac{accuracy*(1-accuracy)}{n}}$ where n is the size of the samples, and z is a critical value from the Gaussian distribution (1.96 for 95%), because each prediction is a binary trial (Bernoulli trial). We define the positive class is $'1'$ (meditation), so the confusion matrix and different statistic associated are defined as follows:

<table>
<tr><td colspan="3"></td><td colspan="2">Reference</td></tr>
<tr><td colspan="3"></td><td>condition positive</td><td>condition negative</td></tr>
<tr><td colspan="3"></td><td>1</td><td>0</td></tr>
<tr><td rowspan="2">Prediction</td><td>predicted condition positive</td><td>1</td><td>True Positive</td><td>False positive</td></tr>
<tr><td>predicted condition negative</td><td>0</td><td>False Negative</td><td>True Negative</td></tr>
</table>

Table 8: Confusion matrix definition

$$Sensitivity = \frac{true\,positive}{condition\,positive}$$

$$Specificity = \frac{true\,negative}{condition\,negative}$$

$$Pos\,Pred\,Value = \frac{true\,positive}{predicted\,condition\,positive}$$

$$Neg\,Pred\,Value = \frac{true\,negative}{predicted\,condition\,positive}$$

Accuracy	0.9772
95%Confidence intervals for accuracy	(0.9536, 0.9908)
No Information Rate	0.5831
p-value[Acc>NIR]	<2e-16
Kappa	0.9533
Sensitivity	0.9721
Specificity	0.9844
Pos Pred Value	0.9886
Neg Pred Value	0.9618

Table 9: Accuracy of training set with SVM classifier using skin conductance and heart rate variability

We noticed that the skin conductance sensor had some recording issues for certain subjects, so we discard these datas. For example, for subject $PRBA$, (see Figure30), the skin conductance did not change during the meditation phase which might be due to the fact that the contact between sensor **SC-GSR** and the finger is poor. Such situation can happen on real-life, because of the frequent movement of fingers, causing the contact issue of **SC-GSR**.

Now we test our SVM model in test set (thirty percent of samples for each subject). Here we present the boxplot (Figure29) of 6 indicators of 14 subjects in order to see the medium and variance of these indicators. However, the accuracy is not always good for all the subjects.

As we can see in Figure29, the accuracy is generally around 90% which is quite good, so as Kappa. We can also see that the Specificity (the ability to rightly predict meditation) is more steady than Sensitivity. It needs to mention that for some subjects, the hyperplan of HRV and SC is complete separable (see Figure31)

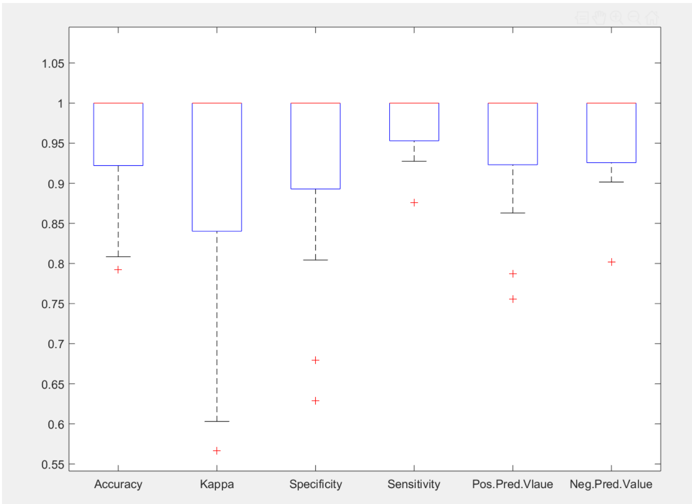


Figure 29: Boxplot of 6 Indicators of 14 Subjects

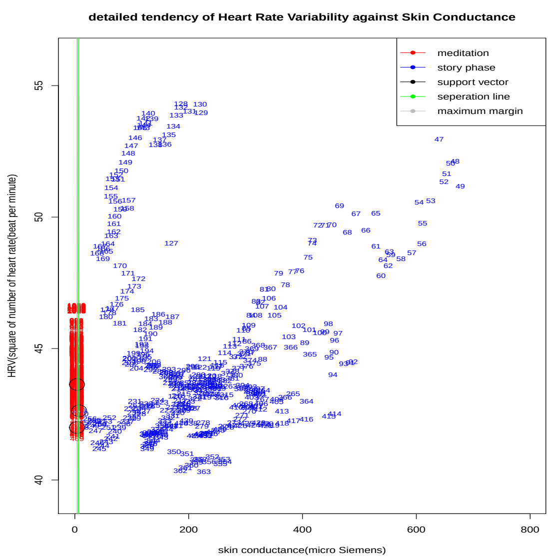


Figure 30: Example of Malfunction of SC/GSR Sensor

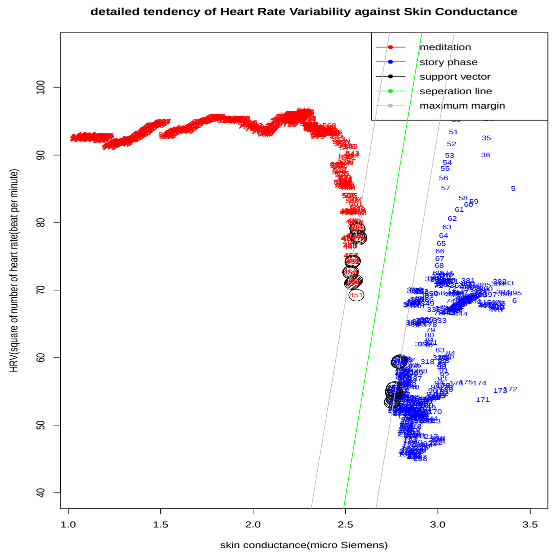


Figure 31: example of completely separable HRV-SC hyperplan

In order to show that the SVM-model using HRV and SC does not depend on the experiment. We asked subject **ETLA** to do two set of experiment. And we used the data from the first set of meditation and story phase for training, and the data from the second set for testing. Here is the result in form of confusion matrix (see Figure32). In this figure, '0' stands for story phase, we spot that this classifier do better in classifying story phase than classifying meditation. This can be also interpreted as the ordinary performance of subject **ETLA** during meditation.

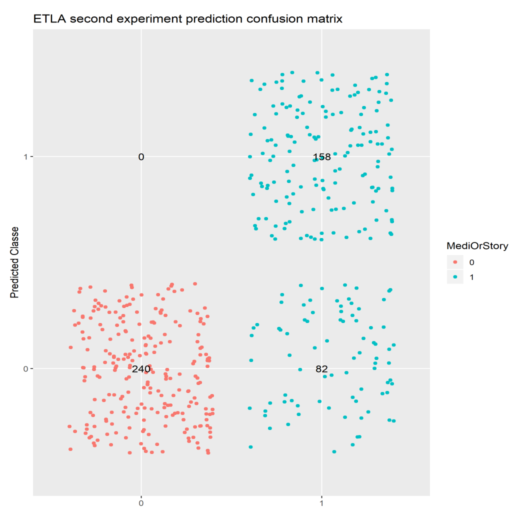


Figure 32: Confusion Matrix of ETLA

Thus we are going to dive into classification using EEG signals, which is also our objective of this project. Besides, in terms of sensors, the meditator has to put on BVP sensor (Blood Volume Pulse) and SC/GSR sensor (Skin Conductance/Galvanic Skin Response) if the classifier is built base on SC and HRV, while the meditator use a cap in case of EEG signals-based classifier. These two classifier will more or less have impact on the comfort of meditation. So there is no difference between the choice of signals considering the effect on the comfort. To sum up, since it is trivial to choose either traditional biomarks or EEG signals, we are motivated to build a EEG signals-based classifier in order to reach a better and steady accuracy while using as less electrodes as possible.

In the following section, our objective is to propose a only EEG-signal using classifier and select limited number of features to train the model and test its accuracy.

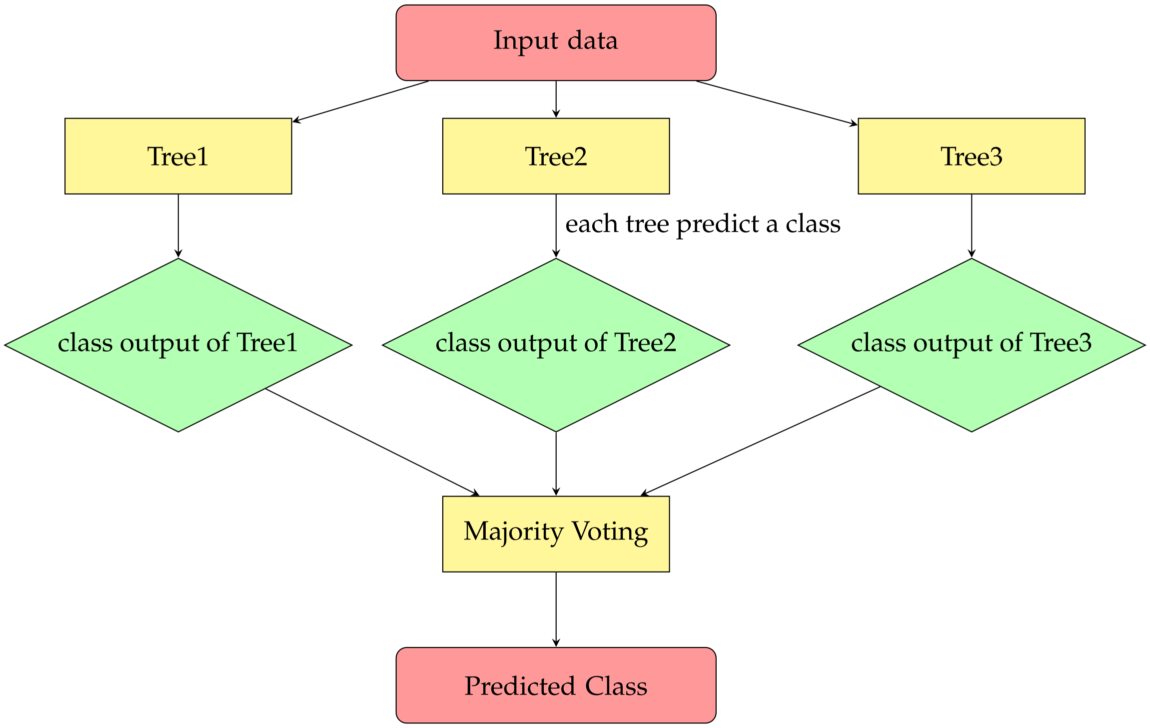
6.2 Proposed Approach

6.2.1 Overview

This module includes three parts - classification model, feature engineering and validation. We used the *df* matrix described above as our input data (Table3). The first step is feature extraction wherein we extract mathematical operation features such as the ratio between different wave band for each electrode. Secondly, we do the classification using a random forest classifier because random forest classifier show its efficiency in classifying EEG signals[17].

