Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

School of Engineering and Sciences

One Step Closer to Mental Health: Resilience to Mental Stress Index in the Face of Analytical Problems, A Machine Learning Approach

A thesis presented by

Ramón Eduardo Díaz Ramos

Submitted to the
School of Engineering and Sciences
in partial fulfillment of the requirements for the degree of

Master of Science

in

Computer Science

Monterrey, Nuevo León, May, 2021

Instituto Tecnológico y de Estudios Superiores de Monterrey

Campus Monterrey

School of Engineering and Sciences

The committee members, hereby, certify that have read the thesis presented by Ramón Eduardo Díaz Ramos and that it is fully adequate in scope and quality as a partial requirement for the degree of Master of Science in Computer Science.

PhD Luis Ángel Trejo Rodríguez
Tecnológico de Monterrey
Principal Advisor

PhD Miguel Angel Medina Pérez
Tecnológico de Monterrey
Committee Member

PhD Miguel Gonzalez Mendoza
Tecnológico de Monterrey
Committee Member

PhD Carlos Gonzalo Figueroa López
Universidad Autónoma de México
Committee Member

PhD Rubén Morales Menéndez
Dean of Graduate Studies
School of Engineering and Sciences

Monterrey, Nuevo León, June, 2021

Declaration of Authorship

I, Ramón Eduardo Díaz Ramos, declare that this thesis titled, ”One Step Closer to Mental Health: Resilience to Mental Stress Index in the Face of Analytical Problems, A Machine Learning Approach” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this dissertation is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Ramón Eduardo Díaz Ramos
Monterrey, Nuevo León, May, 2021

Dedication

This thesis is dedicated to my wife, Daniela Alejandra. Thanks to your support and motivation, this work was made possible. I am thankful for having you in my life. You showed me true patience and love during these difficult times.

Acknowledgements

I would like to express my profound acknowledgment to God and all the blessings that He has brought to my life. Also, I want to express my gratitude and most grateful acknowledgment to my parents, Francisco Ramón and Cecilia, who have shown me unconditional loved and support.

I want to thank my sister, Ana Cecilia. You taught me to be positive even in the most difficult situations. Your faith and joy are contagious and worthy of admiration.

I would like to recognize my thesis supervisor, Dr. Luis A. Trejo Rodríguez. His constant support and patience help me greatly to develop this work. Also, I would like to thank Dr. Carlos Figueroa and Dr. Miguel Medina for committing their time and efforts to this work.

I would also like to thank my thesis committee members, Dr. Miguel Medina, Dr. Miguel González, Dr. Fresia Hernández, and Dr. Carlos Figueroa for sharing their knowledge.

I would like to acknowledge the support of the research group, José Javier Rodriguez Mota, who developed the mobile application of this study, and Danna González and Isabel Rojas for helping in the collection of data. Also, to the Department of Psychology of Tecnológico de Monterrey for developing the psychophysiological mental stress test.

I would like to express my thanks to the Computer Science Department of Tecnológico de Monterrey, especially Dr. Hugo Terashima and Dr. Francisco J. Cantú Ortíz.

Finally, I would like to acknowledge the funding I received from Tecnológico de Monterrey and CONACyT.

One Step Closer to Mental Health: Resilience to Mental Stress Index in the Face of Analytical Problems, A Machine Learning Approach

by

Ramón Eduardo Díaz Ramos

Abstract

Stress and depression are two major topics of concern for higher education institutions. Studies have shown how mental health problems can decrease students' ability to function efficiently during their education life and how depression can risk their physical well-being. To aid students in coping with the challenging experience of higher education and therefore enable them to perform better in stressful situations post-graduation, researchers recommend increasing their level of resilience.

In an attempt to measure a person's resilience, previous studies have developed and analyzed self-rating questionnaires. While these studies have provided a way to assess people's psychological responses and provided a significant amount of insight, they do not provide an objective measurement for resilience to mental stress. There have been related studies that have evaluated physiological signals in individuals and have identified relationships with people's stress. Based on previous literature and applying machine learning, this thesis aims to demonstrate the feasibility of measuring an individual's resilience to mental stress and proposes a Resilience to Mental Stress Index (RMSI). In addition to this, this thesis presents an experiment's methodology to collect physiological and psychological data using smartwatch embedded sensors and psychological tools to study depression prediction.

This research performed data analysis of 71 individuals subjected to a 10-minute psychophysiological test to study resilience to mental stress. The data collected considers five physiological features: (a) muscle response (electromyography), (b) blood volume pulse, (c) breathing rate, (d) peripheral temperature, and (e)skin conductance. We utilized unsupervised learning techniques to visualize and identify the relationship between these feature variability. As a result of the analysis, we created three different methods for the RMSI. The results' analysis between the different methods showed no statistically significant difference ($p>0.05$). However, we recommend using the Mahalanobis distance (MD) method because of its relationship with the validation methods. Even though there exists no standard method to quantify resilience to mental stress, our results indicate a positive relationship to the Resilience in Mexicans (RESI-M) psychological tool.

For the study of depression prediction, during this research, five variables were selected for the study: (a) personality traits, (b) RMSI, (c) heart rate variability (HRV), and (d) sleeping disorders. To collect these variables, we developed a methodological framework and built a mobile application. We hope that this research serves as a solid baseline to understand resilience to mental stress and collect valuable information to predict depression.

Keywords: Machine Learning, Stress, Depression, Resilience to Mental Stress Index, Collection of Depression Data, Unsupervised Learning.

List of Figures

List of Tables

Contents

Chapter 1

Introduction

1.1 Background

The World Health Organization ranked depression as the single most significant contributor to global disability and anxiety as sixth. Globally, over 320 million people are estimated to suffer from depression. Furthermore, 56% of the people with depression do not receive any treatment [1]. Only 16% received treatment in Mexico in the first year for any mood disorder [2].

Worldwide, people suffer from depression, governments and administrators struggle to improve depression care [3]. Women are twice as likely to experience depression than men. Moreover, 15 to 20% of adults will experience major depression at some point [4]. Automated devices' need to detect depression becomes evident with the number of people suffering from it and the low rate of people receiving treatment. Brachet-Marquez [5] stated that 4 out of 10 most disabling diseases in Mexico are neuropsychiatric, such as schizophrenia, depression, obsessive-compulsive disorder, and alcoholism.

Another study was conducted in Mexico in 2002 to detect major depression disorders in a population aged from 18 to 65. The diagnostic instrument was the Composite International Diagnostic Interview (CIDI). Furthermore, 2% of the Mexican population experience a major depression in their life, being 9.8% women and 4.4% men, with an average of 7 episodes during their lifetime. The first episode lasted 31 months and generally did not receive any treatment [6].

A model of the impact of depression on health behavior and longevity was developed by Holger [7] in Germany. The model explains that people with depression reduced their life satisfaction because of symptoms of depression. People tend to consume more unhealthy products, save less money, spend less on their health, and exercise less. As a consequence of the decrease in life satisfaction, the individual has fewer incentives to be involved in healthy activities to increase their life quality. Holger concluded that depression impacts a 20% decline in life satisfaction, reducing the length of life by about four years.

Holger [7] stated that delays in depression diagnosis lead to losses in health that are not fully recovered in the treatment period. Most individuals do not manage to completely compensate for the unhealthy behavior that depression causes in early adulthood.

Several studies [8, 9] have shown how mental health problems can decrease students' ability to function efficiently during their education life and endanger their physical well-being. The study from Western Michigan University [9] investigated the relationship between depression and the academic performance of undergraduate students. They concluded that depression was associated with a 0.49 point decline in student GPA. With appropriate treatment in the early stages of depression, approximately 0.44 point grade was protected at the end of the semester.

Depression negatively impacts the way a person feels, thinks, and acts. It causes feelings of sadness or a lack of interest in activities once commonly enjoyed. It can appear at any time, but the first time it is generally experienced is in the late teens to mid-20s. More symptoms can include the following [10, 11]:

- reduced concentration and attention,
- reduced self-esteem and self-confidence,
- ideas of guilt and unworthiness,
- pessimistic views of the future,
- thoughts of suicide or self-harm,
- disturbed sleep,
- reduced appetite.

It was estimated that almost 90% of people with sleeping disorders suffer from depression [12]. Additionally, depression often comes with symptoms of anxiety [13]. These problems can appear anytime and could be chronic, leading to impairments in an individual's daily activities. When students show symptoms of depression, such as difficulty thinking, concentrating, or making decisions, they will not perform as usual in their academic activities.

There are various methods of classification of depression levels. The ICD-10 Classification of Mental and Behavioral Disorders was selected for this study because of its three characterization levels. There can be a single or first depressive episode in which mild, moderate, and severe categories are used. Chronic states of depression are equally classified as a single depressive episode. The symptoms' minimum duration must be two weeks until it is considered a recurrent disorder [11].

For a person to be categorized in mild depression, it should have at least two of the most common symptoms of change in mood, loss of interest and enjoyment, increased fatigue, and an additional two of the symptoms described before. A person with mild depression has some difficulty in everyday work and social activities but will possibly not cease to behave ordinarily. In moderate depressive diagnosis, at least two of the three most typical symptoms are noted for mild depressive, and at least three of the other symptoms should be present. A person with moderate depression will frequently have trouble continuing with social, work, or domestic activities. In severe depression, the patient shows distress or agitation unless retardation is present. All three manifestations for mild and moderate depression are present, plus at least four common symptoms. It is doubtful that the person will maintain regular interaction with social, work, or domestic activities.

