

# IMPROVING EMOTIONAL WELL-BEING THROUGH NATURE

Ikhwan bin Amir Hamzah<sup>1</sup>, Nik Nur Wahidah binti Nik Hashim<sup>2\*</sup>, Nadzirah binti Ahmad Basri<sup>3</sup>

<sup>1</sup>Department of Mechatronics Engineering,  
Kulliyyah of Engineering,  
International Islamic University Malaysia

<sup>2</sup>Department of Mechatronics Engineering,  
Kulliyyah of Engineering,  
International Islamic University Malaysia

<sup>3</sup>Department of Psychiatry,  
Kulliyyah of Medicine,  
International Islamic University Malaysia

[ikhwan.amirhamzah@gmail.com](mailto:ikhwan.amirhamzah@gmail.com),  
[niknurwahidah@iium.edu.my](mailto:niknurwahidah@iium.edu.my), [nadzirahbasri@iium.edu.my](mailto:nadzirahbasri@iium.edu.my)

## Article history

Received: 10 Oct 2022

Received in revised form: 15 Oct 2022

Accepted: 2 Nov 2022

Published online: 10 Dec 2022

\*Corresponding author:  
[niknurwahidah@iium.edu.my](mailto:niknurwahidah@iium.edu.my)

## ABSTRACT

Nature may be used as a treatment method for patients with mental disorders which has yet to be implemented despite numerous evidence showing that being in touch with nature links to increased happiness, positive affect, positive social interactions, and a sense of meaning and purpose in life, as well as reduced mental distress. This project focuses on demonstrating the benefits of nature towards improving emotional well-being of people based on the detection of Electrodermal Activity (EDA). EDA signals will be collected from the controlled and experimental participants for the identification of nature's positive impact towards participants' mental well-being. Raw signals will require pre-processing and feature extraction which will use an open-source Python toolkit called PyEDA. The project will also incorporate machine learning for EDA signal classification which also utilizes python open-source software. The results will be observed and studied based on its accuracy in classifying the EDA signals between the controlled and experimental conditions. This project will be heading towards improving mental health treatment by proving that incorporating the world's most abundant resources, nature together with the incorporation of biosensors for mental health detection will help close the gap between mental health patients and clinicians.

**Keywords:** EDA signal, mental health, nature exposure and KNN algorithm

## 1.0 INTRODUCTION

Stress is defined as a state of mental tension and worry. However, in medical context, stress is described as a physical, mental or emotional factor that contributes to bodily or mental tension which can be triggered by either external or internal means [1]. As for mental health, the World Health Organization described it as "a state of well-being where a person realizes their own abilities, being able to cope with normal stresses of life, being able to work productively and is able to contribute to their community." This simply refers to how a person thinks, feels and

acts throughout their life which plays a huge role in our daily life and are interrelated to each other [2]–[4]. Stress is also much easier to treat compared to mental problems.

Since the pandemic, the number of people who suffers from mental health problems increases drastically [5]. Astro Awani news also reported that there were 336 suicide cases in the first quarter, doubling the previous years and 122,328 calls in the first 6 months of 2021 in which 89.4% calls were focusing on mental health issues demonstrating that the pandemic has greatly affected the public's mental state [6]. Consequently, this has affected the country in various institutions as well [7][8]. Also, people were disconnected with natural environment causing them to be nature-deficit. This prevents them from enjoying the physical and psychological benefits of nature including showing positive effects towards mental health and well-being [9]–[12].

There are multiple biosensors that can be used to detect bio signals of the human body. This includes Salivary Alpha Amylase (SAA), Electroencephalogram (EEG), Electrodermal Activity (EDA) or Galvanic Skin Response (GSR), Electrocardiogram (ECG) [13]–[20]. However, the biosensor that was chosen for this project is the Electrodermal Activity (EDA) which is due to its high accuracy despite using only two electrodes [21], [22].

The problem statement of this project is to prove the positive effect of nature towards the emotional well-being of individuals which will be use as a treatment method while also integrating biosensors as a mean of detecting mental illness. This allows patient to have the opportunity of better mental health treatment which caters with the low psychiatrist to population ratio as well as the lack of accuracy and continuous observation or assessment of patients' mental health.

The objectives of this project include to collect data on Electrodermal Activity (EDA) on the effect of nature towards well-being. Next, to analyse the signal differences between controlled and experimental conditions. Finally, to develop an algorithm that can identify the improved state of well-being through the effect of nature.

## **2.0 METHODOLOGY**

### **2.1 Experimental Setup**

Large number of data samples are needed to ensure higher accuracy in data recording thus the proposed number of data samples for this study will be at least 15 participants. Each participant will have their EDA reading recorded in both the controlled environment, a closed room and the experimental environment, nature environment. An elaborate detail of the data collection process in both conditions are shown below. In both conditions, participants are required to be relaxed to ensure maximum accuracy of the data. Participants will also be asked to fill in a survey to record their emotional experience before and after the whole process. The survey will be used to gather statistical data of the participants. Figure 1 below shows the example of controlled and experimental conditions of the project.

#### *CONTROLLED CONDITION (CLOSED ROOM)*

1. Take a walk inside a building for two minutes or more.
2. Fill in the questionnaire on participant's emotional state.
3. Record EDA reading for 1 minute.
4. Enter the closed room and sit for three minutes while facing an empty wall.
5. Record EDA reading while still facing the wall.
6. Fill in questionnaire on participant's current emotional state.

#### *EXPERIMENTAL CONDITION (IN NATURE)*

1. Take a walk inside a building for two minutes or more.
2. Fill in questionnaire on participant's current emotional state.
3. Record EDA reading for 1 minute.
4. Walk barefoot on a field while looking at the surroundings for three minutes.
5. Record EDA reading while sitting near the field while staying barefoot.
6. Fill in questionnaire on participant's current emotional state.



**Figure 1:** Example of experimental & controlled condition grounds

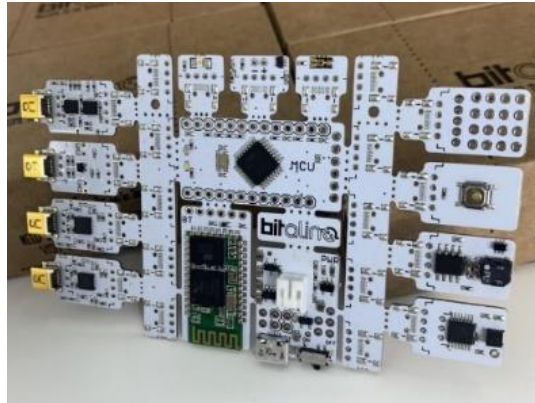
For this project, the participants will consist of university students in their 20s with an above average level of stress and no sign of mental illness. This is to ensure consistency of the recorded data and to ensure a distinct difference in data gathered in both conditions. The participants' mental health state will also be noted to guarantee no abnormalities of the data recorded and to ensure consistency of the experiment.