6.2.2 Random Forest Classifier

EEG signals data is quite complex, for example, in the math-solving mental state, we can not recognize by eyes a clear pattern in real-time recording device. Classification in our model is binary with 1 standing for meditation and 0 standing for story phase. Since different MBA value can increase or decrease during meditation compared with baseline state[18], the appropriate classifier should be decision tree classifier. However, decision-tree learners can create over-complex trees that do not generalize the data well which is so-called over-fitting. In addition, while working with continuous numerical variables, decision tree loses information. Thus we propose to uses an ensemble learning approach – random forest classifier which works in a similar way as the decision tree classifier only with an ensemble learning approach added to it. It will firstly create many random decision trees each predicting a particular class according to the features given to it. Once each tree predicts a class, a voting is carried out to take into consideration the final class according to majority. The output is then the class which has the majority voting as explained in the flow chart below.



6.3 Feature Engineering

Firstly, we want to use only EEG signals, so we delete SC, ST, HR, HRV and Resp from the df matrix. Then we introduce PAC and values of MBA Ratios, and use feature importance to select the best features to train the model.

6.3.1 Ratios of EEG signals

In the software of Biotrace+ we find some predefined ratios of EEG signals that can be exported directly. After doing some research, we decide to use the two most important ones: Theta/Beta, Alpha/Theta. The first one (TBR) is shown to be an objective marker of executive cognitive control (and more specifically attention control; AC) in healthy adults [19]. The Second one (ATR) is used to assess the level of relaxation [20]. We calculate the ratio of TBR and ATR of each electrodes so that we have 42 more features to be used in the Random forest model.These two ratios have been proved useful in identifying the Story phase and Meditation phase (see section 6.5).

6.4 Model Training

The training consists of three parts that are parameter tuning, evaluating and selecting features as well as using a subset of features to train and test the model. After feature engineering, we got a larger df, as our goal is to use as less features as possible to achieve a high accuracy, we will train the random forest classifier with all the feature including MBA values of each band for each electrode and these two ratio that we mentioned above. Then we evaluate the feature importance and select a subset of features which are relevantly important. Finally, after selecting a limited number of features, we try different parameter of the model and select the best one and test it on the test set. In the following subsections we will present the first two parts in details.

6.4.1 Choice of Parameters

In our protocol, the subjects perform one meditation and one story phase. In order to simulate the real-life process, we segregate the acquired experiment data (df) into training set and test set with a proportion of training set P. The training set is considered the first recording under instruction of experienced meditator and the test set is considered one self-practice experiment.

Then we performed the training using the training data with label 1 standing for meditation and 0 standing for story phase. In this stage, we have $21 \times (4+2) = 126 features$ incluing 4 MBA values for each of 21 electrodes and 2 derived ratio of MBA of these 21 electrodes. We choose $ntree$ number of trees as 500, and $mtry$ number of variables randomly sampled as candidates at each split as the square root of number of variables which is 11. The minimum size of terminal nodes $nodesize$ that we choose is 1 for the binary meditation state classification, that is to say, we let each tree fully grow.

All the parameters used for feature evaluation are listed below, if not mentioned, the parameter is the default parameter of $randomForest$ package in **Rstudio**.

Parameter Name	Description	chosen value for feature evaluation
P	proportion of training set	square root of the number of predictor variables
$ntree$	number of trees	500
$mtry$	number of variables used at each split	11
$nodesize$	size of terminal nodes	1
$splitrule$	split criterion	Gini index

Table 10: Parameter overview

6.4.2 Feature Evaluation

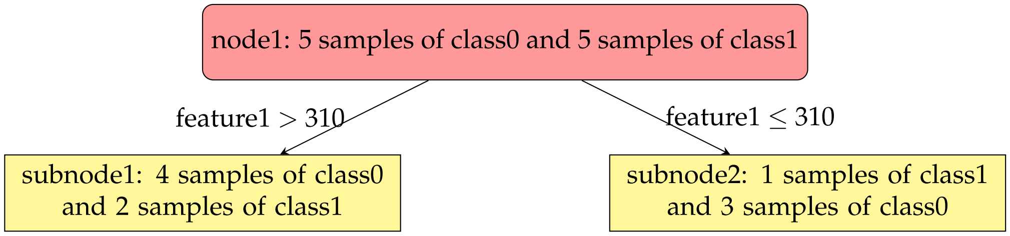
6.4.2.1 Definition of Feature Importance

After the first training, we use *MeanDecreaseGini* to measure the feature importance. To illustrate this measuring method, we will firstly present the *Gini Index* [21]. The *Gini Index* is calculated by subtracting the sum of the squared probabilities of each class from one. For each split (i.e. a node in one decision tree),

$$Gini = 1 - \sum_{i \in C} (p_i)^2$$

where C is the set of class, in our case $C = \{0, 1\}$, while p_i is the proportion of class i at this split.If in this split, samples are perfectly classified, *Gini Index* would be zero. When samples are evenly distributed, *Gini Index* would be $1 - \frac{1}{number\ of\ Classes}$.

And, *MeanDecreaseGini* is the total decrease of $GiniIndex$ from splitting on the variable, averaged over all trees. For the calculation of the decrease of $GiniIndex$, we present the following example:



the *Gini Index* of node1 is $1-(0.5)^2-(0.5)^2=0.5$, and the *Gini Index* of split nodes of node1 is calculated by weighting and suming each of the splits based on the proportion of the data each split takes up, that is $0.6\times(1-(4/6)^2-(2/6)^2)+0.4\times(1-(1/4)^2-(3/4)^2)=0.6\times(1-0.56)+0.4\times(1-0.625)=0.264+0.15=0.414$. So the decrease is $0.5-0.414=0.086$.

Once we know about the *Gini Index* definition, it's worthy to explain more about the $mtry$ parameter in RandomForest of Rstudio, now let us imagine that a single tree is added to a Random Forest model. The standard recursive partitioning algorithm would start with all the data and do an exhaustive search over all variables and possible split points to find the one that best "explained" the entire data - reduced the node impurity the most[22]. The data are split according to the best split point and the process is repeated in the left and right leaves in turn, recursively, until the size of terminal node is 1. The key thing here is that each time the recursive partitioning algorithm looks for a split all the variables are included in the search. Where RF models differ is that when forming each split in a tree, the algorithm randomly selects $mtry$ variables from the set of predictors available. Hence when forming each split a different random set of variables is selected within which the best split point is chosen. Hence for large trees, it is at least conceivable that all variables might be used at some point when searching for split points whilst growing the tree.

6.4.2.2 Feature Importance

When we trained the random forest model, we calculate the importance of each feature for each test subject and adopt a statistical method to analysis the importance of each feature. The original goal was to find out the best features that can be used for each test subjects in a general classifier. We calculated the average importance for each feature in 12 test subjects. The 5 most important feature are shown in the following table below.

Feature	Average Importance	Standard Deviation of Importance
O1Beta	21.69377962	33.17132665
T3Beta	14.39940671	22.34634284
O2Beta	12.82971241	17.27317681
C3Beta	9.746442211	17.64865159
Cz alpha/theta	9.351644362	13.27905536

Table 11: Feature Importance Analysis

As we can see from the table, the average importance of beta band was high, but the standard deviation of importance is very high too. Actually when we do the random forest model, O1beta worked well for only two subjects in which cases the importance was very high. However for other people, the importance of O1beta was quite low. That explain why O1beta have a very high variance among subjects.

After doing a statistical analysis of the importance of all the 126 features used in Random Forest Model, we found out that the importance of features really depends the test subjects. There doesn't exist feature which work well for every test subjects. That is why an individual classifier are needed in random forest model in order to have a good accuracy.

6.4.3 Candidate Model Selection

Except seleting five most important features, we can also tune the parameters such as *ntree* and *mtry* or even *split rule*. Due to the computational complexity and limited time, in current stage, we search 3 times using 3-folds *cross validation* in the training set to select the best *mtry*. The *mtry* value is randomly selected at each time, the final model will use *mtry* the one that has higher accuracy. Of course, in our final classifier, we can definitely greedily search for the best parameters using the data of clients doing meditation under instruction.

Here we take one subject **MOHU** as an example to present how the parameters influence the model accuracy. We use more features (10 most important features) to enable a more extensive research on the parameters. At the first time, we use a 10-fold cross validation to test *mtry* varing from 1 to 9 with *ntree* fixed at 500, the result is shown in Figure33. We can see that accuracy reach to summet when *mtry* equals to 2. At second step, we will change ntree to 300 (see Figure34) and 700 (see Figure35) to demonstrate the inpact of *ntree* parameter.

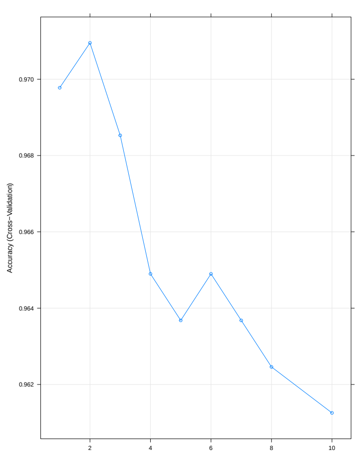


Figure 33: How mtry influence random forest modelaccuracy (ntree=500)

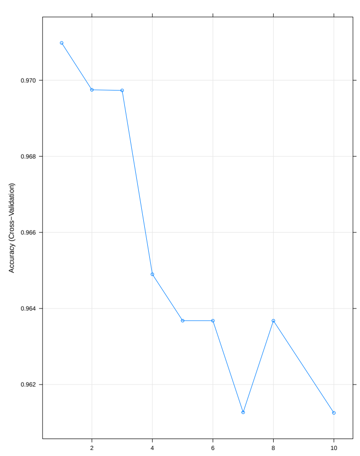


Figure 34: How mtry influence random forest modelaccuracy (ntree=300)

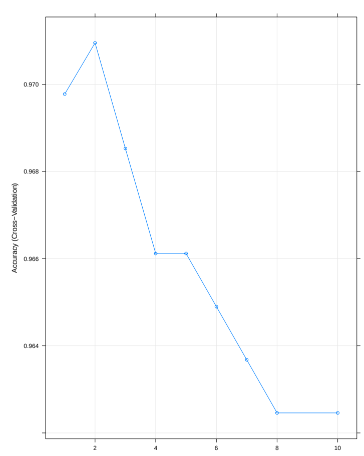


Figure 35: How mtry influence random forest modelaccuracy (ntree=700)

6.5 Result and Discussion

In this section, we firstly present the evaluating statistic of random forest model, then we show the improvement that using PAC as features brings. We also show that the model holds water for several re-experiments of one subject then come to the conclusion and discussion at the end.