According to Statista[14], the use of smartwatches has increased from 18.7 million units to 62.6 and will continue to grow with a forecast for 2022 of 113 million units. The widespread use of smartwatches facilitates capturing physiological data and increases the opportunities to study individuals' data. In this study, the use of smartwatches aids the collection of some of the features of concern.

Different causes can trigger major depression disorder. Genetic may determine a person's susceptibility to depression. The levels of the brain's neurochemicals, such as the decrease of dopamine, norepinephrine, and serotonin, have a relation to clinical depression. Also, psychosocial stressors that a particular individual has in her/his environment can trigger depression, such as financial or emotional stress, traumatic events, or unemployment [4].

The features of interest that could impact detecting depression are heart rate variability, personality traits, physical activity, sleeping disorders, and resilience to stress. The latter variable is the feature of the focus of this thesis.

Stress has become a primary topic in the pursuit of mental health for modern society. The main reason for this is that stressful events can lead to major psychiatric conditions, such as anxiety and depression [15]. Consequently, these mental illnesses can lead to physical problems, such as an increased risk for cardiovascular diseases [15, 16]. Given the mental and physical problems that stress can cause, there has been an increase in the interest in studying ways to detect and prevent stress in individuals promptly [17, 18, 19, 20].

Researchers [21, 22] suggest increasing students' resilience to aid them to succeed with the demanding experience of higher education and, as a result, perform better in stressful circumstances after graduation.

Resilience to stress is defined as the properties contributing to the speed and amount of possible recovery after exposure to a stressful event [23]. Since it is challenging to avoid stressful situations in a fast-paced environment, it has become increasingly important for individuals to develop a mindset to overcome stress [23, 24, 25]. We strongly believe that a metric for resilience to mental stress can provide a measure to identify potential hazards and quantify the improvement that an individual makes when getting a step closer to mental health.

Many psychophysiological tools measure resilience based on questionnaires. These studies assign a scale depending on the individual's response to the questions. The Connor-Davidson Resilience Scale (CD-RISC) [26], the Resilience Scale for Adults (RSA) [27], and the Resilience in Mexicans (RESI-M) [28] are examples of these studies. These tools assign a scale based on how the subjects have felt in the previous weeks. The variety of these scales gives psychologists and psychiatrists a wide range of tools to measure resilience, considering features that may reflect resilience's underpinnings. Some of the features are the following: adaptability when coping with stress, strong self-esteem/confidence, adaptability when coping with change, social problem-solving skills, humor in the face of stress, strengthening the effect of stress on others. These tools are focused on measuring the full spectrum of resilience, but they do not provide resilience to mental stress, specifically.

Physiological changes, such as the variance in the electromyography, electrocardiogram, skin conductivity, respiration changes, and eye closing rate, have long been used as reliable indicators to recognize emotion [29, 30]. Different classification problems have been resolved and analyzed to detect emotion by exploiting these physiological responses from an individual [29, 31].

1.2 Problem Statement

As previously stated, depression and stress are part of the top mental health issues in modern society.

Many suicides and disabilities are led by depression alone. Designing and collecting data intended to predict depression is of great importance to mental health. Also, as stress could be directly correlated with depression, it becomes paramount to analyze it. However, in this fast-paced environment is challenging to avoid stress. Therefore, enhancing stress resilience can decrease the likelihood of developing stress-induced depression or anxiety [32, 33].

"You can't manage what you don't measure." This management quote exemplifies the gap of research to improve resilience to mental stress. There is a need to quantify this psychophysiological response in the face of an analytical challenge; having a point of reference can help understand and measure improvements in psychological treatments for individuals.

Although multiple works have examined stress detection, to the best of our knowledge, we have identified that the features explored in such studies have not measured direct resilience to an individual's mental stress. We have identified no quantifiable index to monitor the degree of an individual resilience to mental stress based on physiological responses to an analytical challenge. This work proposes an approach to measure resilience to mental stress based on biofeedback sensors in students.

1.3 Hypothesis

The hypothesis is that resilience to mental stress can be quantified by measuring an individual's body responses to mental stress in the face of analytical challenges using unsupervised machine learning techniques from physiological features—specifically, utilizing inter-cluster and intra-cluster metrics.

1.4 Research Objectives and Questions

This study has two main objectives:

1. Design an experiment to collect data that includes the identified features that could relate to depression (heart rate variability, sleep disorders, physical activity, personality traits, and resilience to mental stress).

2. Create a new index that quantifies resilience to mental stress when inducing a stressor based on analytical challenges, and create a new factor that quantifies the alteration suffered by the physiological features.

In particular, this work has the following specific objectives:

1. Create a novel dataset by designing and implementing the technological infrastructure based on smartphones and specialized devices. To collect the identified features that could help predict depression.

2. Select the best methodology to compute the resilience to mental stress index and alteration factor. Validate the correctness of the two new metrics: resilience to mental stress index and alteration factor.

The research questions are the following:

1. Which methodology is the most appropriate to build a resilience to mental stress index and an alteration factor?
2. Can the (dynamic) time-series data be transformed into static data to include in a classification problem without vectorizing or training ensemble models?
3. Can the resilience to mental stress index and the alteration factor be compared among individuals?
4. Can the resilience to mental stress index and the alteration factor be independent of socio-demographic features?
5. Does the resilience to mental stress index is correlated with the resilience in Mexicans (RESI-M) psychological tool?

1.5 Significance of Study

Designing an experiment to collect data to detect depression can help by adding a collection of scientific literature of the features that could increase the likelihood of detecting depression promptly. Additionally, collecting sample data increases the knowledge discovery of mental health by providing sample data that could be analyzed.

Muscle response, blood volume pulse, peripheral temperature, breathing rate, and skin conductance will be evaluated with unsupervised machine learning algorithms to quantify the resilience to an individual's mental stress. We hope that this research is of interest to academic institutions and health organizations. Moreover, this study contributes to the field of machine learning research applied to psychiatric computation. It adds to existing research of similar physiological responses to explain psychological behaviors by exploring Mexico's students' related features.

This study adds to the application of unsupervised machine learning by defining an appropriate methodology to quantify physiological features' alteration. The research focus on cluster distances to quantify the alteration of the five measured physiological features.

Additionally, it provides a different approach to combine static and dynamic features by transforming the dynamic features, in this case, a time series, without vectorizing sequential data or using an ensemble method to use conventional classification models. Vectorizing features usually increases the codependency between attributes, and it can create a large p (number of features), small n (number of instances), $n << p$ issue, with insufficient degrees of freedom to estimate the model [34]. Meanwhile, the ensemble methods require high computational expenses. It introduces ambiguity by requiring model selection and experimentation of additional features for model tuning [35].

1.6 Outline

This section describes the organization of this thesis with the following chapters:

Chapter 1: Introduction. This chapter describes the problem of depression and the lack of a measure for resilience to mental stress. It describes the objectives and significance of this research.

Chapter 2: Research Framework. The second chapter presents the research background, as well as the related work. It describes the current state-of-the-art techniques used in this thesis. Moreover, it also briefly describes each related feature of interest in this study.

Chapter 3: Methodology. This section presents the proposed methodology to collect data for the classification problem of depression. Next, it presents the methodology for signal processing and description of data. Lastly, it describes three different proposals to calculate distances that capture the behavior of the features.

Chapter 4: Results. The Results chapter describes the proposed resilience to mental stress index equation, the alteration factor and analyzes the resulting resilience to stress index with two comparative cases.

Chapter 5: Discussion and Conclusion. The final chapter presents the thesis's discussion, conclusion, limitations, and future works.

Chapter 2

Research Framework

2.1 Introductory Comments

The beginning of the chapter briefly describes the data science methodology, unsupervised machine learning techniques, and the related methods employed in this thesis to answer the research questions. Then, we reviewed the current tools to quantify depression and resilience and the personality traits tests. Lastly, the final section describes the physiological variables background identified to help predict depression, the physiological features that contribute to creating resilience to the mental stress index, and the related work. Throughout the document we use the words variables and features to refer to the same concept (physiological data).

2.2 Data Science Methodology

Data Science is the science of acquiring a specific type of knowledge through data. It is all about acquiring, analyzing the data, and using the acquired knowledge to make predictions and understand history. As it can be observed in Figure 2.1, Data Science is a set that must contain three essential areas of knowledge, computer science skills can program algorithms to create predictive models, math and statistics allow to evaluate the algorithms and the data, and a dominium in the area allow to apply the concepts and results in a meaningful way [36].

The data extracted must be analyzed, and it is necessary to understand the type of data to choose the correct method to analyze it. There are three primary classifications of data: structured and unstructured, quantitative and qualitative, and four levels of data. Structured data is usually organized; it can be ordered as a table in rows and columns. Specific observations carefully recorded so that they can be easily digested and visually analyzed. Most machine learning models were developed with structured data and statistical analysis.

The unstructured data is unorganized; it does not follow any standard position or grouping. Text data are generally unstructured data; tweets, e-mails, and literature are generally unstructured data. If the data is not ordered, a preprocessing procedure should come before the study is carried out, attempting to organize an unstructured data collection. Examples of how unorganized data can be converted to structured data create categories, including word counts, the numbering of unique characters, the text's length, or the main topic.

Quantitative data can be represented using numbers and basic mathematical methods.