## **2.2 Bitalino (R)evolution Board Kit**

The sensor used for this project is the Bitalino (R)evolution Board Kit shown in Figure 2 below. It consists of multiple sensors such as ECG, EEG and EDA. However, as mentioned the EDA sensor was chosen due to its convenience and accuracy in recording emotional changes among its users. Bitalino is a sensor that does not require any coding or additional components such as a microcontroller or a display to be connected to operate. Instead, this sensor connects to a laptop or computer installed with the OpenSignals software via Bluetooth. The OpenSignals software will then display the obtained signal which will be immediately saved into the computer according to the time and date of the data recorded as shown in Table 1 below.

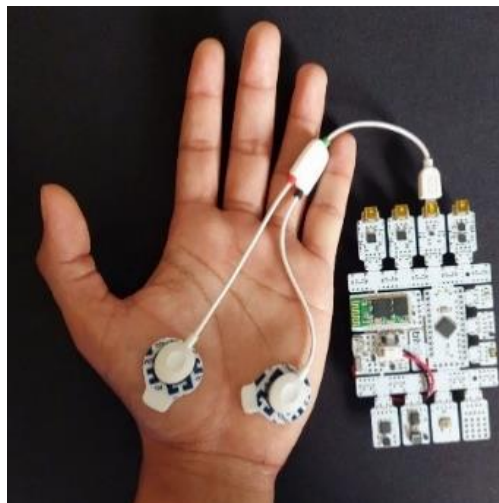
**Table 1:** File saving format

File name	Time	Date
opensignals_98D381FD5FBE_2021-12-13_23-15-15	11:15 p.m.	13.12.2021
opensignals_98D381FD5FBE_2021-11-01_22-36-08	10:36 p.m.	01.11.2021



**Figure 2:** Bitalino (r)evolution Board Kit

### 2.3 Electrode Placement



**Figure 3:** Placement of electrode patches on participant's palm

As shown in Figure 3 above, the electrode patches will be placed on the palm of the participant's non-dominant hand to ensure systematic data to be recorded between participants. The palm of the hand also has high sweat glands density making it a suitable placement for the electrode patches. It is also the most convenient place to apply the patches as it is not a body part that is usually covered unlike the shoulders or the foot.

### 2.4 Data Pre-Processing

In real world application, any recorded data that is based on a signal contain noises. These noises are important to be filtered as it may disrupt the result obtained from the data. This is also the reason why a proper data recording method is established to ensure minimum noises is captured during the process. Even with all the controlled experimental conditions, noises are still captured especially when it comes to physiological data such as the EDA. Any slight movement of the body effects the reading of the EDA creating noises which needs to be filtered [16].

Furthermore, different analysis algorithm requires different features of data which needs to be filtered allowing it to extract information effectively. Some of the suggested pre-processing methods are shown below. However, the methods listed below are not in any order and some may be omitted from the project depending on its signal to noise ratio after the pre-processing process.

#### *Median Filter*

Typically, the noises present in the raw data are non-negotiable and is difficult to control or minimize during the data collection process [16]. The noises are also harder to contained due to the high sampling rate used to record the EDA signal. To eliminate these noises, the median filter is applied which cleans the obtained signal while also retaining its features. Median filter is a filtering method that computes each sample of the recorded data into its median. This allows any extreme unwanted signals to be discarded thus providing a clean signal.

#### *Resampling*

For this project, the signal is sampled with a frequency of 1000 Hz as it matches the suitable sampling rate for the signal. The suitable sampling rate or the EDA signal varies from  $< 10$  Hz, 200 – 400 Hz and 1000 – 2000 Hz [23], [24]. This depends on the required features as different sampling rates provide different accuracy of the features. Since the sampling rate of the signal is quite high, resampling is done to reduce the signal's processing time and the memory footprint that the signal possesses.

#### *Moving Average*

Finally, moving average filter is also applied. Like the median filter, moving average also eliminates non-negligible noises that are present in the signal caused by body movements including respiration [25]. The moving average is the process of calculating the average of a signal across a certain period. For this project, the period used in the moving average will be 0.75 second. This process cleans the obtained signal while also help smoothens upcoming processes such as the feature extraction and classification.

## **2.5 Feature Extraction**

After the data went through pre-processing, the result obtained data is a clean data. Now, these data only contain its characteristics that needs to be extracted. These characteristics or features consume a lot of computational resources to process them into meaningful information which also consumes a lot of time. Therefore, to simplify the process, feature extraction is done which helps to group the large data and its characteristics into manageable sizes called features. The features to be extracted for this project consist of two aspect which are the EDA features and the statistical features which will be explained below.

**Table 2:** List of feature extraction process

No.	Features	
	EDA	Statistical
1.	Trough-to-peak	Mean
2.	Signal decomposition	Standard deviation
3.	Frequency-based	Skewness
4.	Time-frequency-based	Coefficient of variance
5.	-	Peak amplitude

### 2.5.1 EDA Features

Firstly, the signal undergoes the trough-to-peak analysis [26]. Trough and peak refer to the lowest and highest point of each cycle in a signal. After pre-processing, the derivative of the clean signal is applied with a threshold. This process reveals new signal response in the case where the incidents occur within an interval of more than one second from other signal responses. The trough and peak of the new signal responses will be labelled as the new zero crossing derived for the signals before and after the response respectively. This process helps to extract four features which are the number, overall magnitude, duration and area under the graph of the EDA signal.

The next feature extraction method will be the EDA signal decomposition. This is done to obtain the signal's tonic and phasic components. In this feature method, the components that will be calculated is the Discrete Deconvolution Analysis (DDA) utilising nonnegative deconvolution and the Continuous Deconvolution Analysis (CDA) which minimizes negativity [26], [27]. By applying these methods, the continuous tonic and phasic of the signal is obtained. Furthermore, statistical features for each of the components are also extracted. For instance, the average, standard deviation, minimum and maximum of the components. This has helped complete both the statistical and EDA features of this project altogether.

The use of frequency-based features to allow comprehensive understanding of the signal's transient characteristics can be achieved by analysing the signal's frequency domain representation [26], [28]. The utilization of Welch's algorithm for the estimation of the Short-time Fourier Transform (STFT), Fast Fourier Transform (FFT) and power spectral density (PSD) are the most common method. After transforming the processed signal into the



frequency domain, the data will undergo median filtering if it is deemed necessary [16]. To ensure that the frequency of the EDA signal will be analysed in detail, the frequency oscillation of the signal will be divided into several frequency sub bands. This is due to the frequency variation of the signal which is caused by the differences in physiological process that each participant experience despite being under the same conditions.