6.5.1 Overall Accuracy

As is mentioned above, we used 80% of data as the training data for the model and 20% of data as test data to test the accuracy of our model. In the first case, we only used 4 different bands of EEG in each electrodes as features, so there are a total of $21 \times 4 = 84$ features. In this case, the accuracy is 92.0%. In the second case, we used TBR (Thelta Beta Ratio) and ATR (Alpha Thelta Ratio) of each electrodes as additional features. Thus, we have a total of $84 + 21 \times 2 = 126$ features. With the help of new features, the overall accuracy increased a little bit to 92.4%.

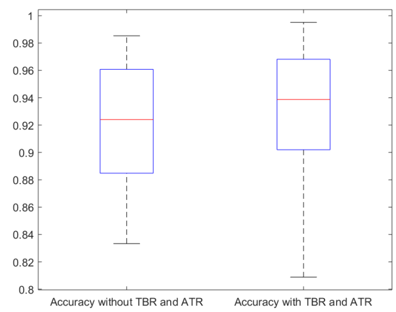


Figure 36: Overall Accuracy without and with TBR and ATR

As we can see from the BoxPlot in figure37, The use of the ratio of EEG bands as new biomarker not only make the box position higher, but also make the box smaller. This means these features make the classifier more accurate and reliable in general. However, this doesn't means adding new features can benefit everyone. There are 3 out of 12 subjects whose accuracy dropped when we used TBR and ATR.

In order to have a clearly view of general accuracy, we also plot the confusion matrix for all the the experiment data, the result is shown below. This over all confusion matrix was plotted using TBR and ATR as features.

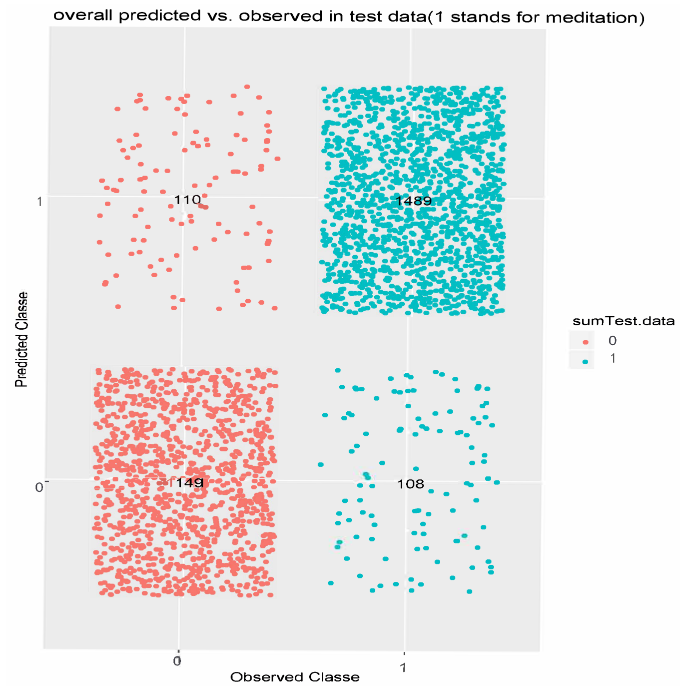


Figure 37: Overall Confusion Matrix

We now compare the statistic result of Random Forest Model on test set (see Figure 38) and the SVM model using SC and HRV on test set (see Figure 29). It's true that SVM model has higher median value. Both models have an accuracy generally higher than 90 percent. And these two model have almost the same variance without considering the extreme value.

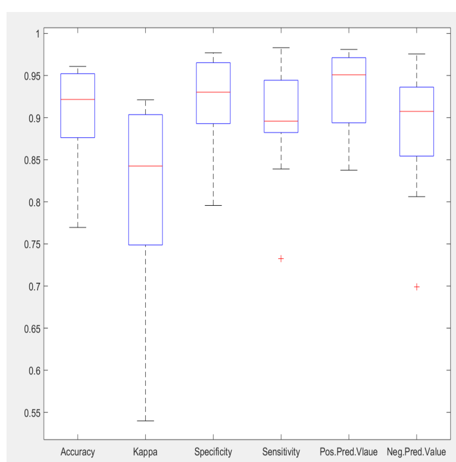


Figure 38: Overall Statistics of Random Forest Model

6.5.2 Result Using PAC as Feature

Limited by our time, we use PAC as Feature for only two subjects. For the first subject, the accuracy was already 95% with the help of TBR and ATR. When we added PAC as new features (147 features in total), the accuracy increased to 96%. When we look into the importance of features, the PAC of two electrodes entered top 5. Because we only use the 5 most important features in the classifier, this means we actually used the PAC of two electrodes as features in the random forest model. In the second subject, the accuracy we reached using conventional feature with BTR and ATR is 89.7% after we added PAC as features, The accuracy increased to 91.7%. The improve by using PAC is more evident and it can also be seen in the confusion matrix in the figure39.

Figure 39: Confusion Matrix without PAC (left) and with PAC (right) on CACA

Statistic	Without PAC	With PAC
Accuracy	0.897	0.917
Sensitivity	0.9491	0.920
Specificity	0.828	0.912
Pos.Pred.Value	0.881	0.929
Neg.Pred.Value	0.923	0.902

Table 12: Improvement Brought by PAC

6.5.3 Re-experiment Validation

The objective of Re-experiment validation is to prove the individual classifier does not depend to the experiment. That is to say the classifier trained in one experiment for one subject works well in other experiments for the same subject. We have done two validation experiments for two different test subjects. In the first validation experiment, we ask the subject to do another 5 minute of meditation after finish all the procedures in the experimental protocol. So the model was still trained with the data from the experiment following the protocol, and the 5 minutes of additional meditation are used as test data to test the model.

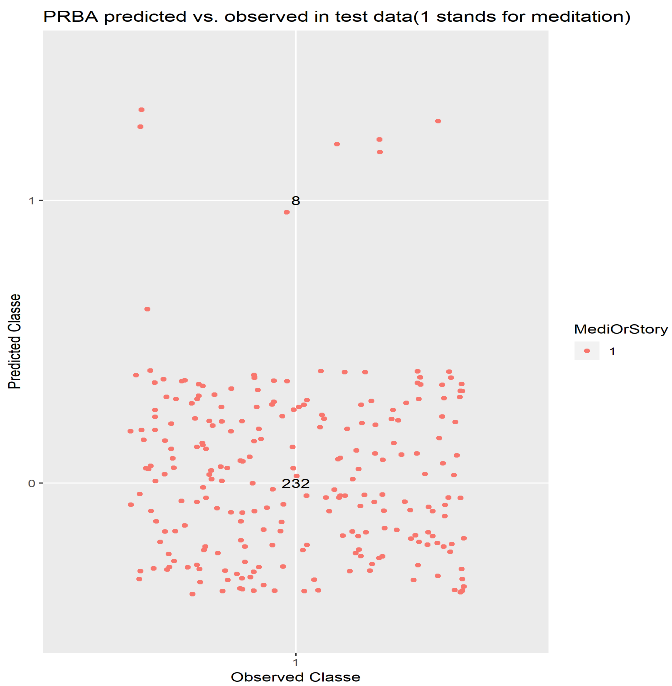


Figure 40: Result of Validation Experiment on PRBA (Meditation only)

The results are shown in the figure 40. We take 4 minutes of sample from 5 minutes of second meditation data. One second is a point, so we have a total of 240 points. As is shown in figure 40, only 8 points out of 240 point were wrongly identified as story, which gives us a total accuracy of 96.7%. Although the accuracy seems very high, we quickly realized that we couldn't totally trust this validation experiment. We can not reproduce the story phase because we only have one story and the subject have already know the story and the answer to the questions. That is the reason why we asked the test subject to do only the mediation phase for the validation experiment. So in this case we are not sure whether our classifier over-evaluates the meditation phase or not. Imagining a bad classifier which classifies every point of data as meditation, in which case, it will give us 100% of accuracy.

In order to avoid this kind of situation, in the second validation experiment, we required the test subject not only to do meditation once again, but also we prepare mathematical calculation and guided imagination to keep the test subject mentally active. The confusion matrix are shown in the figure 41. As we can see all the points from the story phase are correctly classified and 66 % of the points from the meditation phase are correctly classified. However, this relatively low sensitivity may not due to the problem of the model. In this experiment, the surroundings are quite noisy, which may have bad influences on the quality of meditation.

Figure 41: Result of Validation Experiment on ETLA

Statistic	Value
Accuracy	0.83
Sensitivity	0.66
Specificity	1
Pos.Pred.Value	1
Neg.Pred.Value	0.754

Table 13: Statistic of Validation Experiment

6.5.4 Conclusion and Discussion

In the above section, we develop a Random Forest Classifier based on EEG signals and compare it with SVM classifier using SC and HRV as input. The accuracy of SVM classifier is slightly better than Random Forest Classifier but with more extreme cases, that is, SVM classifier has bad performance on more extreme cases. These cases are indicated by a red cross in the boxplot figure. We also present the strength of using PAC as features, and we recommend to use them in future research. Last thing to mention, we prove that the individual classifier does not depend on experiment and the classifier can be real-time once we can obtain the EEG signal instantly. The real-time BCI is workable by using the Random Forest Model.

7 ”Real-time” General Classifier Using EEG Signals

7.1 Motivation

The global results presented in Section 5.1 give ground for the training of a general classifier. If a high precision can be reached with this classifier, this will prove that the two states are EEG-differentiable in a similar way for all the subjects. We will try to build a ”real-time” classifier with features computed on small time intervals: 3 classifiers are tested, based on 60s, 30s or 10s samples.