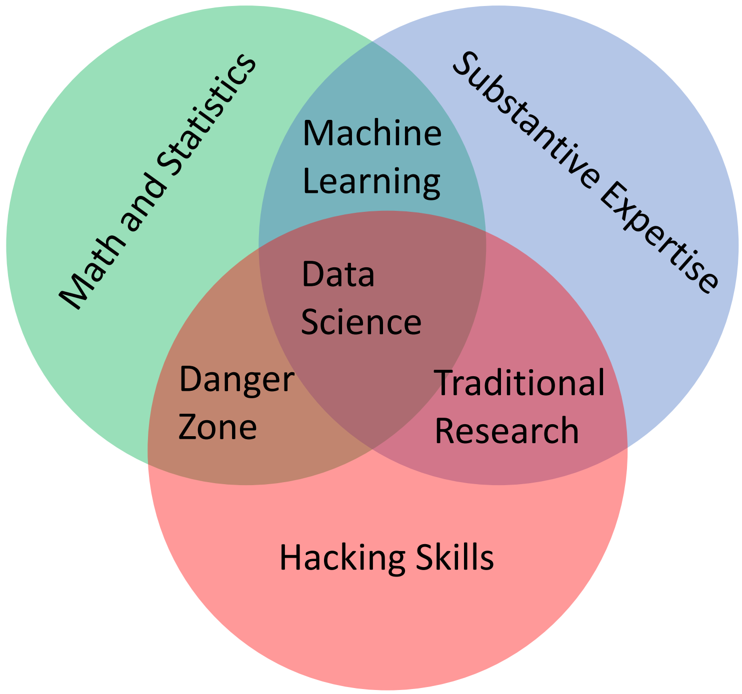


Figure 2.1: Essential principles of Data Science Venn Diagram, where three essential areas of knowledge are needed.

It is usually a structured data set that can include a table with the row and column format. Quantitative information can be categorized into discrete and continuous data. Numbers do not always represent quantitative data; if the result of applying an essential mathematical operation is inconsistent, then it should be classified as qualitative data, and a set of zip codes is an example.

Discrete quantitative data are information that can be counted, which has an actual number to count. Continuous data, on the other hand, describes information measured on an infinite range of values. The height of a building is an example as an infinite scale of decimals can describe the height. Qualitative data is generally described in categories and language. Whenever a word is used to describe a characteristic, it is classified as qualitative data.

The four levels of data are nominal, ordinal, interval, and ratio levels. The nominal rate consists of knowledge defined by name or classification only, classified as qualitative data, and cannot be represented as numbers. Nominal data can be analyzed as equality and set membership. Ordinal level data provide a rank order, but it does not provide a relative information comparison. Therefore, the information can be ordered, but it will not give meaning compared to the observations.

An example is the ranking of the questionnaires. Interval data can be expressed with quantifiable information. Therefore, it allows addition and subtraction between data points. The ratio level is data that allows mathematical operations, addition or subtraction, and multiplication and division.

The book of Ozdemir, Principles of Data Science, states five essential steps to perform data science: 1 Asking an exciting question, 2 Obtaining the data, 3 Exploring the data, 4

Modeling the data, and 5 Communicating and visualizing the results. Asking an exciting question relates to brainstorming ideas. The questioning of the issue and the statement to be explored is a method for setting the analysis's objectives. Then the data based on solving the goals and the data categorizing the data form to apply the data modeling. At the end of the process, the findings are compiled and visualized to digest the data to infer the goals [36].

2.3 Unsupervised Machine Learning

Machine learning aims to design computer algorithms and techniques to allow the computer to display a behavior learned from empirical information. As shown in the Venn diagram in figure 2.1, machine learning consists of scientific computing, maths, and statistics. Domain knowledge is not taken into account. Machine learning algorithms can be classified based on how the machine is trained: supervised and unsupervised learning [37]. This section explores unsupervised learning in specific visualization and dimensionality reduction methods.

The data in unsupervised learning is a collection of unlabeled instances. Unsupervised learning algorithms aim to construct a model that takes a function as input and transforms it into a value that can be used to solve a problem. The model returns the cluster id for each function in the dataset while clustering. The output of a dimensionality reduction model is a function with fewer features than the data. Outlier identification generates a Real number that shows whether the sample varies from a standard case in the dataset.

Data visualization is the most common use of dimensionality reduction. However, it also eliminates functions that are redundant or strongly connected. It can also help to reduce the amount of noise in the data. Principal component analysis (PCA), uniform manifold approximation and projection (UMAP), and autoencoders are the three primary techniques for dimensionality reduction. This study will focus on PCA due to its fast evaluation characteristics and the transformation that removes multicollinearity [38].

2.4 Principal Component Analysis

Principal component analysis (PCA) is an unsupervised learning technique used for dimensional reduction. It identifies the orthogonal dimensions in a dataset to identify the principal components and explain the relationships between variables [39]. The orthogonal transformation [40] identifies the linearly correlated variables and transforms them into new uncorrelated variables that successively magnifies the variance. Unlike the other feature selection methods described in this section, the PCA method does not eliminate any variable but takes the valuable information of all the variables and incorporates that into new variables. The method obtains the eigenvalues and eigenvectors of the covariance matrix and orders the eigenvalues from those that give the maximum information to those that provide the minimum information. The first two transformed features, represented by the red and orange lines in Figure 2.2, are the ones that provide the most information and so are considered the principal components.

An advantage of this is method is that it can reduce correlated features and improve visualization by transforming high-dimensional data into low-dimensional data. The disadvantage, however, is a loss of explainability as the new variables cannot be interpreted as the original variables [41].

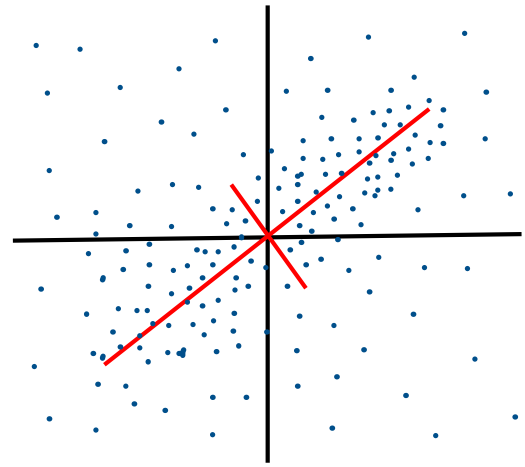


Figure 2.2: Two transformed features, represented by the red and orange lines of the original space.

Specific conditions of the dimensional space must be met to obtain confident results of the PCA [40]. The assumptions of PCA are the following:

- Multiple variables are continuous.
- Linearity. It assumes that the features have linear combinations. This is because PCA is based on Pearson correlation coefficients. Hence, there must be a certain degree of linear relationship between the variables.
- Large enough samples are required to produce reliable components. Usually, a minimum of 150 instances per variable is required[a].
- The results depend on the scaling of the data. To acquire reliable components in the n-dimensional space, they must have the same scale.

After applying PCA, the first component often describes a more significant proportion of the overall variation in the results than the second component. After obtaining the principal components, it is possible to calculate the Euclidean distance or the Mahalanobis distance in this principal component space instead of the original space [42]. These distances are discussed in the following sections. Performing Euclidean distance in the principal components eliminates multicollinearity [43].

2.5 Cluster Validity Indices

Cluster validation is a challenging task that lacks the scientific basis that other disciplines have, such as supervised learning [44]. There is no absolute “best” criterion independent of the clustering’s target [45]. Fortunately, the clusters are already assigned for this work, and

[a]https://statistics.laerd.com/spss-tutorials/principal-components-analysis-pca-using-spss-statistics.php

each label corresponds to each phase of the psychophysiological stress test. The objective of the psychophysiological stress test is to induce mental stress with analytical problems, and the cluster validity indices help us compare cohesion and separation of each subject relative to itself and other subjects in standardized data.

Previous research has shown that no single cluster validity index outperforms the others [44]. Therefore, we can consider the cluster validity index most appropriate to our problem as the cluster distances are of great interest to this thesis.

The majority of indices calculates the cluster cohesion (intra-cluster metric) and the cluster separation (inter-cluster metric), and the combination of both is usually calculated by a division or a sum [46]. For this thesis, the Silhouette index was used for its adaptability of choosing its distance metric.

2.5.1 Silhouette Index

Peter Rousseeuw proposed the Silhouette index [47]. Each cluster is defined by a silhouette created by comparing the cohesion and separation of the clusters. This silhouette portrays which objects are well-placed within their clusters and which are merely in between clusters.

This index is a summation index that has been normalized. The distance between all the same cluster points is the cluster's cohesion, and the separation is calculated based on the separation to the nearest neighbor [46]. The Silhouette Coefficient can be calculated using python library Scikit-Learn[b].

The Silhouette Coefficient is computed with the mean intra-cluster distance (a) and the mean nearest-cluster distance (b) for each instance that the sample is not a part of its assigned cluster. Each instance is calculated in Equation 2.1.

$$SC = \frac{(b-a)}{max(a,b)} \tag{2.1}$$

A Silhouette value of zero indicates that the clusters are overlapping. Negative values usually mean that a sample was wrongly assigned and could belong to a different cluster. A value close to one indicates that the clusters are well-matched [47].

2.6 Distance Metrics

Cluster analysis tries to find valuable groupings that could provide insights or characteristics of each forming group. Finding helpful groups means that they can be identified when comparing them (homogeneity) [48]. This means that the clusters have cohesion and separation from each other. These are metrics that are measured by distance. A metric is a formula that calculates the distance between any two points or groups in the metric space R [49]. A set of points within the function or distance d constitutes the metric space.

Several distance metrics have been proposed in the literature for calculating the similarity between feature vectors, including the Manhattan distance metric, the Euclidean distance metric, and the Vector Cosine Angle Distance. Manhattan and Euclidean distance metrics are commonly used to evaluate the correlations between the data objects in a cluster [45].

[b]https://scikit-learn.org/stable/modules/generated/sklearn.metrics.silhouette_score.html

Another metric distance is the Mahalanobis distance, which is frequently used in cluster analysis problems. The Mahalanobis distance differs from Euclidean because the dispersion of the features is taken into account [49].