Finally, the time-frequency based feature is extracted from the signal which are the Discrete Wavelet Transform (DWT) and Stationary Wavelet Transform (SWT) [28]. This method is widely applied to analyse time and frequency information of bio signals [29]. DWT is used by sampling the wavelets discretely as the name suggest. However, it has its limitation whereby as the level of decomposition increases, the resolution of the signal reduces. This is where the SWT comes into place providing better results compared to the DWT in terms of the sampling rate and can retain the signal's full resolution during the decomposition process.

### 2.5.2 Statistical Features

There are several statistical data that will be extracted from the signal gathered which are listed in Table 3 below. The features will be extracted to help perform the statistical analysis of the project which will also be analysed together with the statistical data extracted from the questionnaire.

**Table 3:** List of statistical features to be extracted

Features	Definition	Standard Formula
Mean	Average/central value of a data	$\mu = \frac{\sum X}{N}$
Standard deviation	Indicator of how far a signal deviates from its average [30]	$\sigma = \sqrt{\frac{\sum (x_i - \mu)^2}{N}}$
Skewness	Measurement of deviation between the probability distribution and the normal distribution	$\tilde{\mu}_3 = \frac{\sum (X_i - \mu)^3}{(N - 1) * \sigma^3}$
Coefficient of variance	Measure of relative variability/ratio between standard deviation & mean [31]	$CV = \frac{\sigma}{\mu} \times 100$
Peak amplitude	Maximum deviation of the obtained signal either positive or negative from its zero-reference point	-

### 2.6 Statistical Analysis

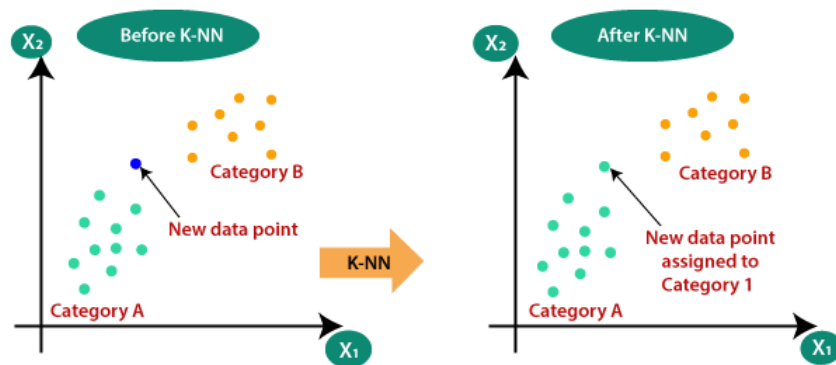
Next process will be the statistical analysis. It is a process of identifying trends and patterns of a data. Firstly, this process involves the analysis of the questionnaire which the participants were required to fill in during the experiment. This includes all the standard statistical analysis such as the mean, standard deviation, maximum and minimum of each question provided in the questionnaires. Similarly, the behaviour of all the participants in both controlled and experimental conditions will also be observed.

Secondly, the statistical analysis of the obtained signal will also be conducted. Like the questionnaires, this process also involves the group analysis and individual analysis of participants in both controlled and experimental conditions. This allows the participant's physiological readings in all conditions to be carefully analysed providing useful insights on their responses. The method of conducting this analysis is by using python programming utilising pyEDA library. By applying this method, the participant's physiological differences with and without nature interaction will be attained.

## 2.7 Classification

To complete the experiment, the classification process involving machine learning will be conducted. The classification method used for this project is based on supervised learning which is an algorithm that utilizes a well labelled data for its training process. The data that will be used for the learning method will be the recorded signals that had already been processed. These signals will also have its features extracted based to ensure the classification process runs smoothly.

This process uses the K-Nearest Neighbour (KNN) algorithm which is also one of the simplest classification algorithms. The way this algorithm works is that whenever a new data is input, it will compare the new data with the existing data and will categorized it according to the closest existing data. This also means that the algorithm does not learn from the existing data upon input but instead will only perform the classification when a new data is inserted which is also why it is called lazy learner algorithm. Additionally, KNN is also a suitable algorithm for dataset that contains many noises making it suitable for this project's application. The working principle of the algorithm is shown in Figure 4 below.



**Figure 4:** Example of the working principle of KNN [32]

To determine which category is the nearest to the new data point, the Euclidean distance formula is used which is shown below:

$$\text{Euclidean Distance} = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (1)$$



Based on the obtained results, the category that has the greatest number of nearest data points will be the category that the new data point will be assigned to. As shown in Figure 4 above, the new data point is assigned to Category A indicating that it has the greatest number of nearest data points to the new data.

### 3.0 RESULTS AND DISCUSSION

#### 3.1 Data Collection

The data collection process was conducted among fifteen (15) participants for the controlled and experimental conditions. The controlled condition took place in a hostel bedroom while the experimental condition took place in the hostel compound. The nature interaction method used for the experimental condition is known as grounding or earthing where participants were asked to walk on grass barefoot [33]. Their emotional states before and after their interaction with both conditions were also recorded using the EDA sensor and by filling in a questionnaire. Figure 5 to 7 below illustrates the data collection process including the participants' interaction with each environment. On the other hand, Figure 8 demonstrates the EDA signal displayed during the data recording process.



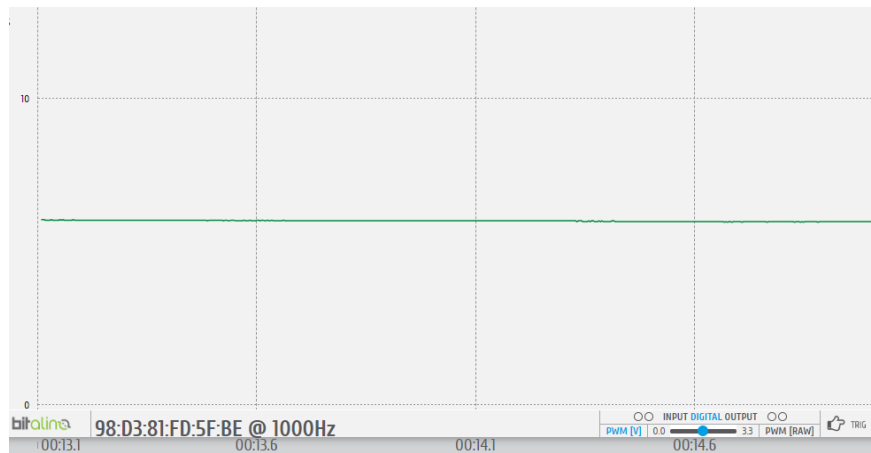
**Figure 5:** Data collection before both conditions