This section presents the choice of features, the algorithm used and discusses the results and validation.

7.2 Choice of Features

7.2.1 Definition of Normalized Features

MBP60, MBP30 and MBP10 from Story and Meditation can’t be directly injected in the classifiers. Indeed, their values can be dissimilar from one subject to another: what really matters is the variations, or the ratio between Story or Meditation. So we need reference values for each subject. We arbitrarily chose that the MBP of the whole story phase, called MBPref, would stand as reference. For a real application, this would correspond to asking the subject to perform the whole Story phase first so as to calibrate his features and be able to classify his state afterwards. This Calibration phase could also be done with the Meditation phase. The study of the time needed to obtain a trust-worthy reference has not been done, and it is probable that a calibration of only a few minutes would be sufficient. Here it should be noted that since there is only one Story phase here, MBPref is actually also a mean of all the MBPduration from Story.
Formally the definition of features is:

$$Feature(band, area) = \frac{MBPduration(band,area)}{MBPref(band,area)}$$

With 4 frequency bands in the 6 brain areas, we have a total of 24 features.

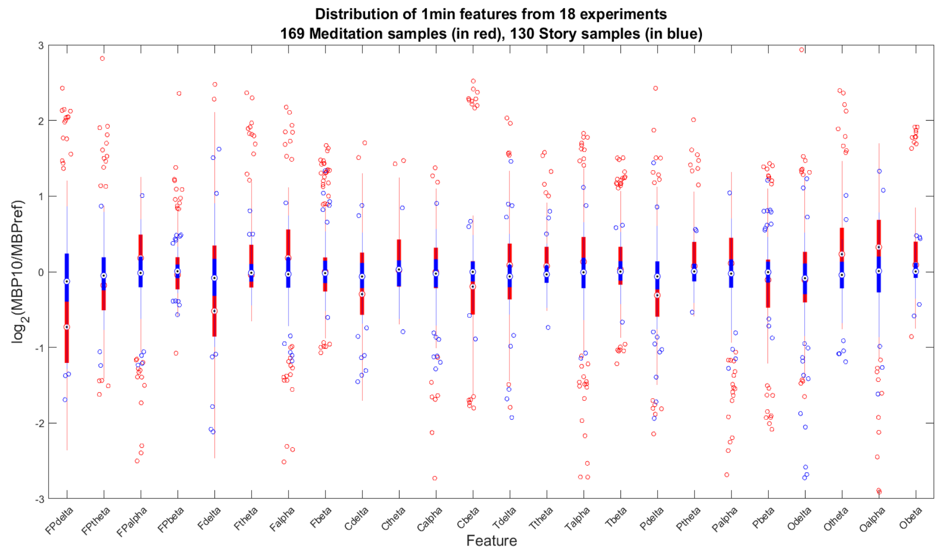


Figure 42: Distribution of 60s features (log2 scale)

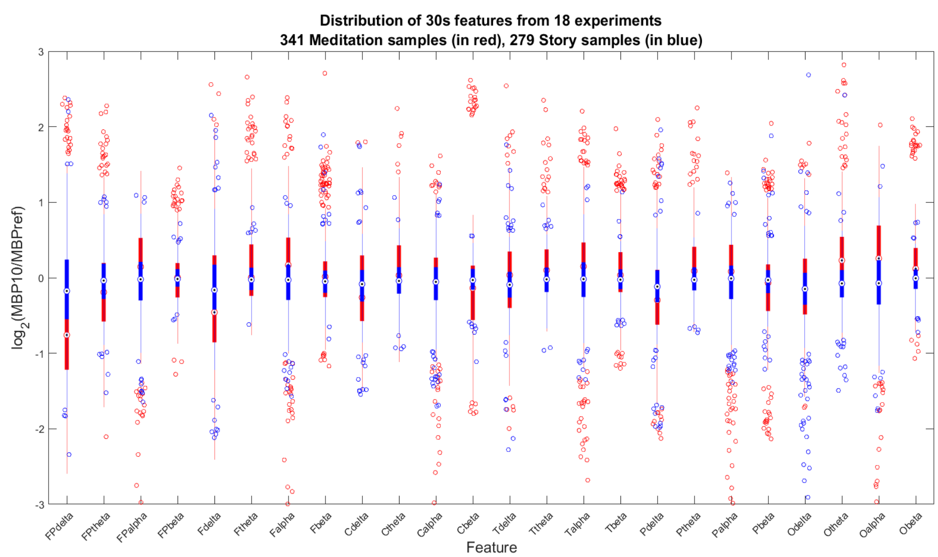


Figure 43: Distribution of 30s features (log2 scale)

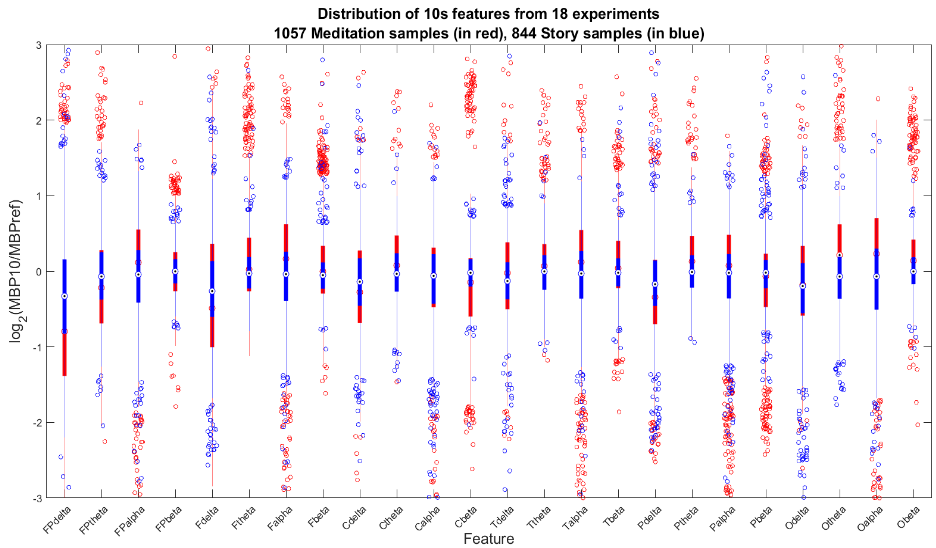


Figure 44: Distribution of 10s features (log2 scale)

Figures 42, 43 and 44 show the distribution of 60s, 30s and 10s features from the 18 training experiments. First important observation, the boxes for Story (in blue) and Meditation (in red) don't completely overlap, which means the values of these features should allow us to discriminate the two states (this had already been observed in section 5.1). The fact that no feature has completely distinct Story and Meditation boxes (not mentioning the outliers) justifies the use of several features, because one is not enough. The repartition remains similar between the 3 time scales, which is not surprising, however we can notice that the boxes broaden as the duration of samples decreases ; indeed there are more points so more potential variations. We may wonder what is the best time scale : a longer time scale for MBP allows to "smooth" the predicted state and observe more distinct values ; but on the other hand, it is more strongly affected by parts of the signal which lead to mistakes of the classifier, and moves away from a real-time application (even though we can use overlapping samples to boost the "refresh rate").

7.2.2 Principal Component Analysis

Principal component analysis (PCA) is a statistical procedure that uses an orthogonal transformation to convert a set of observations of possibly correlated variables (entities each of which takes on various numerical values) into a set of values of linearly uncorrelated variables called principal components. If there are n observations with p variables, then the number of distinct principal components is $min(n-1,p)$. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible under the constraint that it is orthogonal to the preceding components. The resulting vectors (each being a linear combination of the variables and containing n observations) are an uncorrelated orthogonal basis set [23].

We applied PCA to the 30s features, and 18 PCA components were retained for a total explained variance of 99%. Then we used the coordinates of these PCA components in the original feature space (24 coordinates) to estimate the importance of each feature, since the value of each coordinate represents the ”participation” of each associated feature to the PCA component, which accounts itself for a certain variance of the data. More precisely, the contribution of a given feature to a PCA component is obtained by dividing the absolute value of its coordinate by the sum of absolute values of all the coordinates of the PCA component. These contributions (in %) were regrouped by frequency band and by brain area (each feature corresponds to a pair [band, area]) for each PCA component, and then multiplied by the explained variance of this component to give the following figures. The tables show the total contribution (sum of the contributions multiplied by variances for all PCA component) per frequency band and per brain area.