The Mahalanobis and Euclidean distance can be calculated in the original dimensional space. The Euclidean distance is easy to compute and interpret; on the other hand, the Mahalanobis distance is more complicated [50].

2.6.1 Euclidean

Euclidean distance remains the traditional form for calculating the distance between two points in space when the conditions are satisfied. It is used to determine the similarity between a cluster's features, and it is a metric in the space R [49].

The Euclidean distance of two n-dimensional vectors is presented in the following Equation [51].

$$d(x, y) = \sqrt{\sum_{i=1}^{n}(x_{i1} - y_{i1})^2 + (x_{i2} - y_{i2})^2 + ... + (x_{in} - y_{in})^2} \tag{2.2}$$

Some form of standardization is required in n-dimensional space to balance the contributions of each dimension of the Euclidean space. A normalized distance is often used to compare functions measured in different units with substantially different values [49].

2.6.2 Mahalanobis

Besides the Euclidean distance, the Mahalanobis distance is another commonly used technique to measure distances that could be calculated in the original feature space of a principal component space. The variance-covariance matrix C_x is shown in Equation 2.3 [50].

$$C_x = 1(n-1)(X_c)^T(X_c) \tag{2.3}$$

Where X is the data matrix containing n instances and X_c is the column-centered data matrix. Once the variance-covariance matrix is calculated, we can calculate the Mahalanobis distance in Equation 2.4.

$$MD(x, y) = \sqrt{(x-y)^T C_x^{-1}(x-y)} \tag{2.4}$$

Since the Mahalanobis distance is determined using the inverse of the variance-covariance matrix of the data collection of interest, the Mahalanobis distance considers data similarity in the initial variable space. The computation of the variance-covariance matrix, on the other hand, maybe troublesome. When there is a large number of features, they could contain correlated information. This multicollinearity could lead to a singular variance-covariance matrix that cannot be inverted. Another limitation is that the variance-covariance matrix calculates that the amount of instances in the data needs to be larger than the number of features [50]. Nevertheless, these limitations could be circumvented by applying techniques of dimensional reduction, as PCA.

2.7 Depression Tests

Different tests have been developed and tested in different populations to aid the diagnosis and detection of depression. These assessments are usually conducted with an interview or by inventories using a self-report, clinician-report, or family report. There are several ways to detect depression by written text; the following are the most commonly used [52].

2.7.1 Hamilton Rating Scale

Hamilton Rating Scale (HRS) [53] is one of the most used tests to recognize depression. Max Hamilton initially published it, and its purpose is to measure the severity of depression in subjects previously diagnosed with the disease. It consists of 17 items graded in the Hamilton Scale (e.g., guilt, suicide, insomnia, agitation, anxiety). Four items were added, but Hamilton stated that the last four items: diurnal variation, depersonalization, paranoid symptoms, and obsessive and compulsive symptoms, should not contribute to the score because they were not considered part of the symptoms. The Hamilton Depression Scale requires trained staff to be administered [54].

2.7.2 Beck Depression Inventory

Beck Depression Inventory (BDI-II) [55] is likewise one of the most used self-rating scales for measuring depression. The BDI discriminates between subtypes of depression and differentiates between depression and anxiety [56]. There have been two major revisions in 1978, and in 1996, it has 21 items (e.g., agitation, worthlessness, difficulty concentrating, and energy loss) and use in more than 7,000 studies. The test's theoretical assumption is based upon the idea that negativistic distorted cognitions would be the core characteristic of depression [57].

2.7.3 Geriatric Depression Scale

Geriatric Depression Scale (GDS) [58] was developed by Jerome Yesavage and his colleagues and consists of a 30 question-based test. A short version of the test consists of 15 questions with yes or no answers. Suppose the patient scores more than four points against the score sheet of the test. In that case, minor depression is identified, and if more than nine answer matches, major depression is diagnosed and should have a complete psychiatric evaluation [59].

2.7.4 Zung Self-Depression Scale

Zung Self-Depression Scale [60] consists of 20 items with a 4-point scale constructed with clinical diagnostic criteria to diagnose depression. Of the 20 items, ten are worded positively and ten negatively concerning symptoms. The total score is used as the index of severity of depression. A higher score indicates greater severity of depression [61].

2.8 Personality tests

A personality test is an approach to categorize the traits of human personality to explain human behavior. There has been much research and applications using the scales to measure distinctive characteristics where people can be ordered. These characteristics are consistent patterns that define how individuals understand, behave, and react in the world. A personality rating scale describes the characteristics of the same personality with a score and scales that better describe the behavior of the individual [62].

2.8.1 Big Five

Cattell [63] developed a classification of individual differences that consisted of 16 primary elements and eight second-order elements. Tupes [64] investigated the classification of Catell and found that five factors have a reasonable correlation: Surgency, Emotional Stability, Agreeableness, Dependability, and Culture. A study conducted by Norman [65] confirms the classification of Tupes and Christal and his labels: Extraversion, Emotional Stability, Agreeableness, Conscientiousness, and Openness to Experience are most commonly used in the literature and have been called the "Big Five" [66]. Each "Big Five" attribute consists of another attribute within its domain.

People who fall into the category of *extraversion* distinguishes by being talkative, energetic, and bold; on the opposite side are those described as quiet, shy, and withdrawn. They are more likely to feel comfortable around other people. They do not mind being the center of attention and usually start a conversation. *Agreeableness* people distinguish as cooperative, sympathetic, and kind; on the opposite side of the trait are described as cold, rude, and unsympathetic. Individuals that got a high score in *agreeableness* tend to respect others, treat them as equals, and are concerned about them. The *conscientiousness* classification characteristics are responsible, efficient, organized, and thorough, and people who score lower in this classification are described as disorganized, careless, sloppy, and inefficient.

Conscientious people tend to be prepared, pay attention to details, and make and follow schedules. The four main features are industriousness, reliability, orderliness, and impulse control. The *openness to experience* feature characteristics is described by being imaginative, philosophical, creative, and profound. People with these characteristics tend to enjoy thinking about things, hearing new ideas, having larger vocabularies, and value artistic expression. On the other hand, people who score lower are described as uninquisitive, unimaginative, unsophisticated, and shallow. All five features also correlate. They are not mutually exclusive. Emotional stability correlates with extraversion, agreeableness, and conscientiousness; extraversion correlates with openness to experience and conscientiousness with agreeableness [67].

2.8.2 Minnesota Multiphasic Personality Inventory

The Minnesota Multiphasic Personality Inventory (MMPI), developed by Starke R. Hathaway and J. C. McKinley. The scales can be mainly divided into two parts: the validity part, which has three scales, and the clinical part, with scales such as hypochondriasis, depression, hysteria, psychopathic deviate, masculinity/femininity, paranoia, psychasthenia, schizophrenia, hypomania, and social introversion [68].

A second revision of the test was made in 1989 with a publication of the MMPI-2. In this new revision, a new representation of the sample was re-standardized to have a more significant group representation of the United States population. Additional validity measures and newly developed scales were incorporated in the second revision [69]. The MMPI-2 is a standardized questionnaire consisting of 567 affirmations with true or false answer statements. Interpretation of the results depends on the configuration of scales on the person's profile sheet and demographic characteristics. The profile sheet is made up of nine validity measures and ten traditional clinical scales. The MMPI-2 is valid only for adults, but a shorter version of the test (MMPI-A) is for adolescents [70].

Another revised publication of the MMPI-2 was made for a restructured form, which is 338 statements, and the assigned scores are based on a set of clinical scales. This revision was made to reflect the current knowledge of the phycological problems reported by the test [71].

2.8.3 California Psychological Inventory

The California Psychological Inventory (CPI) test is designed to analyze traits that describe an adult's personality. It was designed by Harrison [72]. Four hundred sixty-two statements are evaluated with true or false answers that consist of personal feelings and opinions, ethical and social attitudes, characteristic behavior, and personal relationships.

The characteristics that these test measures are sociability, dominance, independence, responsibility, self-control, tolerance, and achievement. The test can identify and classify 30 scales (Dominance, Capacity for Status, Sociability, Social Presence, Psychological-mindedness) [73].

2.8.4 Personality Assessment Inventory

Lastly, we will review the personality assessment inventory[c] (PAI). The PAI measures psychopathological syndromes and offers evidence for psychiatric diagnosis, recovery preparation, and psychopathology screening. The test is available in Spanish and has been validated [74]. Significant elevations over various scales characterize the PAI clinical profile, suggesting a broad spectrum of clinical characteristics and raising the risk of multiple diagnoses [75]. It has 22 non-overlapping scales, which promotes high discriminant validity.

The test consists of 344 items that have an average completion time of an hour. The 22 scales provide four validity scales, five treatment scales, two interpersonal, and 11 clinical scales [75]. From there, it assigns a score based on the multiple-choice answer. For this study, the PAI tool is selected to characterize the individual's personality and possible mental disorders because of its validity in Hispanic regions.

2.9 Resilience Tests

Even though there is no direct measure of resilience to mental stress, this work identifies several tools that quantify individuals' resilience as a broad spectrum. These tools serve

[c]https://www.parinc.com/Products/Pkey/287

as a validation point of our work. We will compare the results of the resilience test right after the experiments to collect the physiological variables. Four types of tools stand out in the literature for their validation in different communities, their development in different languages, and the number of citations. The four tests are briefly reviewed in this section. These tools are The Connor-Davidson Resilience Scale (CD-RISC) [26], the Resilience Scale for Adults (RSA) [76], the Resilience Scale (RS) [77], and the Resilience in Mexicans scale [78].