**Figure 6:** Participant's interaction with a closed room and data collection process (controlled condition)



**Figure 7:** Participant's interaction with nature via grounding and data collection process (experimental condition)



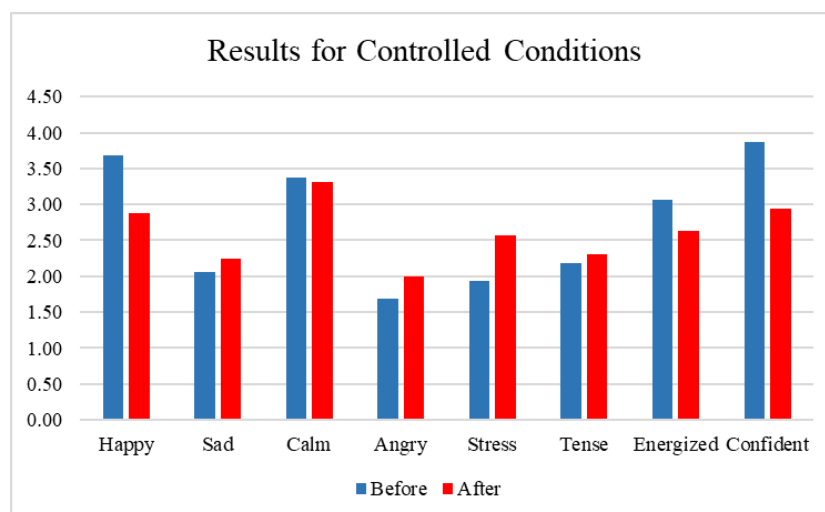
**Figure 8:** EDA signal displayed in OpenSignal software

### 3.2 Questionnaire Analysis

The statistical data of the signal gathered during the data collection process will be elaborate and tabulated. These data were obtained through questionnaires given before and after the participants interact with the tested surroundings. The questionnaire is conducted by inquiring the intensity (1 - 5) of participants' emotions such as happy, calm, angry and tense. The obtained results are averaged among all participants allowing the data to be statistically analysed.

To further explain the questionnaire, the rating of five (5) shows maximum intensity of all the listed emotions. For instance, very happy, very confident, very angry and very sad. On the other hand, this also demonstrates the rating of one (1) shows negative rating indicating lowest intensity of the emotions. For example, least happy and least angry.

#### 3.2.1 Controlled condition



**Figure 9:** Statistical data obtained from questionnaire for controlled condition

**Table 4:** Tabulated data for controlled condition

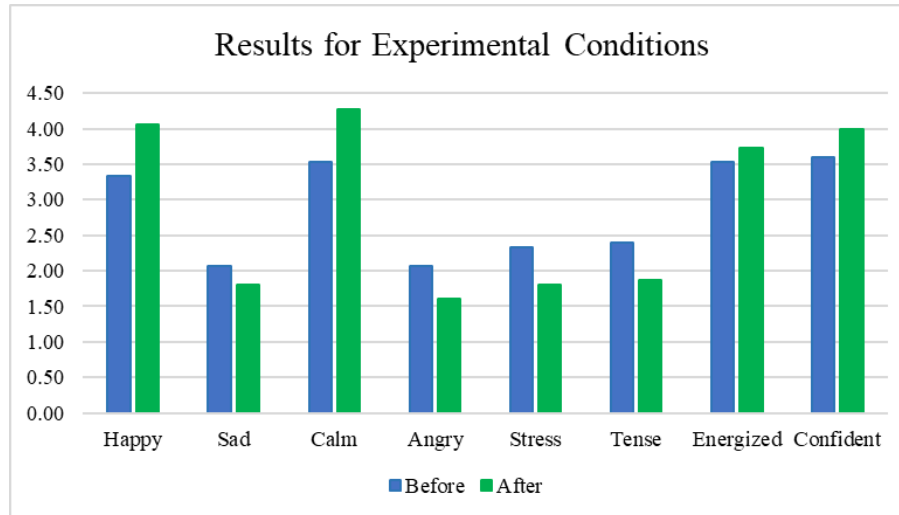
<b>ANALYSIS FOR CONTROLLED CONDITION</b>						
Emotions	Before			After		
	Mean	Variance	skew	Mean	Variance	skew
Happy	3.69	0.63	-0.25	2.88	0.78	-0.39
Energized	3.06	1.00	0.32	2.63	0.78	0.23
Confident	3.88	0.52	0.19	2.94	0.86	-0.43
Calm	3.38	0.78	0.43	3.31	1.56	-0.21
Sad	2.06	0.73	-0.13	2.25	1.67	0.74
Angry	1.69	0.63	0.66	2.00	1.87	1.08
Stress	1.94	0.86	0.71	2.56	2.13	0.45
Tense	2.19	1.10	0.38	2.31	1.96	0.52

Figure 9 and Table 4 above shows the statistical data for the controlled condition of the project. The data is separated into two emotions, positive and negative. Based on the data gathered before participants' interaction with the closed room, their rating of positive emotions was quite neutral while also leaning slightly towards intense emotions ranging from **3.06** to **3.88**. On the other hand, negative emotions indicated the opposite trend of slightly low emotion ratings ranging from **1.69** to **2.19**.

Post interaction with the closed room, participants' positive emotion ratings decrease to a mild rating, **2.63** to **3.31**. As for negative emotions, it increases slightly with a range of **2.00** to **2.56**. This indicates that participants' interaction with the closed room has trigger negative emotions which can be seen based on the slight increase in negative emotions as well as a slight decrease in positive emotions of participants. Additionally, the variance and skew before participants' interaction with the closed room is ranged from **0.52** to **1.10** and **-0.25** to **0.71** respectively. While after participants' interaction with the closed room, their variance and skew are ranging from **0.78** to **2.13** and **-0.43** to **1.08** respectively.

Furthermore, individual observation of the emotions demonstrates a similar trend as describe above. Each positive emotions shows a decrease in the averages while the negative emotions increase after the closed room interaction process. However, each emotion experience different degree of increment or decrement. Some might have a higher or lower degree compared to the other. This depends on each participant's opinion on their emotional experience throughout the experiment. Some might have a higher intensity of emotions while some might have lower intensity. Nonetheless, the result shown above clearly explain that confinement of participants inside a closed room has shown to trigger negative emotions.