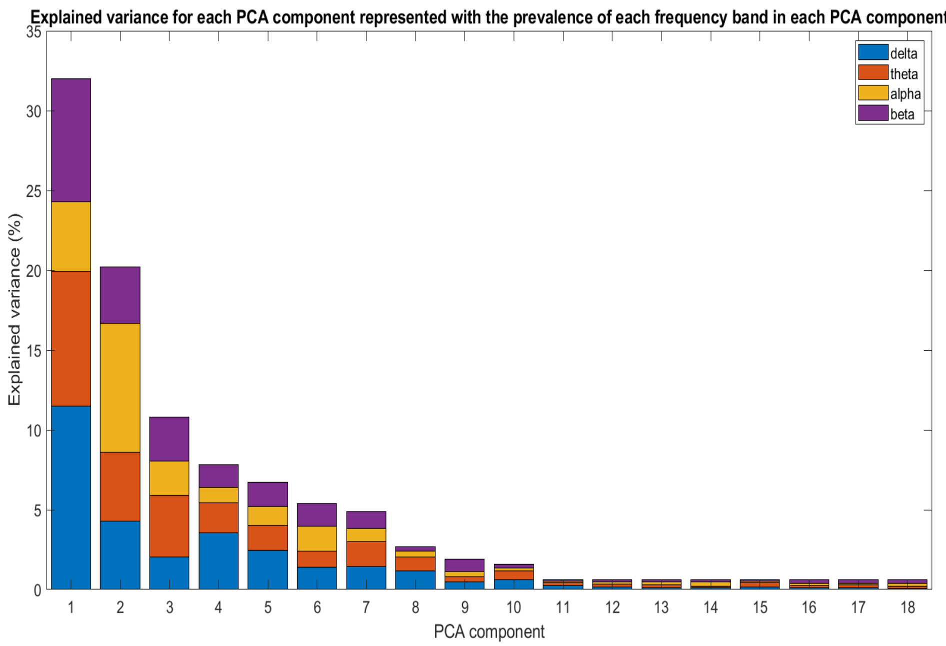


Figure 45: Contribution to the variance per frequency band for each PCA feature

Band	Delta	Theta	Alpha	Beta
Total contribution (%)	30.0	25.7	21.3	22.0

Table 14: Total contribution to the variance per brain area

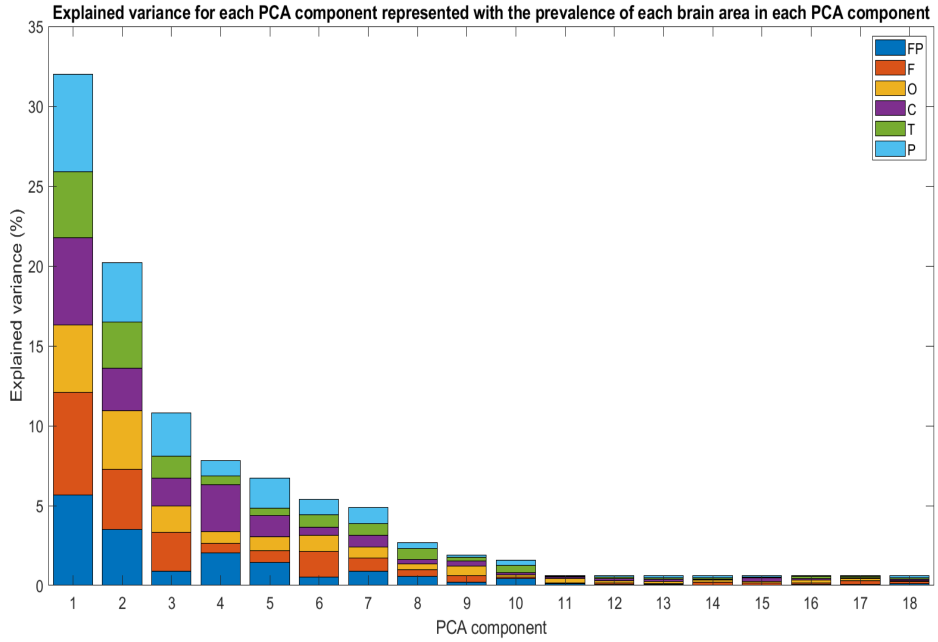


Figure 46: Contribution to the variance per brain area for each PCA feature

Area	FP	F	O	C	T	P
Total contribution (%)	16.8	18.2	15.0	17.0	13.0	19.0

Table 15: Total contribution to the variance per brain area

Interestingly enough, all the frequency bands and all the brain areas seem to contribute to the variance of data. In terms of band, Delta accounts for the most part of the variance ; in terms of area, it is the Parietal area. However it doesn't mean that Pdelta is the feature that accounts for most of the variance. Neither does it mean that Delta features or Parietal features are the best ; they might account for a lot of variance, but this variance might also not be linked to the nature of the phase! In other terms, a variable which takes a lot of different values is not necessarily related to the fact that the subject is meditating or not.

7.3 Subspace-KNN Classifier

7.3.1 Classical KNN Algorithm and the Curse of Dimensionality

In pattern recognition, the k-Nearest Neighbors algorithm (KNN) is a common non-parametric method used for classification. An object is classified by a plurality vote of its neighbors, with the object being assigned to the class most common among its k nearest neighbors in the feature space (k is a positive integer, typically small). Here the Euclidean metric is used to compute the distances.

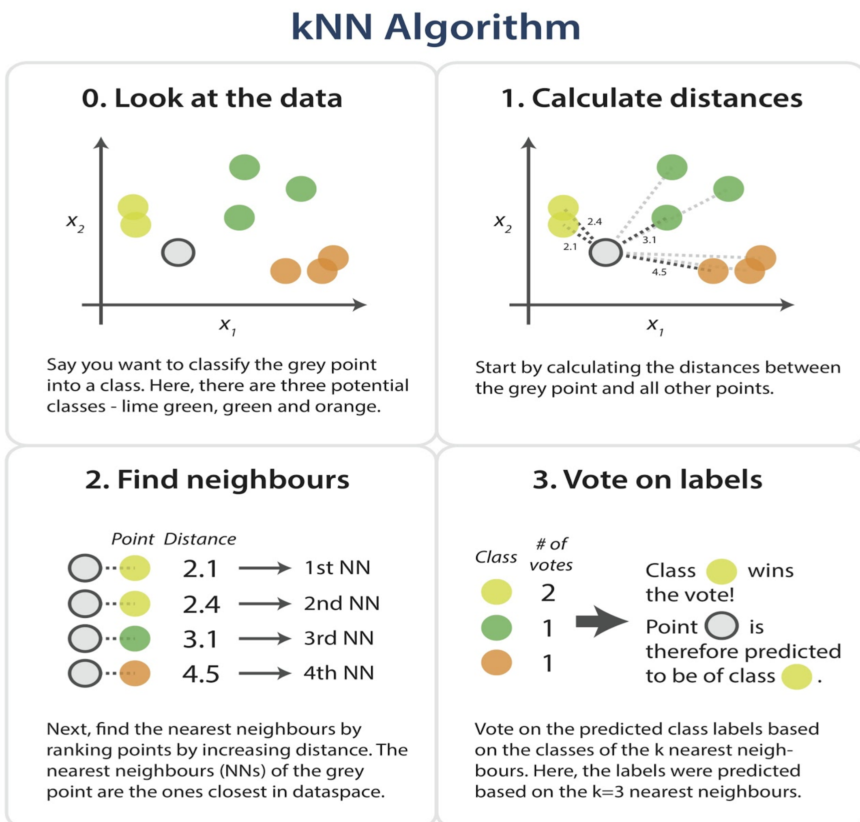


Figure 47: Operating principle of the KNN algorithm [24]

KNN is easy to implement and to interpret, and has some strong consistency results. As the amount of data approaches infinity, the two-class KNN algorithm is guaranteed to yield an error rate no worse than twice the Bayes error rate (the minimum achievable error rate given the distribution of the data).

However in high dimension the "curse of dimensionality" can limit the efficiency of this algorithm. Indeed as the number of features increases Euclidean distance becomes unhelpful because all vectors are almost equidistant to the search query vector (imagine multiple points lying more or less on a circle with the query point at the center; the distance from the query to all data points in the search space is almost the same) : all objects appear to be sparse and dissimilar in many ways [25]. Several methods exist to solve this issue, among which the Random Subspaces method used here.

7.3.2 Random Subspaces of Features

The random subspace method relies on a stochastic process that randomly selects a number of components of the given feature vector in constructing each classifier. In the case of k-nearest-neighbor classifiers, that means when a test sample is compared to a prototype, only the selected features have nonzero contributions to the distance. Geometrically this is equivalent to projecting all the points to the selected subspace, and the k nearest neighbors are found using the projected distances. Each time a random subspace is selected, a new set of k nearest neighbors are computed. The k nearest neighbors in each selected subspace are then assembled for a majority vote on the class membership of the test sample. The same training sample may appear more than once in this ensemble if it happens to be among the k nearest neighbors in more than one selected subspace [25].

7.3.3 Parameters

The number of learners n and the dimension of subspaces s, as well as the number of neighbors k used in KNN should be optimized to improve the accuracy. Here the simplified tools of MATLAB used give no information about the number of neighbors k. The basic parameters $n = 30$ and $s = 12$ (half of the features) are common values and seemed to achieve the best accuracy.

7.4 Results

7.4.1 Accuracy

We used a 5-fold cross validation to compute the accuracy of these classifiers. That means the data is partitioned into 5 "folds" ; then 5 classifiers are trained, using 4 of the folds as training data while the last one is kept for validation (a different fold each time). The accuracy computed for the general classifier is the average of the accuracies of these 5 classifiers. The advantage of this technique is that in the end all the data is used for both training and validation, which is particularly useful for small data sets.

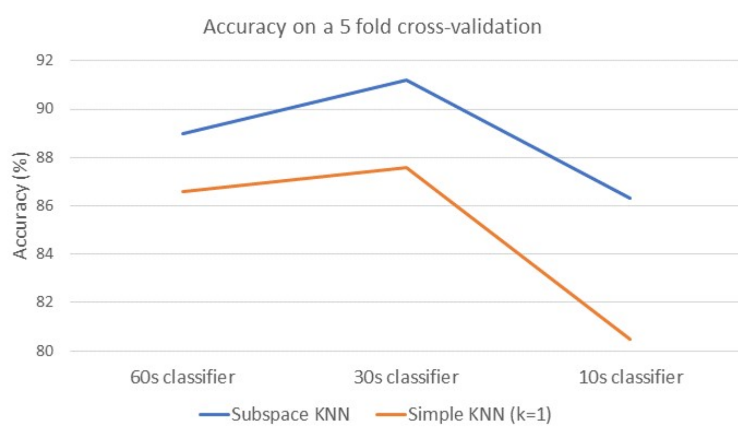


Figure 48: Accuracies of the 60s, 30s and 10s classifiers

The cross validation performed doesn't take into account the subject from which comes each data vector, which means that for each classifier samples from different times of a same experiment can be found both in the training folds and the validation fold. So even though a validation data vector has not been used for training, the classifier might already be "familiar" with the experiment it comes from. Because of this, and since the number of experiments is actually low compared to the number of samples, the accuracy computed is probably over-estimated.
In the confusion matrices below, 1 means Meditation, 0 means Story ; rows correspond to the real state and columns to the state predicted by the classifier.