2.9.1 Connor-Davidson Resilience Scale (CD-RISC)

The CD-RISC is a rating scale to assess resilience. It consists of 25 items, each rated on a scale from zero to four, where a higher score reflects a higher trait of resilience. The scale had good psychometric properties, and factor analysis revealed that there were five factors. These five factors are personal skill, high standards, and tenacity reflected in Factor 1. Factor 2 is associated with trusting one's instincts, coping with negative emotions, and enhancing the effects of stress. Factor 3 is about positive change approval and stable relationships. Factors 4 and 5 were linked to control and spiritual factors, respectively. The CD-RISC has good psychometric properties and can differentiate between more resilient people and those who are less resilient. [26].

It is sorely tested in both the general population and clinical samples. Internal consistency and test-retest reliability are both good. However, it lacks a comprehensive process description to assess it and a detailed scoring method [79].

2.9.2 Resilience Scale for Adults (RSA)

The RSA was developed by Friborg et al. [76] for assessing the presence of protective resources that constitutes resilience in adults. The scale consists of 37 items covering five factors: social support, personal competence, family coherence, personal structure, and social competence. In health and clinical psychology, the RSA is a legitimate and trustworthy tool for determining protective factors critical for regaining and maintaining mental health. There is concern about the scale's validity due to the non-random sample and low response rate [79].

Because the RSA subscales have low-to-moderate intercorrelations, the dimensions should be viewed as subscales measuring distinct but all positive aspects of the notion of resilience. This low-to-moderate intercorrelation supports the notion of resilience as multidimensional [76]. Hence, we can consider resilience as a whole set and resilience to mental stress as a subset of the resilience set.

2.9.3 Resilience Scale (RS)

The RS consists of a 25-item questionnaire that rates using a 7-point score. Scores range from 25 to 175 points. The higher the scores, the more the trait of resilience [80]. The test has two factors that measure the construct of resilience: acceptance of self and life and personal competence. This tool was initially tested with adults, but several studies have validated the scale with multiple samples of several ages and ethnic groups [79]. Perseverance, or the act of persisting despite adversity or discouragement, is one of the five traits of resilience that serve as the Resilience Scale's conceptual foundation. It denotes a desire to continue the struggle to rebuild one's life and stay involved in the face of adversity [80].

2.9.4 Resilience in Mexicans (RESI-M)

Palomar and Gómez [78] developed a resilience scale tested in the Mexican population. It considers two resilience scales in adults: The CD-RISC of Connor - Davidson and the RSA of Friborg et al. as a reference in its construction. The test has 43-items that assign a score using a 4-point rate scale. A higher score indicates a higher trait of resilience.

The scale measures five factors related to resilience: strength and self-confidence, family support, structure (order and organization in their life), social support, and social competence. This instrument was primarily validated in Mexican adults. Therefore, the scale is in Spanish. The integration of both scales: CD-RISC and RSA, allowed selecting the best items from both scales to assess resilience in the Mexican adult population. According to the authors, the scale has adequate validity and reliability indexes. Due to its validation in Mexican culture and its Spanish availability, this tool was considered for execution in this work.

2.10 Identified Variables for Depression

A search was conducted to identify features related to depression. Previous studies have analyzed several variables that could trigger a depression event. This study used these variables that could precisely be monitored, measured, and related to depression in students. The identified variables are described in this section.

2.10.1 Heart Rate Variability

The study of Agelink et al. [81] compared time and heart rate variability (HRV) between 32 patients with major depression assessed with the Hamilton Depression Scale and 64 non-depressed controls. The patients of the experiment were controlled by age, gender, and smoking. The study concluded that patients with severe depression on the Hamilton Scale showed a higher heart rate and a significantly lower cardiovagal activity modulation than the control group. Patients with mild depression did not differ much from control, but they were in the expected direction. Agelink et al. concluded a significantly negative correlation between the Hamilton Scale scores and the variability in heart rate, suggesting a direct link between the severity of depressive symptoms and cardiovagal activity modulation.

Another study [82] did a meta-analysis to measure the impact of depression on HRV in depressed patients without cardiovascular disease. The meta-analyses were based on 18 articles with 673 depressed participants and 407 healthy compared controlled groups. The results of the study showed that patients with depression had lower HRV than the healthy control group. They concluded that depression without cardiovascular disease is related to reduced HRV, decreasing depression severity. Also, tricyclic medication decreased HRV, and serotonin reuptake inhibitors, mirtazapine, and nefazodone had no significant impact on HRV.

Koening et al. [83] meta-analysis compared the HRV of 99 children and adolescents clinically diagnosed with depression with a controlled sample of 160. The meta-analysis results showed a lower resting state of HRV among clinically depressed adolescents/children than the healthy controlled group. They did not find a consistent correlation between the severity of depression and HRV.

These previous studies showed a correlation between the HRV in patients clinically diagnosed with depression and with and without cardiovascular disease. So, this study included the variable of HRV of patients without cardiovascular disease.

For this study, the recommended smartwatch to monitor HRV is the Samsung Galaxy Active. The embedded sensors of the smartwatch allow having a constant monitor of the heartbeat. Smartwatches are non-invasive methods that allow having a constant HRV measurement of the subjects. These smartwatches are recommended because of the several embedded sensors and their open-source content.

2.10.2 Personality

Klein, Kotov, and Bufferd [84] conducted extensive research that reviews the literature on the association between personality and depression. Their study summarized cross-sectional associations between depression and three personality traits (Neuroticism/Negative Emotionality, Extraversion/Positive Emotionality, and conscientiousness) and personality types (e.g., depressive personality).

Bagby et al.[85] examined the association between personality and depression to understand, assess, and treat major depression. They concluded that there are personality features that reliably differ from those of healthy samples.

Santor, Bagby, and Joffe [86] assessed the personality relative stability among individuals with depression. They concluded that neuroticism and extraversion measures showed relative stability and absolute change, and changes in neuroticism and extraversion were modestly or not accounted for by changes in depression scores.

The relation between personality and depression is controversial, but this study aims to identify the personality characteristics that are more prominent in students with depression. This study used the Personality Assessment Inventory (PAI) with 22 non-overlapping complete scales, including four validity scales, 11 clinical scales, five treatment scales, and two interpersonal scales. Moreover, the PAI is a validated instrument with a sample of 1051 students with a high degree of consistency. The non-overlapping scales help the machine learning model to discriminate against the levels of the category. The interpretation and administration of the test are conducted by the Psychology department of Tecnologico de Monterrey University.

2.10.3 Physical Activity

There have been many studies trying to analyze the relationship between physical activity and depression. Camacho et al. [87] studied the relationship between physical activity and the risk of being diagnosed with depression. They concluded that changes in activity level could alter the risk of depression by changes in exercise habits; other habits such as social supports, life events, physical health, socioeconomic status, and other health habits did not affect the baseline.

The study conducted by McKercher et al. [88] measures physical activity with an individual self-report and a pedometer (steps/day). The study results were that women with more than 7,500 steps/day were associated with approximately 50% lower prevalence of depression than those with less than 5,000 steps/day. The self-reports concluded that those who performed more than 1.25 hours/week of physical activity were approximately 45% lower prevalence of depression. On the contrary, they did not find any significant associations for men.

The research study presented by Ströle [89] collected previous research studies that gave evidence that physical activity and exercise can also be used to treat depression and anxiety disorders. It concluded that there was no general concept for the therapeutic administration of physical activity, but usually, 3-4 training sessions/week should be performed with a duration of at least 20-30 minutes.

Physical activity is associated with depression, so in this study, the physical activity is relatively monitored with a written test and objectively monitored with the Samsung smartwatches' embedded sensors. The smartwatch register variables are the time and duration of the physical activity, the traveled distance, the burn calorie count, the HRV, the minimum heart rate, and the maximum heart rate.

2.10.4 Sleep Disorders

There is a relation between sleep disorders and depression, particularly in insomnia, that 60%-80% of depressed patients have insomnia, as described in the study conducted by A. Luca, M. Luca, and C. Calandra [90]. They described common sleep disturbances during the depression; particularly, pavor nocturnus, nightmares, hypersomnia, and insomnia are present in clinically depressed patients.

Another study conducted by Vandeputte and Al de Weerd [91] estimated the prevalence of depressive feelings in the various sleep disorders diagnosed in a Center for Sleep and Wake Disorders with the ASDA international classification of sleep disorders. They included 917 patients evaluated with the BDI scale. They concluded that the prevalence of depressive feelings was high in patients diagnosed with a sleep disorder, with no significant differences in age and gender. In the study, only 3.5% of the patients were diagnosed with moderate to severe depression.

Peppard et al. [92] studied the relationship between sleep-related breathing disorder and depression. A sample of 1,408 subjects evaluated with the Zung depression scale with a score of 50 or higher. The sleep-related breathing disorder was characterized by the apnea-hypopnea index (AHI; events per hour). They concluded that compared to participants with no sleep-related breathing disorder, the odds of developing depression were increased by 1.6 to 2.6 folds depending on the severity of the disorder.

Individuals that had sleep disorders are more prompt to develop depression. In this study, we relatively measured the quality and time of the individual sleep and objectively measured the four stages of REM sleep and non-REM sleep (awake, light, and deep). The time that the individual remains in the four stages helped to measure the quality and time of the subject sleep to identify if there is an active sleep disorder or not.

2.10.5 Stress

According to H.M. van Praag [93], stress can cause depression when changes in the stress hormone produced by sustained stress mimic the disturbances observed in depression. The study also concludes that stress hormone disturbances in depression are not a consequence of the depressed state.