### 3.2.2 Experimental condition



**Figure 10:** Statistical data obtained from questionnaire for experimental condition

**Table 5:** Tabulated data for experimental condition

ANALYSIS FOR EXPERIMENTAL CONDITION						
Emotions	Before			After		
	Mean	Variance	skew	Mean	Variance	skew
Happy	3.33	0.81	0.58	4.07	0.50	-0.09
Energized	3.53	0.98	0.15	3.73	1.07	-1.18
Confident	3.60	0.83	0.31	4.00	0.86	-0.62
Calm	3.53	0.84	-0.11	4.27	0.50	-0.43
Sad	2.07	0.78	-0.14	1.80	0.74	1.20
Angry	2.07	0.78	-0.14	1.60	0.40	0.55
Stress	2.33	1.38	0.16	1.80	0.31	-0.11
Tense	2.40	0.97	0.06	1.87	0.41	0.10

Figure 10 and Table 5 above demonstrate the statistical data obtained from the experimental conditions. Similarly, the data is divided into positive and negative emotions. The obtained results before participants' interaction with nature demonstrates that positive and negative emotions have similar trend to the controlled conditions before interaction with the closed room. The positive emotions have a rating ranging from **3.33** to **3.60** while negative emotions are on the low range between **2.07** to **2.40**. These indicate participants' emotions were relatively neutral and are feeling slightly positive.

After interaction with nature, participants' emotional rating increases in positive emotions ranging from **3.75** to **4.27** while negative emotions decrease with a range of **1.60** to **1.87**. These results indicated that positive emotions were triggered after the participants interacted with nature. This demonstrated that interaction with natural environment i.e., grass,

earth and sky has proven to positively impact their emotions and reduce their negative emotions. This aligns with previous studies conducted on this topic demonstrating the positive impact of nature towards human emotional wellbeing after their interaction with nature. In addition, the variance and skew of participants before their interaction with the closed room is ranged from **0.78** to **1.38** and **-0.14** to **0.58** respectively. While after participants' interaction with the closed room, their variance and skew are ranged from **0.31** to **1.07** and **-1.18** to **1.20** respectively.

Next, individual trends of emotions before and after the participants interacted with nature shows that each positive emotion experiences a significant increase in averages while negative emotions decrease. Correspondingly, the intensity of the increment and decrement of these emotions differs from each other which depends on the participants' sensitivity towards the experiment. The most significant difference in these averages comes from 'calm' and 'happy' with an increment of **0.74** while the lowest comes from 'energized' with an increment of **0.2** only.

### **3.3 Data Pre-Processing & Feature Extraction**

This section of the report demonstrates the results of the feature extraction process conducted using the gathered EDA signals. The process took place using a python library obtained from GitHub known as pyEDA which performs both the EDA and statistical feature extraction. Before performing these processes, the program will also perform data pre-processing allowing clean signals to be input for the feature extraction processes.

Before performing these processes, the data is separated into four (4) which helps with the processing of the signals as the whole data of 1 minute was not able to be processed due to hardware constraints. Furthermore, the instead of using the local IDE software, Google Colab was chosen as it provides better processing power to perform the program due to the application of cloud-based GPU and RAM. This allows the program to be executed smoothly without the hardware limitations.

#### *3.3.1 Data Pre-Processing*

The pre-processing process involves three methods which are the data resampling, moving average filter and the extraction of phasic and tonic components of the EDA signals. For resampling, the signal is down sampled from 1000 Hz at which the signal is recorded, to a lower frequency of 250 Hz. This method helps to speed up the signal processing time of the program while also retaining its features. Moving on, the next process involves the moving average filter. This process is conducted by setting the window size of the signals in which the window size is set to be 0.75 seconds. The program will then calculate the average filter by multiplying the rescale sampling rate with the window size. These processes were performed using the built-in library of pyEDA which are executed during the statistical and EDA feature extraction process.



### 3.3.2 EDA Feature Extraction

Moving on, the program then executes the EDA feature extraction process. As stated earlier, the feature extraction is performed immediately after the data pre-processing processes is conducted. In this section, the data is trained using deep learning to identify the best feature to be extracted for the signal. This is done by performing the learning process for 100 times or in deep learning terms, 100 epochs. The sample of the training process is shown in Figure 11 below.

[35 35 35 ... 33 34 33]	epoch : 34/100, loss = 10.497496	epoch : 67/100, loss = 0.268094
epoch : 1/100, loss = 1186.589355	epoch : 35/100, loss = 5.069577	epoch : 68/100, loss = 0.398649
epoch : 2/100, loss = 817.381697	epoch : 36/100, loss = 1.455881	epoch : 69/100, loss = 0.441344
epoch : 3/100, loss = 466.063812	epoch : 37/100, loss = 0.290306	epoch : 70/100, loss = 0.380674
epoch : 4/100, loss = 188.853256	epoch : 38/100, loss = 1.322827	epoch : 71/100, loss = 0.260490
epoch : 5/100, loss = 33.725910	epoch : 39/100, loss = 3.630501	epoch : 72/100, loss = 0.123451
epoch : 6/100, loss = 2.765636	epoch : 40/100, loss = 5.974854	epoch : 73/100, loss = 0.040933
epoch : 7/100, loss = 54.880554	epoch : 41/100, loss = 7.330404	epoch : 74/100, loss = 0.031009
epoch : 8/100, loss = 130.565948	epoch : 42/100, loss = 7.179513	epoch : 75/100, loss = 0.077560
epoch : 9/100, loss = 184.869598	epoch : 43/100, loss = 5.762581	epoch : 76/100, loss = 0.140781
epoch : 10/100, loss = 198.744446	epoch : 44/100, loss = 3.669482	epoch : 77/100, loss = 0.180589
epoch : 11/100, loss = 176.688324	epoch : 45/100, loss = 1.689039	epoch : 78/100, loss = 0.176215
epoch : 12/100, loss = 131.617874	epoch : 46/100, loss = 0.447125	epoch : 79/100, loss = 0.132979
epoch : 13/100, loss = 80.031342	epoch : 47/100, loss = 0.193656	epoch : 80/100, loss = 0.075003
epoch : 14/100, loss = 36.242584	epoch : 48/100, loss = 0.764338	epoch : 81/100, loss = 0.029990
epoch : 15/100, loss = 8.516217	epoch : 49/100, loss = 1.708912	epoch : 82/100, loss = 0.013935
epoch : 16/100, loss = 0.708885	epoch : 50/100, loss = 2.512834	epoch : 83/100, loss = 0.026311
epoch : 17/100, loss = 9.600560	epoch : 51/100, loss = 2.811468	epoch : 84/100, loss = 0.052301
epoch : 18/100, loss = 27.974636	epoch : 52/100, loss = 2.519264	epoch : 85/100, loss = 0.073157
epoch : 19/100, loss = 46.960854	epoch : 53/100, loss = 1.783342	epoch : 86/100, loss = 0.080444
epoch : 20/100, loss = 59.225418	epoch : 54/100, loss = 0.950053	epoch : 87/100, loss = 0.064481
epoch : 21/100, loss = 61.316284	epoch : 55/100, loss = 0.327100	epoch : 88/100, loss = 0.038888
epoch : 22/100, loss = 53.428284	epoch : 56/100, loss = 0.093196	epoch : 89/100, loss = 0.016860
epoch : 23/100, loss = 38.949928	epoch : 57/100, loss = 0.237564	epoch : 90/100, loss = 0.007472
epoch : 24/100, loss = 22.817768	epoch : 58/100, loss = 0.595109	epoch : 91/100, loss = 0.011609
epoch : 25/100, loss = 9.486883	epoch : 59/100, loss = 0.941695	epoch : 92/100, loss = 0.022958
epoch : 26/100, loss = 1.894878	epoch : 60/100, loss = 1.099807	epoch : 93/100, loss = 0.032672
epoch : 27/100, loss = 0.770482	epoch : 61/100, loss = 1.007328	epoch : 94/100, loss = 0.034570
epoch : 28/100, loss = 4.884544	epoch : 62/100, loss = 0.725071	epoch : 95/100, loss = 0.027976
epoch : 29/100, loss = 11.487015	epoch : 63/100, loss = 0.388536	epoch : 96/100, loss = 0.017066
epoch : 30/100, loss = 17.436239	epoch : 64/100, loss = 0.134539	epoch : 97/100, loss = 0.007695
epoch : 31/100, loss = 20.704351	epoch : 65/100, loss = 0.049945	epoch : 98/100, loss = 0.003861
epoch : 32/100, loss = 19.967884	epoch : 66/100, loss = 0.119626	epoch : 99/100, loss = 0.005840
epoch : 33/100, loss = 16.079187		epoch : 100/100, loss = 0.010706