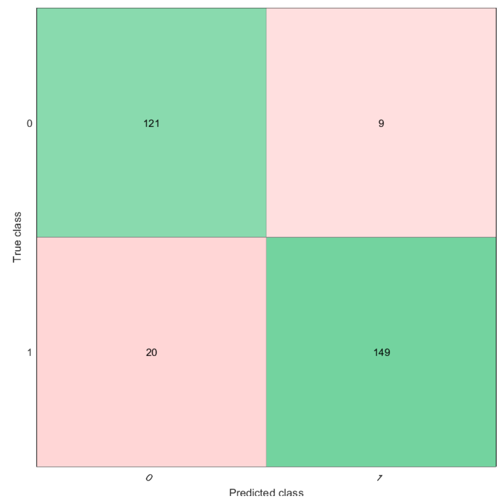


Figure 49: Confusion matrix of the 60s classifier

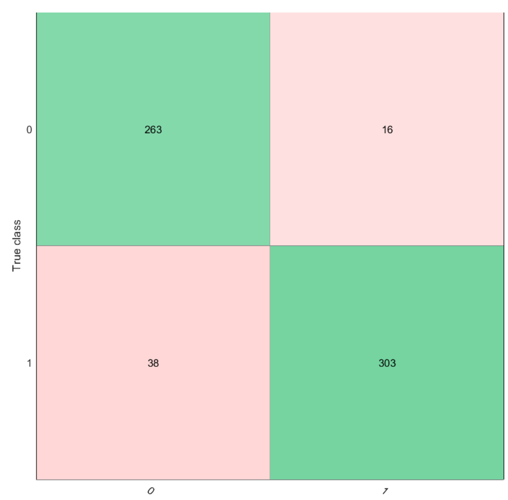


Figure 50: Confusion matrix of the 30s classifier

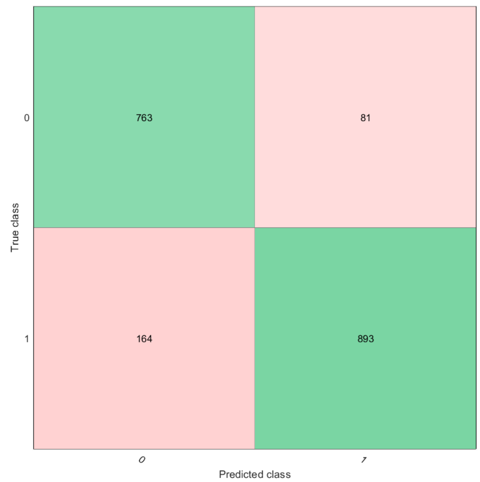


Figure 51: Confusion matrix of the 10s classifier

From these 3 confusion matrices, we can deduce the performances of the classifiers regrouped in Table 17. The 30s classifier gives the best accuracy. The fact that the Pos.Pred.Value is quite high compared to the Sensitivity, means that the classifiers don't correctly label all the Meditation samples as Meditation, but when a sample is labeled as Meditation there is a high probability that it is, indeed, a Meditation sample. Consequently, the inverse phenomenon happens for Story: looking this time at Specificity and Neg.Pred.Value, the classifiers have a tendency to over-classify samples as Story.

Classifier	**60s**	**30s**	**10s**
Accuracy	0.890	0.912	0.863
Sensitivity	0.868	0.889	0.845
Specificity	0.878	0.943	0.904
Pos.Pred.Value	0.934	0.949	0.917
Neg.Pred.Value	0.841	0.874	0.823

Table 16: Performances of the 3 classifiers

7.4.2 Validation

An additional step of validation will let us witness how the classifier actually performs on the two phases of an unknown experiment.

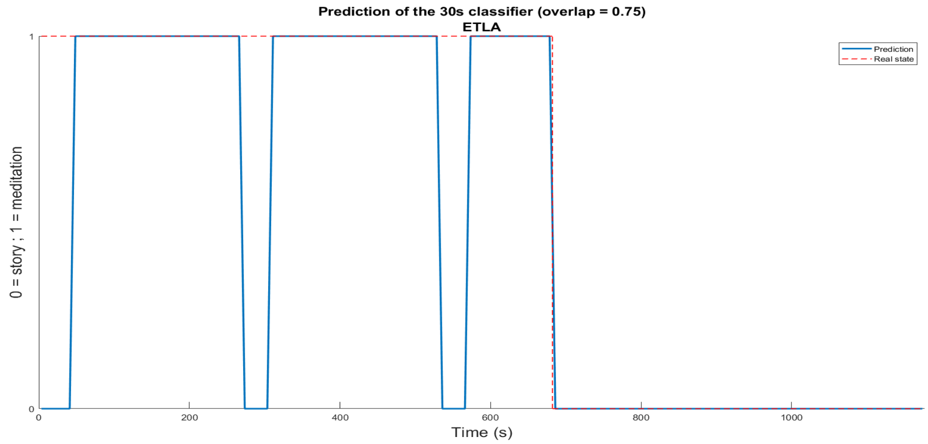


Figure 52: Prediction of the 30s classifier (in blue) and real state (in dashed red), with Meditation then Story

Figure 52 was obtained after computing the 30s features of the experiment ETLA, which was not used in the training. An overlapping rate of 0.75 was applied, which means the refresh rate of the prediction is 7.5s. The computed features were then given as input to the trained classifier, which labelled the corresponding each of them either as ”Meditation” (1) or ”Story” (0).
The predicting accuracy on this experiment is 0.898. In agreement with the confusion matrix of this classifier, a tendency to over-classify the Meditation phase as Story (false negative) is observed. The Story phase is completely correctly labeled.

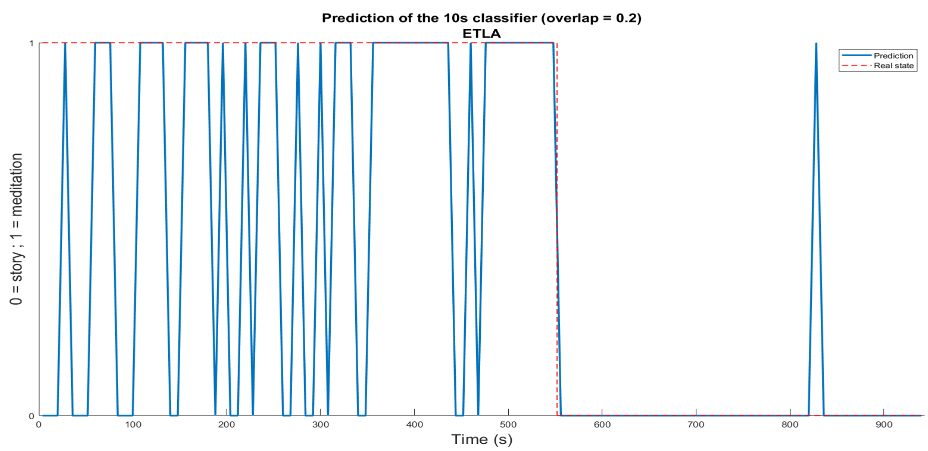


Figure 53: Prediction of the 10s classifier (in blue) and real state (in dashed red), with Meditation then Story

Figure 53 evaluates the 10s classifier. A new set of 10s features was computed on the same experiment.The overlapping rate is 0.2, which means the refresh rate is 8s.

The predicting accuracy on this experiment is 0.780. Again the Story phase is almost completely correctly classified, but most of the precision is lost on the Meditation phase. On the basis of this experiment, the 10s classifier can hardly be considered reliable.

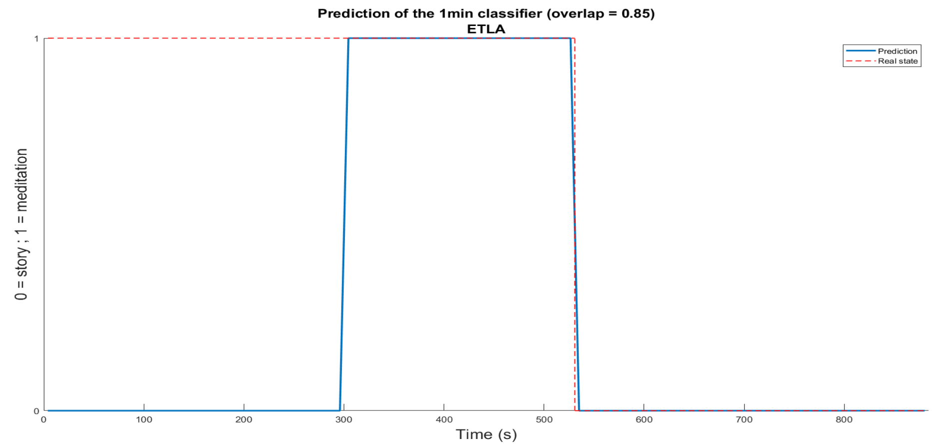


Figure 54: Prediction of the 60s classifier (in blue) and real state (in dashed red), with Meditation then Story

Finally Figure 54 evaluates the 60s classifier. The overlapping rate is 0.85, giving a refresh rate of 9s.

Surprisingly, the 60s classifier has the worst accuracy on this experiment with only 0.667. This differs a lot from the computed accuracy of 0.863, which may be due either to the subject himself or to the quality of the recording, as well as to the bias mentioned above in the computation of global accuracy. Of course this accuracy is not satisfying for a real-life application.

8 Conclusion

8.1 Limits of the Study

8.1.1 Collection of Data

The data collected should be trusted only to a certain point, for the following reasons:

- Experimental conditions : noise from the outside, movements of experimenters or irruptions in the room.
- Experimental setup : only one type of gel applied, no precise positioning of the EEG cap (through the measuring of inion-to-nasion distance, etc).
- Validation of Story for the Individual classifier : not exactly the same conditions, hard to reproduce the same state.
- Different days, times of the day ; however this allows a ”mix” of subject conditions.

8.1.2 Processing

The data has not been completely exploited, and some parts of its processing could be improved:

- Even though we collected information on the age of subjects, meditation practice and gender, these were not taken into account ; even though there didn't seem to be significant differences, it would have been interesting to perform this differentiation, with a larger number of subjects.
- The "quality" of the subjects' state of mind induced by the two phases and evaluated through the questionnaires was also not taken into account.
- Correlations between the variables were studied but not systematically evaluated since they didn't seem to present high interest for the discrimination of the two states (except for PAC). Also for the General classifier, electrodes were regrouped by brain area for the computation of features, hence hiding the possible asymmetries between the two hemispheres induced by the two states.
- The artifacts cleaning algorithm developed in parallel came late in the year, so it has not been used in most of the study. Even though it should not affect too much the general tendencies given the rarity of artifacts, it would certainly improve the accuracy of the two classifiers. The cleaning algorithm itself could be improved to be more sensitive yet spare as much data as possible.
- No optimization of the Global classifier parameters (number of neighbors, subspace dimension, number of learners) has been performed.