A study conducted by Caspi et al. [94] explained why stressful experiences lead to depression in individuals that presented one or two copies of the short allele of the serotonin transporter (5-HTT) gene. Those individuals exhibited more depressive symptoms, diagnosable depression, and suicidality concerning stressful life events.

Another study concluded that stress leads to the releases of hormones and dynorphin, which activates kappa opioid receptors (KORs), prolonged KOR exposure in animal models can lead to the persistent expression of behavioral signs that are characteristic of human depressive disorders [95]

Prolonged exposure to stress could lead to developing depression symptoms. So, for this study, the resilience to stress is measured for each individual. The ProComp equipment measures five variables (electromyography, heart rate, skin conductance, number of breaths, and body temperature) that help determine the student's recovery time to an external stressor used to determine the resilience to mental stress in students.

2.10.6 Conclusion

The identified variables for the proposed methodology are HRV, personality traits, sleeping disorders, physical activity, and stress. The design of the experiment to collect these variables is one of the objectives of this thesis. However, the main focus of this work was on the stress variable, specifically, the creation of a resilience index to quantify mental stress. The following section describes the identified physiological features that aid in capturing the behavior of an individual's mental stress.

2.11 Identified Physiological Features for Resilience to Mental Stress

This section introduces the different physiological features examined in this work, previously linked with psychological stress and resilience to mental stress.

To measure the body's physiological changes in the subject's under study, we employed the biofeedback device: ProComp5 Infiniti System w-BioGraph Infiniti Software T7525 [d]. With this device, we measured five different body responses: peripheral temperature, electromyography, breathing rate, blood volume pulse, and skin conductance. The variables were selected as they have previously been identified to suffer an alteration while an individual is experiencing stress.

[d] https://thoughttechnology.com/procomp5-infiniti-system-w-biograph-infiniti-software-t7525/

2.11.1 Peripheral Temperature

A change in body temperature is experienced in different parts of the body when exposed to stress. Vinkers et al. [96] studied the effects of stress on the core and peripheral body temperatures. The authors concluded that stress exposure decreases the intestines' temperature and specific skin locations, such as fingertips and the finger's base. Vinkers et al. precluded previous studies [97], where they indicated that temperature uniformly rises in response to mental stress. Hence, we can expect that exposure to mental stress lowers the peripheral temperature (fingertip). The study analyzed the data with typical statistics techniques, specifically, ANOVA.

The study induced mental stress with the Trier Social Stress Test [98]. Briefly described, the test consists of a mental challenge where the participants are instructed to prepare a speech (three minutes) and deliver it (five minutes), followed by a mental arithmetic task (five minutes) in front of a committee. The participant is informed that the whole procedure is videotaped and its voice recorded.

2.11.2 Electromyography

Ludenberg et al. [99] conducted a study of the body's biological reaction to stress. The results concluded that the induced stress session significantly increased systolic and diastolic blood pressure, heart rate, urinary catecholamines, salivary cortisol, and electromyography activity compared to the baseline trapezius muscle (indicating muscular tension in the trapezius area).

The induced stressor consisted of mental arithmetic, the Stroop color-word test, the cold pressor test, and standardized test contractions. This research supports the theory that psychological stress contributes to musculoskeletal disorders by enhancing muscular tension in low-load and no-load situations. A more recent study by Hermens et al. [100] conducted similar experiments in 2014 and reached the same conclusion.

2.11.3 Respiratory Rate

The respiratory rate has been of interest, as it increases when an individual is induced to mental stress [101, 102]. The study of Masaoka and Homma [103] concluded that unpleasant emotions alter breathing patterns. The study induced mental stress by amplifying noise using sounds from the environment. They were assuming that noisiness increases with the loudness of the induced sound. These relationships of noise and stress were previously studied by Berglund [104]. For the induced test, the procedure was that each subject was consistently tracked twice during three conditions: baseline of resting-state, three minutes; physical load-mental stress, two minutes; and stress, three minutes. The relationships between the control and the tests were investigated using a repeated-measure analysis of variance (ANOVA).

Besides measuring the muscle response, Ludenberg et al. [99] evaluated the respiration rate. They concluded that the average respiration rate increased after exposure to mental stress and was independent of gender.

2.11.4 Skin Conductance

In regards to skin conductance, the study led by Harker [105] concluded that psychological sweating in response to stress, anxiety, and pain occurs over the whole body. However, it is more evident on the palms, soles, face, and axilla.

Moreover, Ling et al. [106] reached a similar conclusion when inducing mental stress to students between 18-25 years old. A two-way ANOVA was conducted to evaluate the interaction between the responses and the effects of stress. They induced stress with mental arithmetic testing and similar procedures of measuring relaxation and testing phases. They concluded that stress is positively correlated with skin conductance.

2.11.5 Blood Volume Pressure

Hjortskov et al. [107] evaluated how a combined physical and mental workload affected the cardiovascular and subjective stress responses, and the impact of rest. Mental stressors associated with computer work were either added or removed from a standardized computer work session. They concluded that the stressors caused a sustained rise in blood pressure compared to baseline.

Additionally, they observed that while local conditions more influence the blood pressure response in the working muscles, it partly masks mental workload changes. For this purpose, this thesis's experiment is evaluated when the individual is in a resting position.

2.11.6 Conclusion

As per the previous studies, we have identified that the expected variables to increase when an individual is exposed to stress are: heart rate variability, electromyography, skin conductance, and respiratory rate. Moreover, the variable that is expected to decrease is peripheral temperature. Given this, the time a person takes to return its heart rate variability, respiratory rate, muscle tension, temperature, and skin conductance to its baseline parameters can help us define the resilience to the mental stress of a person.

2.12 Related Work

2.12.1 Measuring Depression Related Work

Different studies have shown that depression is associated with more significant cardiac morbidity and mortality, such as Udapa et al., where they recruited 40 patients suffering from major depression, using the DSM-IV-TR criteria. The study used statistical analysis to find a correlation between the measure variables such as heart rate variability, sympathovagal balance, and profound breathing difference. The study shows a correlation between depression and altered sympathovagal balance, linked with higher cardiac morbidity and mortality in myocardial infarction and depression subjects. The impact found in depression was low, but long-term exposure to depression can affect the heart. This study tries to explain the high mortality rate in patients with major depression and cardiac morbidity. However, the study fails to measure a joint neurobiological dysfunction that could impact the heart rate [108].

Similar works were conducted at the University of Melbourne in Australia. The researchers concluded that heart rate variability was reduced with adults with clinical depression compared to healthy controls. Low-frequency heart rate variability was reduced in depressed patients [109].

A study conducted with deep learning algorithms of machine learning has shown a robust predictive model for depression in seafarers from India, with an accuracy of 82.6% with the Catboost algorithm using Python [110]. Various socio-demographic, occupational, and health-related information were collected from 470 individuals. Variables such as age, educational qualification, type of family, marital status, family income, employment, job profile, employment status, rank in the organization, type of vessel, duration of service as a sailor, presence or absence of hypertension, diabetes, ischemic heart disease, and Body Mass Index, were used as input to the model.

A comparison of machine learning methods and statistical linear regression was conducted by Hatton, Paton, et al. in 2019. They found that a machine learning approach of detection was considerably higher than linear regression. The models were developed using baseline demographic data and questionnaire data with 361 participants of 65 and older in England that presented depressive symptoms according to DSM-V criteria and the Mini International Neuropsychiatric Interview. The model obtained only 11% of false-positive [111].

Other methods have been used to detect depression. Dham et al. measured depression with variables such as speech and audio recognition and text analysis [6]. Le [112] created and published a prediction model with a Decision Tree classification and multimodal fusion of audio, video, and text features, which resulted in good accuracy.

Even when similar works have proved a minor link to heart rate variability, no constant measure has been tested. This thesis's focus is the measure of significant data collections, such as the constant monitoring of the heart rate. The initial variables of the study, such as gender, age, personality, and health-related information, are used to construct the predictive model. This study's limitations are that the sample is limited to Monterrey city students, and no other variables are measured, such as socio-demographic variability, older individuals, or different occupations.

2.12.2 Resilience to Mental Stress Related Work

The following literature review concerning mental stress detection with physiological sensors was undertaken to highlight physiological variables related to psychological stress.

The first related work is the one from Healey and Picard. In this study, the authors analyzed physiological variables during real-world driving environments to determine a driver's relative stress level. Electrocardiogram, electromyogram, skin conductance, and respiration were monitored during five-minute intervals of data. This enabled identifying three distinct levels of psychological stress with an accuracy of 97 percent [17]. The researchers concluded that skin conductivity and heart rate were highly correlated with stress levels prediction. Our work uses similar variables to measure relative alteration towards the definition of resilience to mental stress index.

They assess the mental stress level with the questionnaire and video coding. The questionnaire was based on the subject's relative sentiment, where he/she specified the stress level. The questionnaire results were analyzed with ANOVA and determined that the means were significantly different with 95% confidence (p-value > 0.001). The video coding evaluated the subject's movement and assigned a score depending on the number of tasks associated with it. This to estimate the number of stressors per minute for each type of test.

A second study evaluates stress resilience based on similar physiological features during the selection of air traffic controllers [113]. During this study, the authors measured each feature's variability using traditional statistical techniques while assessing the differences between a control and a stress-induced group. The input-induced stimulation included the simulation of different conditions that an air traffic controller is prone to encounter. The study results indicated that there were statistically significant differences ($p < 0.05$) between the groups. This research focused on eight relevant physiological features as tools for stress resilience assessment. While this study provided very insightful information for measuring individual resilience to stress, the study was performed using conventional statistics. In the present study, we use advanced machine learning techniques to capture the subject's response to stress.