**Figure 11:** Training process for EDA feature extraction

In the program, the number of losses is also inserted while running the training process. This is to allow monitoring of the feature extraction training to ensure minimal losses is achieved. Based on the training conducted, the range of losses for the whole data set is less than one (1) which demonstrates that almost all the dataset is being used to train the data. This also portrays the accuracy of the EDA feature that was extracted by the program. The extracted feature is then saved into Excel files which will be used as training data for the machine learning process.

### 3.3.3 Statistical Feature Extraction

The next process is the statistical feature extraction of the EDA signal. Here, the statistical features that was extracted includes the number of peaks, maximum value of the peaks, average of EDA signals and more. This process also extracts the tonic and phasic components of the EDA signal. The coding command for this section is shown below.

```
m,wd,eda_clean=process_statistical(eda,use_scipy =
True,sample_rate=1000,new_sample_rate=250,segment_width=0,segment_overlap=0)
```

From the command above, a new sampling rate will be obtained from 1000 Hz to 250 Hz. The signal segmentation is done for 10 seconds. The output 'm' is the number of peaks, the maximum value of peaks and the mean of the signal, 'wd' contains the filtered phasic and tonic components of the EDA and finally 'eda\_clean' stores the processed signal. These outputs will be displayed in the command prompt of the program and was not saved into the local hardware. The obtained statistical feature of the signal is shown in Figure 12 below. The features were labelled as pcost, dcost, gap, pres and dres whereby, similar to the EDA features, the optimal solution for these features were obtained to ensure the best feature was extracted.

	pcost	dcost	gap	pres	dres		pcost	dcost	gap	pres	dres
0:	-7.0881e+07	-7.0879e+07	1e+04	1e+02	1e+00	0:	-3.9279e+07	-3.9277e+07	7e+03	8e+01	3e+00
1:	-7.0883e+07	-7.0886e+07	5e+03	6e+01	6e-01	1:	-3.9281e+07	-3.9283e+07	4e+03	5e+01	2e+00
2:	-7.0885e+07	-7.0889e+07	5e+03	5e+01	5e-01	2:	-3.9283e+07	-3.9285e+07	4e+03	4e+01	2e+00
3:	-7.0887e+07	-7.0889e+07	2e+03	8e+00	9e-02	3:	-3.9292e+07	-3.9306e+07	1e+04	7e+00	3e-01
4:	-7.0887e+07	-7.0888e+07	8e+02	3e+00	3e-02	4:	-3.9292e+07	-3.9301e+07	9e+03	4e+00	2e-01
5:	-7.0888e+07	-7.0888e+07	3e+02	7e-01	7e-03	5:	-3.9293e+07	-3.9300e+07	7e+03	2e+00	8e-02
6:	-7.0888e+07	-7.0888e+07	1e+02	2e-01	2e-03	6:	-3.9293e+07	-3.9295e+07	2e+03	5e-01	2e-02
7:	-7.0888e+07	-7.0888e+07	5e+01	5e-02	6e-04	7:	-3.9293e+07	-3.9294e+07	6e+02	1e-01	4e-03
8:	-7.0888e+07	-7.0888e+07	2e+01	5e-03	6e-05	8:	-3.9293e+07	-3.9293e+07	1e+02	1e-02	4e-04
9:	-7.0888e+07	-7.0888e+07	1e+01	3e-03	4e-05	9:	-3.9293e+07	-3.9293e+07	4e+01	3e-03	1e-04
/usr/local/lib/python3.7/dist-packages/pandas/core						10:	-3.9293e+07	-3.9293e+07	9e+00	5e-04	2e-05
values = np.array([convert(v) for v in values])						11:	-3.9293e+07	-3.9293e+07	8e+00	4e-04	2e-05
10:	-7.0888e+07	-7.0888e+07	6e+00	1e-03	1e-05	12:	-3.9293e+07	-3.9293e+07	3e+00	1e-04	4e-06
11:	-7.0888e+07	-7.0888e+07	2e+00	2e-04	3e-06	13:	-3.9293e+07	-3.9293e+07	2e+00	7e-05	3e-06
12:	-7.0888e+07	-7.0888e+07	1e+00	1e-04	1e-06	14:	-3.9293e+07	-3.9293e+07	8e-01	2e-05	7e-07
13:	-7.0888e+07	-7.0888e+07	3e-01	2e-05	2e-07	15:	-3.9293e+07	-3.9293e+07	7e-01	1e-05	5e-07
14:	-7.0888e+07	-7.0888e+07	2e-01	9e-06	9e-08	16:	-3.9293e+07	-3.9293e+07	2e-01	3e-06	1e-07
15:	-7.0888e+07	-7.0888e+07	5e-02	2e-06	2e-08	17:	-3.9293e+07	-3.9293e+07	1e-01	1e-06	4e-08
16:	-7.0888e+07	-7.0888e+07	3e-02	3e-07	4e-09	18:	-3.9293e+07	-3.9293e+07	3e-02	9e-08	4e-09
17:	-7.0888e+07	-7.0888e+07	4e-03	5e-08	5e-10	Optimal solution found.					