8.2 Main results

Model	General classifier	Individual classifier
Trained beforehand	Yes	No
Calibration/Training	Story phase (8min)	Meditation (10min) and Story (8min) phases
Algorithm	k-nearest neighbors with 30 subspaces of 12 random features	Random Forest with 500 trees, selecting the best *mtry* from 3 random trials
Sample length	30s	1s
Number of features	24	126
5 most important features	FPdelta, Ptheta, Obeta Cdelta, Pdelta	depend on the subject
Number of training subjects	18	1
Validation	5-fold cross validation	20% holdout validation the statistics below are averaged over 12 subjects
Accuracy (standard dev.)	0.912	0.908 (0.05)
Kappa (standard dev.)	0.501	0.814 (0.11)
Sensitivity (standard dev.)	0.889	0.917 (0.06)
Specificity (standard dev.)	0.943	0.901 (0.06)
Pos.Pred.Value (standard dev.)	0.949	0.884 (0.07)
Neg.Pred.Value (standard dev.)	0.874	0.931 (0.05)

Table 17: Comparison of the General (KNN) and the Individual (RF) classifiers

References

[1] R. Pi, "Eeg-based brain computer interface (bci) for smart home control using," June 2016.

[2] P. Eskandari and A. Erfanian., "Improving the performance of brain-computer interface through meditation practicing," 2008.

[3] N. R. Prymak C. A. Cauthen, "Meditation versus relaxation: An examination of the physiological effects of relaxation training and of different levels of experience with transcendental meditation," *Journal of Consulting and Clinical Psychology*, 1977.

[4] T. H. M. O. H. K. M. K. H. Y. Tetsuya Takahashi, Tetsuhito Murata and Y. Wada, "Changes in eeg and autonomic nervous activity during meditation and their association with personality traits," *International Journal of Psychophysiology*, 2005.

[5] J. P. Banquet, "Spectral analysis of the eeg in meditation." *Electroencephalogr Clin Neurophysiol.*, 1973 Aug; 35(2): 143–151.

[6] L. D. Hebert R, "Theta bursts: an eeg pattern in normal subjects practising the transcendental meditation technique," *Electroencephalogr Clin Neurophysiol*, 1977 Mar;42(3):397-405.

[7] L. Aftanas and S. Golocheikine, "Human anterior and frontal midline theta and lower alpha reflect emotionally positive state and internalized attention: high-resolution eeg investigation of meditation," *Neuroscience Letters*, 2001.

[8] D. Henz and W. Schöllhorn, "Temporal courses in eeg theta and alpha activity in the dynamic health qigong techniques wu qin xi and liu zi jue," *Frontiers in Psychology*, 01 2018.

[9] *Supervised Machine Learning: Classification*. [Online]. Available: https://towardsdatascience.com/supervised-machine-learning-classification-5e685fe18a6d

[10] J. E. T. Akinsola, "Supervised machine learning algorithms: Classification and comparison," *International Journal of Computer Trends and Technology (IJCTT)*, vol. 48, pp. 128 – 138, 06 2017.

[11] I. Saini, D. Singh, and A. Khosla, "Qrs detection using k-nearest neighbor algorithm (knn) and evaluation on standard ecg databases," *Journal of Advanced Research*, vol. 4, no. 4, pp. 331 – 344, 2013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S209012321200046X

[12] B. Farnsworth. (2017) Heart rate variability – how to analyze ecg data. [Online]. Available: https://imotions.com/blog/heart-rate-variability/

[13] J. O. Lisman J E, "The theta-gamma neural code," *Neuron*, 2013.

[14] M. W. Angela C.E.Onslow, Rafal Bogaczb, "Quantifying phase–amplitude coupling in neuronal network oscillations," vol. 105, pp. 49–57, March 2011.

[15] N. Smeeton, "Early history of the kappa statistic," *Biometrics*, 1985.

[16] *kappa statistic*. [Online]. Available: http://precog.iiitd.edu.in/people/anupama

[17] S. Bose, V. Rama, N. Warangal, and C. B. Rama Rao, "Eeg signal analysis for seizure detection using discrete wavelet transform and random forest," in *2017 International Conference on Computer and Applications (ICCA)*, Sep. 2017, pp. 369–378.

[18] K. T. W. H. Hinterberger T, Schmidt S, "Decreased electrophysiological activity represents the conscious state of emptiness in meditation," *Front Psychol*, pp. 5–99, Feb 2014.

[19] S. L. P. P. Angelidis A., van der Does W., "Frontal eeg theta/beta ratio as an electrophysiological marker for attentional control and its test-retest reliability," *Biological Psychology*, 2016.

[20] T. A. Lin and L. R. John, "Quantifying mental relaxation with eeg for use in computer games," pp. 409–415.

[21] *Decision Tree Flavors: Gini Index and Information Gain*. [Online]. Available: http://www.learnbymarketing.com/481/decision-tree-flavors-gini-info-gain/

[22] G. S. (https://stats.stackexchange.com/users/1390/gavin simpson), "Random forest mtry question," Cross Validated, uRL:https://stats.stackexchange.com/q/102870 (version: 2014-06-10). [Online]. Available: https://stats.stackexchange.com/q/102870

[23] [Online]. Available: https://en.wikipedia.org/wiki/Principal_component_analysis

[24] [Online]. Available: https://www.kdnuggets.com/2016/01/implementing-your-own-knn-using-python.html

[25] T. K. Ho, "Nearest neighbors in random subspaces," *Lecture Notes in Computer Science book series (LNCS, volume 1451)*, 1998.

Appendices

A Information of Subjects

CODE	Age	Nationality	Gender	story first	Health issue	last 24h-alcohol/drugs/smokin	Meditation practice	Frequency	Duration	Meditation type
PAMA	20	French	F	Y	N	N	N			
LETH	21	French	M	N	N	N	N			
LOVA	21	French	F	N	N	N	N			
TRWI	19	French	M	Y	N	N	N			
LENO	21	French	F	N	N	N	N			
MATA	20	French	M	Y	N	N	Y	< 3 times/M	30m-1h	Auto hypnosis
DEER	21	French	M	N	N	N	N			
EZLE	20	Moroccan	F	N	N	N	N			
HADA	21	French	M	Y	N	N	N			
GRCL	20	French	F	Y	N	N	N			
BOOU	20	Moroccan	M	Y	N	N	N			
SUCH	21	Chinese	M	Y	N	N	N			
LIKE	21	Chinese	F	Y	N	N	Y	1-2 times/W	10m-30m	Not specified
ZAMI	21	French	M	N	Extrasystoles; Tachycardia	N	Y	1 time / D	10m	Mindfulness meditation (Headspace application)
WAMI	22	Chinese	F	Y	N	N	N			
NIJO	22	Polish	F	N	N	N	N			
MOHU	52	French	M	Y	N	N	Y	1-2 times/D	30m	Taoist meditation; Buddhist meditation
HOPI	21	French	M	N	N	N	N			
ETLA	20	French	F	N	N	N	N			
CACA	20	Chilean	M	Y	N	N	N			
PRBA	20	French	M	N	N	N	N			

Figure 55: Information of all the subjects

where Y means yes, N means no ; as to gender, M means male, F means female ; as to frequency, M means month, D means day, W means week.

B Escape Dungeon

If the subject is a female: prince $< - >$ princess
The princess has fallen into a deep sleeping after touching the cursed spinning wheel. A hundred years have passed and a prince from another family spies the hidden castle during a hunting expedition. His attendants tell him differing stories regarding the castle: within the castle lies a beautiful princess who is doomed to sleep for a hundred years until a king's son comes and

awakes her. The prince then braves the tall trees, brambles and thorns which part at his approach, and enters the castle. However. . . you, the prince, cannot enter the princess's room unless you solve every challenge in the 5 upcoming rooms.

The first room is called: 24 room. You are given 4 cards, with a number on each card. You can only use addition, subtraction and multiplication, exactly one time each, to get the final result that should be 24. Your cards are: 4, 3, 8, 1 *(repeat)*. Of course you can change the order of the cards, but you must use all of them. How can you get 24 with these cards?

—45s—

The prince is very intelligent, and the moment he solves out the problem, the door leading to the next room is opened.

The second room is called: Give me four. What comes to your eyes are two non-graduated glass cups, their volumes are respectively 5L and 3L. Then there is a magical flower nearby that needs to be watered with exactly 4L all at once in order to bloom and trigger the opening of the door. You can use as much water from the tap as you need. How can you get 4L from these cups?

—45s—

The smart prince succeeds. So, the door leading to the next room is successfully opened.

The third room is called: The lost city. Several letters are incised on the wall: lawful, brilliant, intelligent, neat, ubiquitous, dauntless *(repeat)* The name of the lost city is hidden in the first letter of these six letters, which should be put in the right order. What name of a faraway city can you form with these letters?

—45s—

Wow amazing! It is his favorite city. He definitely wants to go there for one exchange semester. The door to the fourth room opens.

The fourth room is called: Unsolved age. There is a lock with a six-digits code on the door leading to the next room. The prince looks around and finds three statues that look like a family: a father, a mother and their daughter. A sign on the statue of the father reads : " I am 6 years older than my wife.". An other one on the statue of the mother reads : "I'm twice the age of my daughter. In 12 years, I will be 3 times her age." *(repeat)*. How old are the the father, the mother and the daughter?

—45s—

A piece of cake for our clever prince. The door opens, and through a glass wall he sees the beautiful princess lying on a bed. She is so beautiful that the prince wants to awake her as soon as possible. If he passes the fifth door, he can save the princess!

This final room is called: The mirror world. No lock here, but only entering on the right time the door can be opened. Be careful : if you are wrong, the princess will be constraint here forever... As you look up, you find a big clock on the wall. Its hour hand seems to point between 5 o'clock and 6 o'clock, and its minute hand points to 8 o'clock. Suddenly you remember the name of this room: the mirror world. At what time should you enter?

—45s—

And finally, the last door opens.

The prince comes across the chamber where the princess lies asleep. Struck by the radiant beauty before him, he falls on his knees before her. By a kiss, the enchantment comes to an end and the princess awakens...

C The Questionnaire for Meditation Phase and Story Phase

Meditation phase : what's your opinion?

Please answer the following questions according to how you felt during the previous phase.
For each affirmation, place the cursor depending on how strongly you agree.

1. **NAME First name**

2. **Respiration count :**

3. **You kept your focus during the whole phase.**
Une seule réponse possible.