The second study applied psychological questionnaires relevant for resilience assessment such as the Connor-Davidson Resilience Scale, Anxiety Sensitivity Index, and Core Self-Evaluations Scale. Nevertheless, the authors do not analyze the relationship between the index and the psychological questionnaires. Instead, they compared it against a control sample of students against air traffic controller candidates. They concluded that the experimental group of air traffic controller candidates are more resilient than the control group, according to their stress resilience assessment.

A final relevant study is the one from Lu, Wang, and You [114]. This study identified resilience based on CD-RISC's traits, studied physiological recovery to stress. The third related study subjects were subjected to stress in a protocol with seven stages: baseline, stress anticipation 1, stress 1, post-stress 1, stress anticipation 2, stress 2, and post-stress 2. The mental stress was induced with the Trier Social Stress Test [98].

The results indicate that high-trait-resilient participants have a complete recovery from the first and second stress of the systolic and diastolic blood pressure compared to low-trait-resilient subjects. The study concluded that an adaptive physiological response pattern to recurrent stress is found in high-trait-resilient individuals based on electrocardiogram data [114].

These studies' findings suggest that there are mental stress and resilience traits to mental stress that can be measured based on physiological measures that are harder to control by people when inducing mental stress.

Chapter 3

Methodology

Since this thesis aims to create both a framework for detecting depression and creating an index that quantifies the resiliency to mental stress, we have documented two different methodologies. The former focuses exclusively on collecting the data, as this is the beginning of developing the following project. The second methodology focused on the analysis of resiliency. It follows the steps from the collection, the exploration of the data, the RMSI evaluation, and the analysis of the results.

3.1 The Methodology for Collecting the Data for the Depression Experiment

This section describes the proposed methodology to collect the identified variables that could aid in detecting depression. Smartwatches' embedded sensors can collect the identified features. For this application, the Samsung Galaxy Active 2 was chosen for its open-source platforms and the technical characteristics of its analog sensors.

The design methodology is centered on a sample of adult students currently enrolled from Tecnológico de Monterrey University. In this experiment, the number of selected individuals was chosen according to the number of smartwatches available at the Computer Science department. However, a sample bigger than 50 individuals is recommended due to the quantity of extracted data. Even when the sample is not statistically representative of the population, the quantity of data per individual can help to detect symptoms of depression instead of the level of depression.

The first step to take is to make the call. A public call must be sent to the University's active students that are physically located in the same region. The purpose of the call is to obtain a sample, from which at least 50% of them must be diagnosed with depression. The criteria for the participants must be restricted to those who do not present a cardiovascular condition, non-smokers, or have no respiratory issues. These criteria are considered because they could impact the HRV measurements. The control group, on the other hand, must be formed of 11 healthy individuals without depression. The gender of the sample must be well distributed to avoid gender bias.

To measure the level of depression, we selected the Beck Depression Inventory (BDI-II) test for this methodology because of its validity in Hispanic samples. The BDI-II was applied online in the ITESM digital platform and could be replicated in online platforms. After identifying those individuals with the level of depression and categorizing it, we can consider those individuals as candidates for the data collection experiment. Gender, age, and physical activity relative baseline must be collected with a simple medical diagnosis. Qualified personnel must perform the medical diagnosis.

The several tests to collect the data can be applied in parallel. The tests are independent of each other. So, the timeline is defined based on the availability of each participant.

The variable of stress, specifically, the resilience to mental stress test objective, is to measure the physiological features to capture an individual's response to mental stress in the face of analytical challenges. A biofeedback device can be employed to capture the identified physiological variables that have proven to have a relationship to mental stress. Five sets of data will be acquired from the device: electromyography, blood volume pulse, breathing rate, skin conductance, and peripheral temperature. The needed test to induced mental stress takes 10 minutes to apply and approximately five minutes to connect the biofeedback equipment sensors. The test details are explained in the next section, as it is the basis to collect the data to create the resilience to mental stress index.

Another variable to collect is the traits of personality. For this, the PAI test could be used to assess the principal components of the subject's personality. The test takes between 50 to 60 minutes to complete, and its application must be supervised by trained personnel.

The smartwatch can capture physical activity, HRV, and sleeping disorders. A mobile Android application was created to capture these variables by the research group. Images from the design of the application are in Appendix A. Not only was the platform developed to collect the smartwatch embedded sensors data, but also to capture the daily log. The daily log was designed to capture the relative sentiment of each participant through the data collection experiment with the smartwatch. The participants were asked to install the mobile app on their cellphones. The data captured by the mobile app will upload the information to the cloud to access and monitor the experiment's development. Derived from the accelerometer, we obtained the physical activity time, calorie burn, speed, and distance the subject performed. The smartwatch was set to capture and record the physical activity after 10 minutes of constant movement sensed by the accelerometer. Samsung smartwatch development interface logged the sleep stages of each individual. Hence, we took advantage of the development and included the start and end times of the subject's sleep and each sleep stage provided by the Samsung interface.

The daily log was created to measure physical activity, sleep time and quality, and the severity of a depressive episode. This allows emphasizing the smartwatch activity data with the logged depression episode described by the subject in the specified period. The test duration to collect the physical activity, sleeping disorders, and HRV with the smartwatch must be at least two full weeks' worth of data. The timeframe was selected because depression is diagnosed only after two weeks of persistent symptoms.

Data collection from the smartwatches of two weeks and initial conditions will generate thousands of data from each individual. The CRISP-DM methodology for data analysis could be implemented to process the data.

The design of the experiment follows the six phases of the CRISP-DM [115] methodology, as shown in figure 3.1. The first phase of the CRISP-DM methodology is understanding the project developed in the section of problem statements and definitions in this thesis. The data will be structured in tables to categorize each individual and its variables acquired from the tests and the smartwatch's embedded sensors. Fundamental statistical analysis needs to be performed to familiarize the information and extract general knowledge from the data.

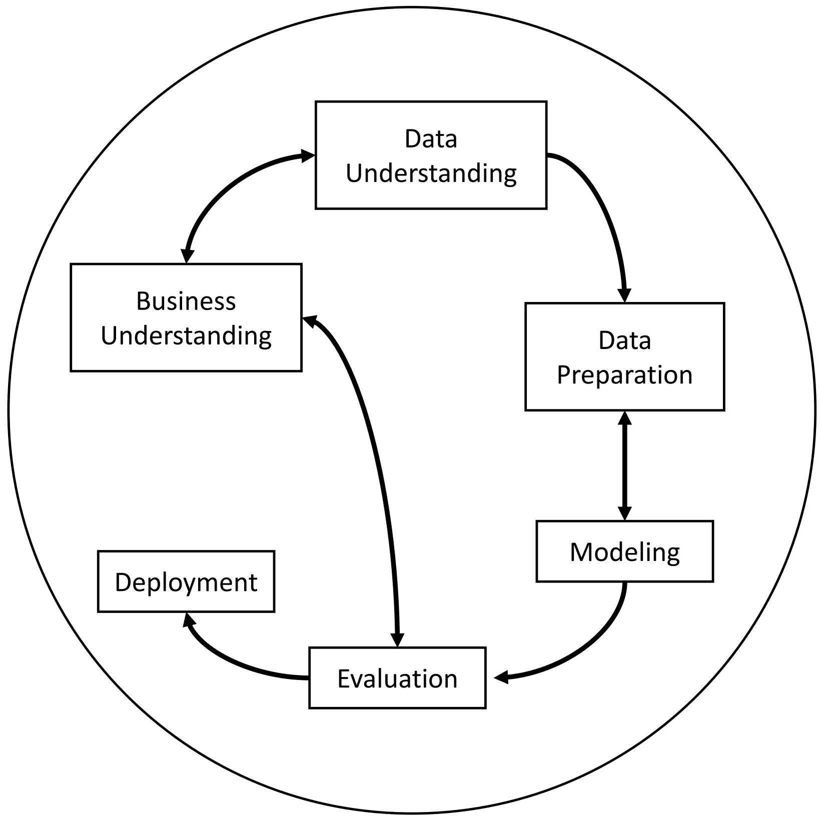


Figure 3.1: Flow diagram of the six phases of the CRISP-DM methodology.

Once the data preprocessing is complete, it can train different machine learning algorithms to create the modeling. The type of data will be structured, and the type of machine learning problem is supervised. Different types of classification and regression algorithms could be tested for acquiring a robust model.

The most accurate algorithm will be used for the platform to predict the level of depression in students. The trained algorithm will acquire the data for each individual and accurately predict the level of depression to alert the student and a university counselor to conduct a deeper analysis and treatment for their condition.

A summary of the experiment design and its expected output variables is presented next.

Objective. Acquire a complete dataset of 22 individuals.

Raw Expected Data. The raw expected data corresponds to all the data analyzed and processed to obtain the features that will train the model. The collected data is from the embedded sensors from the biofeedback device and the Samsung smartwatch to measure biological data, Beck Depression Inventory (BDI) tests, and Personality Assessment Inventory (PAI) test to measure the severity of depression and the type of personality, respectively. One medical test to assess physical health and one daily log to relatively assess of the physical activity, sentiment, sleep quality and quantity, and depression severity of the individual. The specific variables and the type of response are listed next.