**Figure 12:** Statistical feature results from pyEDA

### 3.4 Machine Learning (KNN Algorithm)

Finally, the experiments continue with the machine learning process which uses the KNN algorithm. Similarly, the algorithm is also performed on Google Colab to allow better processing from the cloud-based GPU and RAM. The classification method is performed based on the labelled data whereby both gathered signals, controlled and experimental signals will be labelled before training the algorithm.

After labelling, the data will be trained to classify the differences between the controlled and experimental conditions. The algorithm will also split the data into two which is the training dataset and the testing dataset. To perform the learning process, several variables were set which includes the number of neighbours,  $n\_neighbours=2$ , the test size,  $test\_size=.5$  and the random state,  $random\_state=9$ . This is performed using two lines of command which is shown below.

```
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=.5, random_state=9)
knn = KNeighborsClassifier(n_neighbors=2)
```

The testing data will then be used to identify the accuracy of the algorithm in which for this experiment, obtains 52% only as shown in Figure 13 below. This demonstrates that the classification algorithm was only able to classify the tested dataset correctly only half of the time. This can also be observed by the list of accurate (*acc*) and inaccurate (*unacc*) training results displayed in Figure 13 below which is written to study accuracy of the predictions made by the algorithm.

```
Predicted: unacc data (1107,) actual: acc
Predicted: unacc data (1652,) actual: unacc
Predicted: unacc data (855,) actual: acc
Predicted: unacc data (1804,) actual: unacc
Predicted: unacc data (1214,) actual: unacc
Predicted: unacc data (1198,) actual: acc
Predicted: acc data (2280,) actual: acc
Predicted: unacc data (2532,) actual: unacc
Predicted: unacc data (3796,) actual: acc
Predicted: acc data (3239,) actual: acc
Predicted: unacc data (214,) actual: unacc
Predicted: acc data (1729,) actual: unacc
Predicted: unacc data (3619,) actual: unacc
Predicted: unacc data (439,) actual: unacc
Predicted: unacc data (240,) actual: unacc
Predicted: unacc data (3296,) actual: unacc
Predicted: unacc data (3074,) actual: acc
Predicted: acc data (473,) actual: acc
Predicted: acc data (1941,) actual: acc
0.5262096774193549
```

**Figure 13:** Prediction lists and accuracy of KNN algorithm

## 4.0 CONCLUSION

### 4.1 Conclusion

To conclude, this project has achieved all its objectives. This includes to study nature's effect towards the well-being by using EDA signals. It is also proven that interaction with natural environment provides positive impacts towards the well-being of individuals who interacts with it. Not to mention, this theory is proven with both statistical data through questionnaire and the EDA signal.

Next, the signal differences between the controlled and experimental conditions are also observed and tabulated successfully. Finally, the algorithm that is capable in identifying the improved state of well-being through the effect of nature was constructed successfully. All in all, this project has proven to be successful in achieving all its objectives.

### 4.2 Future Works

For future works, this study should include some of these recommendations. Firstly, since classification process will be performed, a bigger dataset should be gathered as classification relies heavily on number of datasets in which the greater the number of datasets, the better. The dataset should include a wider scope of individuals to help identify the impact of nature's intervention among individuals of various age groups and comprises of both male

and female. Next, multiple classification algorithms should also be used. The results of these algorithms should also be compared with each other to determine the algorithm with the best model and highest accuracy in classifying the EDA signals.

#### 4.3 Limitations

The most prominent limitation for this experiment is the machine learning algorithm as it was only able to display an accuracy of 52%. Next, hardware limitations are also a concern as classification algorithms require huge memory and processing power which the hardware could not cope. Not to mention, the data collection process also contains interference by people passing by as it was performed in public areas.

#### 4.4 Acknowledgement

The authors would like to thank International Islamic University Malaysia for granting the opportunity of performing this project until its completion. Additionally, the authors would also like to convey deep appreciation to those who directly and indirectly involved in this project.

#### REFERENCES

- [1] Adam Felman and Dillion Browne (2020). Anxiety: Symptoms, types, causes, prevention, and treatment. *Medical News Today*. <https://www.medicalnewstoday.com/articles/323454> (accessed Jan. 20, 2022).
- [2] Mind (2017). What is stress? Mind, the mental health charity - help for mental health problems. <https://www.mind.org.uk/information-support/types-of-mental-health-problems/stress/what-is-stress/> (accessed Jan. 20, 2022).
- [3] NIMH. NIMH: I'm So Stressed Out! Fact Sheet. <https://www.nimh.nih.gov/health/publications/so-stressed-out-fact-sheet> (accessed Jan. 19, 2022).
- [4] Sara Lindberg and Timothy Legg (2019). Psychological Stress: Symptoms, Causes, Treatment & Diagnosis. <https://www.healthline.com/health/psychological-stress> (accessed Jan. 19, 2022).
- [5] Li P. Wong, Haridah Alias, Afiah A. M. Fuzi, Intan S. Omar, Azmawaty M. Nor, Maw P. Tan, Diana L. Baranovich, Che Z. Saari, Sareena H. Hamzah, Ku W. Cheong, Chiew H. Poon, Vimala Ramoo, Chong C. Che, Kyaimon Myint, Suria Zainuddin, Ivy Chung. (2021). Escalating progression of mental health disorders during the COVID-19 pandemic: Evidence from a nationwide survey. *Public Library of Science (PLoS ONE)* 16(3).