	1	2	3	4	5	
Not true at all	⬭	⬭	⬭	⬭	⬭	Completely true

4. **Your mind often wandered out of your control.**
Une seule réponse possible.

	1	2	3	4	5	
Not true at all	⬭	⬭	⬭	⬭	⬭	Completely true

5. **You felt increasingly relaxed throughout the phase.**
Une seule réponse possible.

	1	2	3	4	5	
Not true at all	⬭	⬭	⬭	⬭	⬭	Completely true

6. **You felt increasingly self-aware throughout the phase.**
Une seule réponse possible.

	1	2	3	4	5	
Not true at all	⬭	⬭	⬭	⬭	⬭	Completely true

7. **You felt sleepy at times.**
Une seule réponse possible.

	1	2	3	4	5	
Not true at all	⬭	⬭	⬭	⬭	⬭	Completely true

8. **The duration of this phase was :**
Une seule réponse possible.

- ( ) Too short
- ( ) Appropriate
- ( ) Too long

9. **This experiment has made your day better:**
Une seule réponse possible.

	1	2	3	4	5	
Completely true	( )	( )	( )	( )	( )	Completely true

10. **Comments:**

Story phase : what's your opinion?

Please answer the following questions according to how you felt during the previous phase.
For each affirmation, place the cursor depending on how strongly you agree.

1. **NAME First name**

2. **Enigmas solved :**
   Plusieurs réponses possibles.
   - [ ] 1
   - [ ] 2
   - [ ] 3
   - [ ] 4
   - [ ] 5

3. **You kept your focus during the whole phase.**
   Une seule réponse possible.

	1	2	3	4	5	
Not true at all	○	○	○	○	○	Completely true

4. **You did NOT feel lost at any point during the phase.**
   Une seule réponse possible.

	1	2	3	4	5	
Not true at all	○	○	○	○	○	Completely true

5. **All questions were understandable.**
   Une seule réponse possible.

	1	2	3	4	5	
Not true at all	○	○	○	○	○	Completely true

6. **You felt intellectually stimulated by the questions.**
   Une seule réponse possible.

	1	2	3	4	5	
Not true at all	○	○	○	○	○	Completely true

7. **The difficulty and time given for each question were:**
Une seule réponse possible.

- ( ) Generally too hard/not enough time
- ( ) Appropriate
- ( ) Generally too easy/too much time
- ( ) Sometimes too hard, sometimes too easy
- ( ) Autre : ____________________

8. **Comments:**

____
____
____
____
____

D The Questionnaire for Meditation Phase

E Questionnaire Results

E.1 Results of feedback

In order to know how the subjects felt, 2 questionnaires were designed and submitted right after each corresponding phase. The answers to these questionnaires will provide additional data for future analysis as new variables to study the relation between their EEG signals and their personal feeling of relaxation. The questions were mainly Yes/No questions, to have more precision the answers can range from 1 (completely false) to 5 (completely right).

E.1.1 Questionnaire of the Story Phase

This questionnaire contains 5 questions : 1 on the subject's focus throughout the phase; 2 on their comprehension of the story and its enigmas. And 2 on the difficulty of the enigmas and how intellectually stimulated they felt.

Generally speaking, the story just reached to the expectation. All our test subjects rated their level of intellectual stimulation between 3 and 5. More than $\frac{3}{5}$ of them rated to 5.

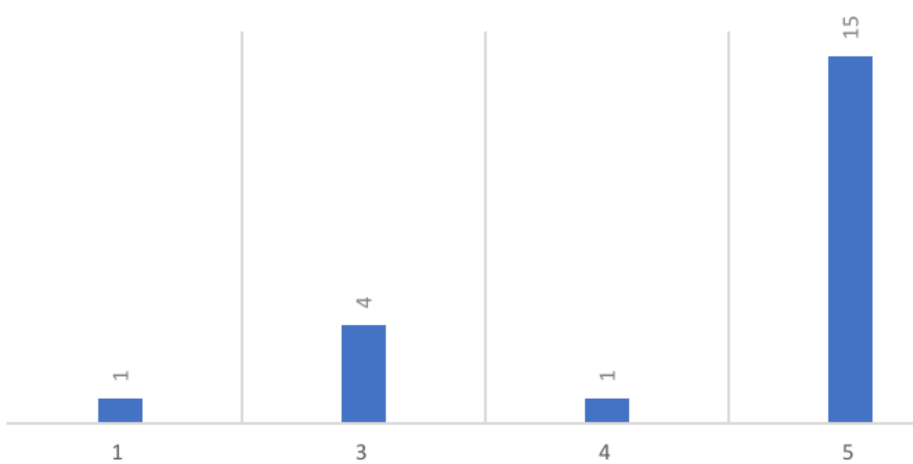


Figure 56: Level of Intellectual Stimulation

We got positive outcomes for the clearness of questions, since more than 95% of the test subjects thought that all the questions were very understandable (rated more than 4/5). The story successfully caught their attention, since about 95% of them considered they remained focused all the time (rated more than 4/5), with only 1 subject losing their focus during the story (rated 1/5).

In terms of difficulty and the time given for each question, more than half of our test subjects thought the difficulty and time are appropriate. If we take a closer look at the percentage of people who solved each question, there are still some difference in the level of difficulty for each question. We can find out that the last question was the easiest, with 60% of the subjects solving it. The 4^{th}

question seems to be the most difficult one with only 25% of subjects solved it. Now the fact that the subjects do not solve all the questions easily is positive, as well as the fact that they can solve at least one. This way, the enigmas are good challenges.

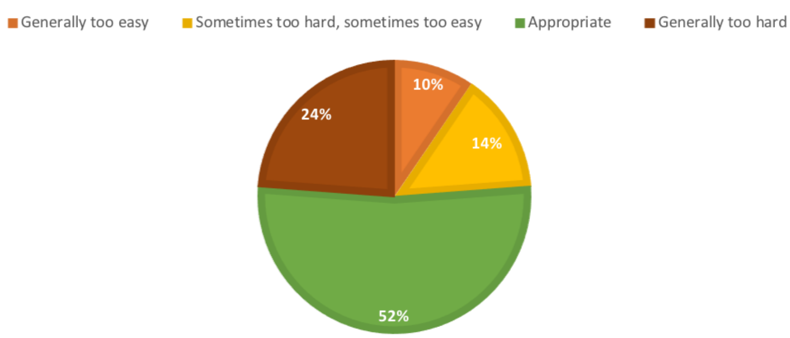


Figure 57: Difficulty of Each Question

E.1.2 Questionnaire of the Meditation Phase

This questionnaire contains 6 questions : 2 on the subject's focus throughout the phase, 3 on their personal feelings of relaxation, sleepiness and self-awareness, and 1 on the duration of the phase.

The duration of 10min for this meditation phase seemed appropriate for 86% of the test subjects.

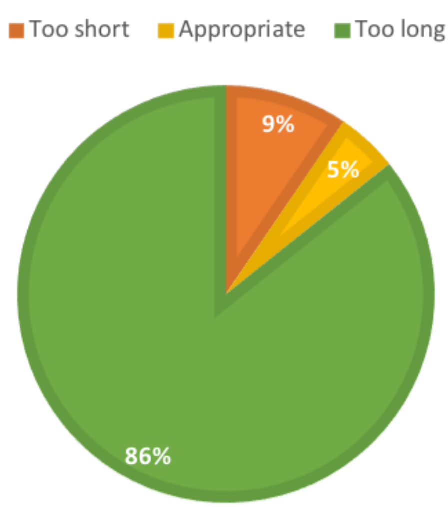


Figure 58: Duration of the Meditation phase

E.1.3 Questionnaire of the Meditation Phase

In terms of focus, all the subjects rated more than 3/5 to confirm that they remained focused during the whole phase. More than 80% of the test subject rated more than 4/5. However, this doesn't mean their mind was perfectly exempt from thoughts all the time. To the question: "You mind often wandered out of control", the answer follows a normal distribution, 3/5 being the most attributed grade. However, we need to be aware of the subjectivity of these questions. Mind wandering is a normal process during meditation, and everyone has a unique manner of handling it, which can lead to refocusing immediately or losing oneself in their thoughts.

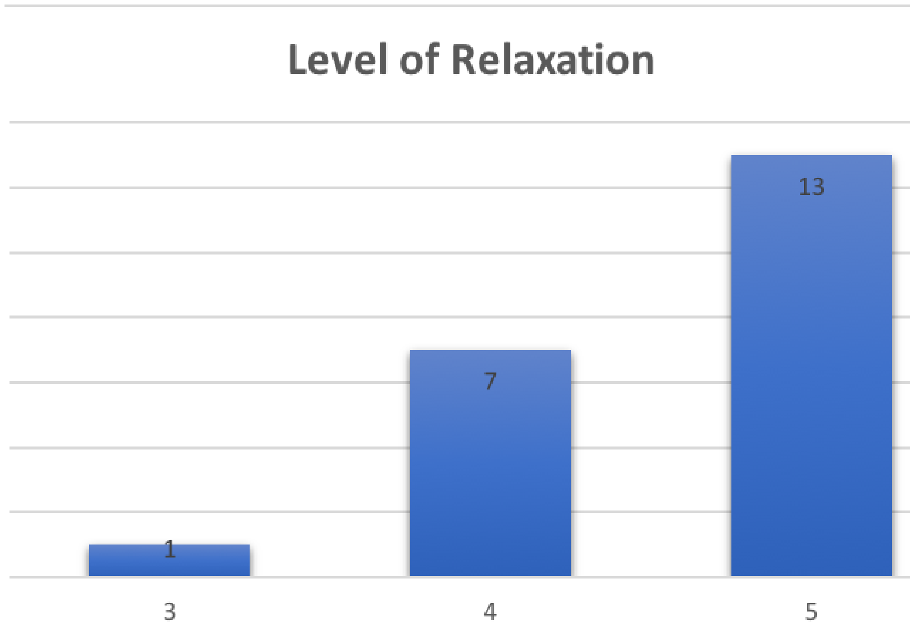


Figure 59: Duration of the Meditation phase

We also asked to the subject if they felt sleepy at times. The answer was not very consistent, but more than half of them reported feeling sleepy to some extent (rated more than 3/5).

We have very consistent result in terms of level of relaxation, as almost all the subjects felt more and more relaxed throughout the phase. This is very positive as for the efficiency of Su-Soku meditation intended for relaxation, even when performed by beginners.