- [6] Amanda Yeoh (2021). Mental health is critical and need to be addressed during trying times | Astro Awani. <https://www.astroawani.com/berita-malaysia/mental-health-critical-and-need-be-addressed-during-trying-times-306105> (accessed Jan. 20, 2022).
- [7] Mental health: strengthening our response, *World Health Organization [WHO]* (2018). <https://www.who.int/news-room/fact-sheets/detail/mental-health-strengthening-our-response> (accessed Jan. 21, 2022).
- [8] Anuradha Mansinghka and Namratha Mohan (2021). The effects of low productivity on business growth - Saviom. <https://www.saviom.com/blog/effects-of-low-productivity-business-growth/> (accessed Jan. 21, 2022).
- [9] Madhuleena Roy Chowdhury (2021). The Positive Effects of Nature on Your Mental Well-Being. <https://positivepsychology.com/positive-effects-of-nature/> (accessed Jan. 21, 2022).
- [10] Michelle Tester-Jones, Mathew P. White, Lewis R. Elliott, Netta Weinstein, James Grellier, Theo Economou, Gregory N. Bratman, Anne Cleary, Mireia Gascon, Kalevi M. Korpela, Mark Nieuwenhuijsen, Aisling O'Connor, Ann Ojala, Matilda V. D. Bosch and Lora E. Fleming (2020). Results from an 18-country cross-sectional study examining experiences of nature for people with common mental health disorders, *Sci Rep*, 10(1).
- [11] Hilal Kahveci and Parisa Göker (2019). Mental health and wellbeing; Ecotherapy. *3<sup>rd</sup> International Symposium on Innovative Approaches in Scientific Studies*, 4(3), 308-311.
- [12] Dorothy Ibes, Isabel Hiram, and Carolyn Schuyler (2018). Greenspace ecotherapy interventions: The stress-reduction potential of green micro-breaks integrating nature connection and mind-body skills. *Ecopsychology*, 10(3), 137–150.
- [13] Nida Ali and Urs M. Nater (2020). Salivary Alpha-Amylase as a Biomarker of Stress in Behavioral Medicine, *Int J Behav Med*, 27(3), 337–342.
- [14] Ravindranath Vineetha, Keerthilatha-M. Pai, Manoj Vengal, Kodyalamoole Gopalakrishna and Dinesh Narayanakurup (2014). Usefulness of salivary alpha amylase as a biomarker of chronic stress and stress related oral mucosal changes - a pilot study,” *J Clinical and Experimental Dentistry*, 6(2), 132-137.
- [15] Houtan Jebelli Sungjoo Hwang and SangHyun Lee (2018). EEG-based workers’ stress recognition at construction sites, *Automation Construction*, 93, 315–324.



- [16] Rui Guo, Shuangjiang Li, Li He, Wei Gao, Hairong Qi, and Gina Owens (2013). Pervasive and unobtrusive emotion sensing for human mental health, *Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2013*. 436–439.
- [17] Antonio Lanata, Alberto Greco, Gaetano Valenza, and Enzo P. Scilingo (2014). A Pattern Recognition Approach Based on Electrodermal Response for Pathological Mood Identification in Bipolar Disorders, *ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings*, 3601–3605.
- [18] Lena Biel, Ola Pettersson, Lennart Philipson, and Peter Wide (2001). ECG analysis: A new approach in human identification, *IEEE Transactions on Instrumentation and Measurement*, 50(3), 808–812.
- [19] Shitong Huang, Jia Li, Pengzhu Zhang, and Weiqiang Zhang (2018). Detection of mental fatigue state with wearable ECG devices,” *International Journal of Medical Informatics*, 119, 39–46.
- [20] S. Tivatansakul and M. Ohkura (2015). Improvement of emotional healthcare system with stress detection from ECG signal, *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 6792–6795.
- [21] Marco Sarchiapone, Carla Gramaglia, Miriam Iosue, Vladimir Carli, Laura Mandelli, Alessandro Serretti, Debora Marangon, Patrizia Zeppegno (2018). The association between electrodermal activity (EDA), depression and suicidal behaviour: A systematic review and narrative synthesis, *BMC Psychiatry*, 18(1).
- [22] Diana Batista, Hugo Plácido da Silva, Ana Fred, Carlos Moreira, Margarida Reis, Hugo Alexandre Ferreira (2019), Benchmarking of the BITalino biomedical toolkit against an established gold standard, *Healthcare Technology Letters*, 6(2), pg. 32-36.
- [23] Jason J. Braithwaite, Derrick, G. Watson, Robert Jones, and Mickey Rowe (2015). A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments via the BIOPAC MP36R & AcqKnowledge software.
- [24] Gersak Gregor and Dmovsek Janko (2020). Electrodermal activity Patient Simulator, *Public Library of Science (PLOS ONE)*, 15(2).



- [25] Seyed A. H. Aqajari, Emad K. Naeini, Milad A. Mehrabadi, Sina Labbaf, Nikil Dutt, Amir M. Rahmani (2021). Pyeda: An open-source python toolkit for pre-processing and feature extraction of electrodermal activity, *Procedia Computer Science*, 184, 99–106.
- [26] Erika Lutin, Ryuga Hashimoto, Walter D. Raedt, and Chris V. Hoof (2021). Feature extraction for stress detection in electrodermal activity, *BIOSIGNALS 2021 - 14th International Conference on Bio-Inspired Systems and Signal Processing; Part of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies, BIOSTEC 2021*, 177–185.
- [27] Alberto Greco, Gaetano Valenza, Antonio Lanata, Giuseppina Rota, Enzo Pasquale Scilingo (2014). Electrodermal activity in bipolar patients during affective elicitation, *IEEE J Biomed Health Inform*, 18(6), 1865–1873.
- [28] Jainendra Shukla, Miguel Barreda-Angeles, Joan Oliver, G. C. Nandi, and Domenec Puig (2021). Feature Extraction and Selection for Emotion Recognition from Electrodermal Activity, *IEEE Transactions on Affective Computing*, 12(4), 857–869.
- [29] Fauzani N. Jamaluddin, Siti A. Ahmad, Samsul B. M. Noor, Wan Z. W. Hassan, Azhar Yaacob and Yunus Adam (2015). Performance of DWT and SWT in muscle fatigue detection, *ISSBES 2015 IEEE Student Symposium in Biomedical Engineering & Sciences*.
- [30] Steven W. Smith. The Scientist and Engineer's Guide to Digital Signal Processing Statistics, Probability and Noise, Chapter 2: Statistics, Probability and Noise, *California Technical Publishing*, 11-34.
- [31] How to Find a Coefficient of Variation - Statistics How To. <https://www.statisticshowto.com/probability-and-statistics/how-to-find-a-coefficient-of-variation/> (accessed Jan. 23, 2022).
- [32] K-Nearest Neighbor (KNN) Algorithm for Machine Learning - Javatpoint. <https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning> (accessed Jan. 23, 2022).
- [33] Eleesha Lockett and Debra R. Wilson (2019). What Is Grounding and Can It Help Improve Your Health? <https://www.healthline.com/health/grounding> (accessed Jan. 23, 2022).