- ProComp Sensor Variables:
  - Surface Electromyography (SEMG)
    - Type: Numerical
    - Unit: millivolts (mV)
  - Heart Rate
    - Type: Numerical
    - Unit: Beats per Minute (BPM)
  - Number of breaths
    - Type: Numerical
    - Unit: Number of Breaths per Minute (BRPM)
  - Skin conductance
    - Type: Numerical
    - Unite: microohms ($\mu\Omega$)
  - Body Temperature
    - Type: Numerical
    - Unit: Fahrenheit Degrees (°F)
- Smartwatch Sensor Variables:
  - Accelerometer
    - Type: Numerical
    - Unit: Meters over squared seconds (m/s^2)
  - Barometer
    - Type: Numerical
    - Unit: millibars (mbar)
  - Heart Rate Sensor
    - Type: Numerical
    - Unit: Beats per Minute (BPM)
- PAI Test Answers:
  - Type: Categorical
  - Unit: 1-“Not true at all, False”, 2-“Slightly true”, 3-“Mainly true”, and 4-“Very true”. (1,2,3,4)
- BDI assessment to label data for supervised problem

  - Type: Categorical
  - Unit: Severe, Moderate, or Mild (SV/MO/MI)

Factors Risks:

- The daily log could be biased by the respondent.
- The medical diagnosis could be biased by the respondent.

Responses or output measures:

- Heart rate variability
  - Type of response: Numerical
  - Unit: Beats per minute (BPM)
- Sleeping disorder
  - Type of response: Binary
  - Unit: Yes or No (1/0)
- Physical Activity
  - Type of response: Numerical
  - Unit: Minutes per day (MIN/DAY)
- Resilience to stress
  - Type of response: Numerical
  - Unit: Without unit (0-1)
- Type of personality
  - Type: Numerical Categorical
    * Somatic Complaints
    * Anxiety
    * Anxiety-Related Disorders
    * Depression
    * Mania
    * Paranoia
    * Schizophrenia
    * Borderline Features
    * Antisocial Features
    * Alcohol Problems
    * Drug Problems

    * Aggression
    * Suicidal Ideation
    * Stress
    * Nonsupport
    * Treatment Rejection
    * Dominance
    * Warmth
    * Inconsistency
    * Infrequency
    * Negative Impression
    * Positive Impression
  - Unit: Range 20-110

- Beck Depression Inventory assessment
  - Type of response: categorical
  - Unit: Severe, Moderate, or Mild (SV/MO/MI)

A daily log of eight questions was designed to keep track of the participants' negative emotions and register relative variations of physical activity and sleep quality, and quantity. The daily log aims to compare the extracted data from the smartwatch embedded sensors to the participant's relative perspective. The relative time of exercise of the participant is asked at the beginning of the medical diagnosis. The daily log is presented in Appendix B (Spanish only).

Initial medical diagnosis. Potential candidates for the experiment are subject to an initial evaluation to assess if the subject is healthy. The exclusion criteria to test in the medical diagnosis are: if the individual has a heart condition that impacts the heart rate, if the individual has diabetes, if the individual has weight problems, or is a smoker. If the candidate presents any of the exclusion criteria, it is removed from the experiment.

Samsung algorithms process the sensors' information to return two crucial variables: the sleep stages and exercise time. The smartwatch record four stages of sleep: awake, REM, light, and deep. Additionally, it records the time in minutes of daily physical activity.

The PAI test will characterize the traits of the personality of each participant. The PAI test considers a Likert scale of 4 points for each question (False, no true at all; slightly true; pretty true; and entirely true), which allows a better precision than other binary scales. The test was designed to cover all items that help to diagnose mental disorders. It consists of five feature scales: validity of responses, clinical syndromes, interpersonal style, complications in treatment, and the characteristics of the person's environment.

The PAI test provides features that are helpful to characterize a person's personality traits, the tendency to clinical mental disorders, and characteristics of the subject's environment that could be essential features that help build the machine learning model. The validity scales measure the respondent's approach to the test. It includes the degree of inconsistency, infrequency, positive impression, and negative impression.

The clinical scales provide the clinical mental disease that shows stability and importance in the contemporary diagnosis of mental health diseases. The subscales of the clinical scale include relevant variables to this study, such as depression and anxiety. However, it can be categorized into three broad classes of disorders: those within the neurotic spectrum, the psychotic spectrum, and those associated with behavior disorder or impulse control problems.

The treatment scales provide potential complications in treatment, such as the potential to self-harm or harm others, the subject's environmental circumstances, and an indicator of the subject's motivation for treatment.

The interpersonal scales indicate information on the respondent's relationships and interactions. Two dimensions are provided: a warmly affiliative or a cold rejecting style and a dominating or a submissive style—additionally, two other scales to assess pathology: borderline features and antisocial features.

The BDI-II results categorize the level of depression. The scale ranges from 0 to 64, where a higher score means a higher degree of depression. Table 3.1 shows the considered thresholds of this study to categorize the severity of depression in minimum, mild, moderate, and severe:

Table 3.1: Beck depression inventory score thresholds to classify the level of depression.

Total Score	Depression Level
0-13	Minimum
14-19	Mild
20-28	Moderate
29-63	Severe

The steps of the experiment for the data collection are as follow:

1. **Call for participants**. Call adult students. The initial call is for every candidate to make an initial assessment through the online BDI-II test. Students who get a mild, moderate, or severe depression level from the BDI-II test are labeled with depression, and candidates that show no symptoms of depression are selected for the control group. Both groups depressive and control groups are invited to participate in the experiment's next steps.

2. **Collection of data**.

   (a) All participants must sign the Agreement of Consent and Non-Disclosure Agreement design by each researcher to protect and use personal data to continue.

   (b) Conduct the medical diagnosis to evaluate the participants' physical health. Previously stated exclusion criteria are applied at this stage. The exercise time per week is also logged as an initial assessment and if the candidate has a clinical diagnosis of depression.

   (c) Apply psychophysiological mental stress test to detect the individual's resilience using the biofeedback device to filter the five physiological variables. The device will obtain 12 minutes' worth of data with a sample rate of 256 observations per second for each measured feature. The test is explained in section 3.2.

(d) Apply PAI.

(e) Apply the online BDI-II test the same day that the collection from the smartwatches begins.

(f) Deliver the smartwatch for the two-week collection of data. The smartwatch's demonstration and training are needed to collect physical activity, sleep disorders, and unregular HRV. The collection of the data is automatic when the cellphone application detects a Wi-Fi connection.

(g) At the end of the two-week collection, apply online BDI-II again.

Table 3.2: Psychophysiological mental stress phases description.

Event Time	Event Label	Event Type
0:00:00	Baseline. Step: Baseline record.	Activity
0:02:00	Stressor phase 01. Step: Registry of the stressor.	Activity
0:04:00	Recovery phase 01. Step: Registry of the recovery phase.	Activity
0:06:00	Stressor phase 02. Step: Registry of the stressor.	Activity
0:08:00	Recovery phase 02. Step: Registry of the recovery phase.	Activity
0:10:00	Final Phase. Step: Final phase record.	Activity

In an investigation conducted by Beck [116] to estimate the effects of memory on the BDI test, it was concluded that there was no relation between memory and the type of answer of the respondents. Nevertheless, Hatzenbuehler, Parpal, and Mathews [117] informed that the test's repetitions resulted in a diminished score in consecutive tests. Hence, consecutive tests are compared to avoid a bias in the BDI-II results, and one week is elapsed until the participant can retake the test.

3.2 Methodology to calculate RMSI

This section explains the dataset analyzed, the methodology employed, and the non-supervised learning techniques used to measure resilience to mental stress. Moreover, in this section, we explore different approaches to transform dynamic data and propose three methods to compute the RMSI.

3.2.1 Dataset

The dataset used for this study is a collection of physiological features from 71 students from Tecnologico de Monterrey University in Mexico. It has a gender distribution of 37 males and 34 females. The age of the participants ranges from 18 to 28 years old. The subjects were excluded if they have a cardiovascular disease that could impact the heart rate.

For each of the subjects, five variables were measured using a biofeedback device. The biofeedback device has a sampling rate of 256 observations per second from a continuous signal. This rate allowed the five sensors (even slow signals, with no high-frequency component) to have sampling rates without precision loss. The analog signal was processed by the encoder (ProComp Infiniti) and finally sent to the computer.

The five variables measured are skin conductance, blood volume pulse, peripheral temperature, electromyography, and breathing rate. These variables were assessed with analog sensors and captured by the analog-digital converter of the biofeedback equipment. The measurement obtained from each variable is described in Table 3.3 and these variables are described in detail in the following section.

Table 3.3: Description of the five measured variables of the biofeedback device.

	Skin Conductance (SC)	**Blood Volume Pulse (BVP)**	**Peripheral Temperature (PT)**	**Electromyography (El)**	**Breathing Rate (BR)**
Description	Measures the peripheral skin conductance	Measures changes in the arterial translucency	Measures changes in corporal temperature	Measures muscle response	Measures the expansion and contraction of the rib cage
Sensor	Temp-Flex/Pro	BVP-Flex/Pro	Temp-Flex/Pro	MyoScan-Pro	Respiration-Flex/Pro
Unit of Measure	micro-Siemens	Volts	Centigrade Degrees	Volts	Ohms
Range of values	$0\,\mu$ S - $30\,\mu$ S	0 - 100 %	10 - 45°C	$0\,\mu$ V - $1600\,\mu$ V	0 - 100 %
Typical Relaxed Measures	2μ S	30 - 60 %	18°C	3 - $5\,\mu$ V	No typical waveform

Figure 3.2 illustrates a single subject's data during the psychophysiological stress test. The five measured features are represented in each slot through time (256 samples per second). We can observe from this figure the presence of clear spikes in electromyography, blood volume pulse, and breathing signals. This indicates that the signals need to be filtered and preprocessed before the machine learning model experimentation.

3.2.2 Physiological Signals

Skin Conductance

The skin conductance sensor measures the skin's ability to conduct electricity. The device determines the change of the peripheral temperature as changes in an electrical current occur

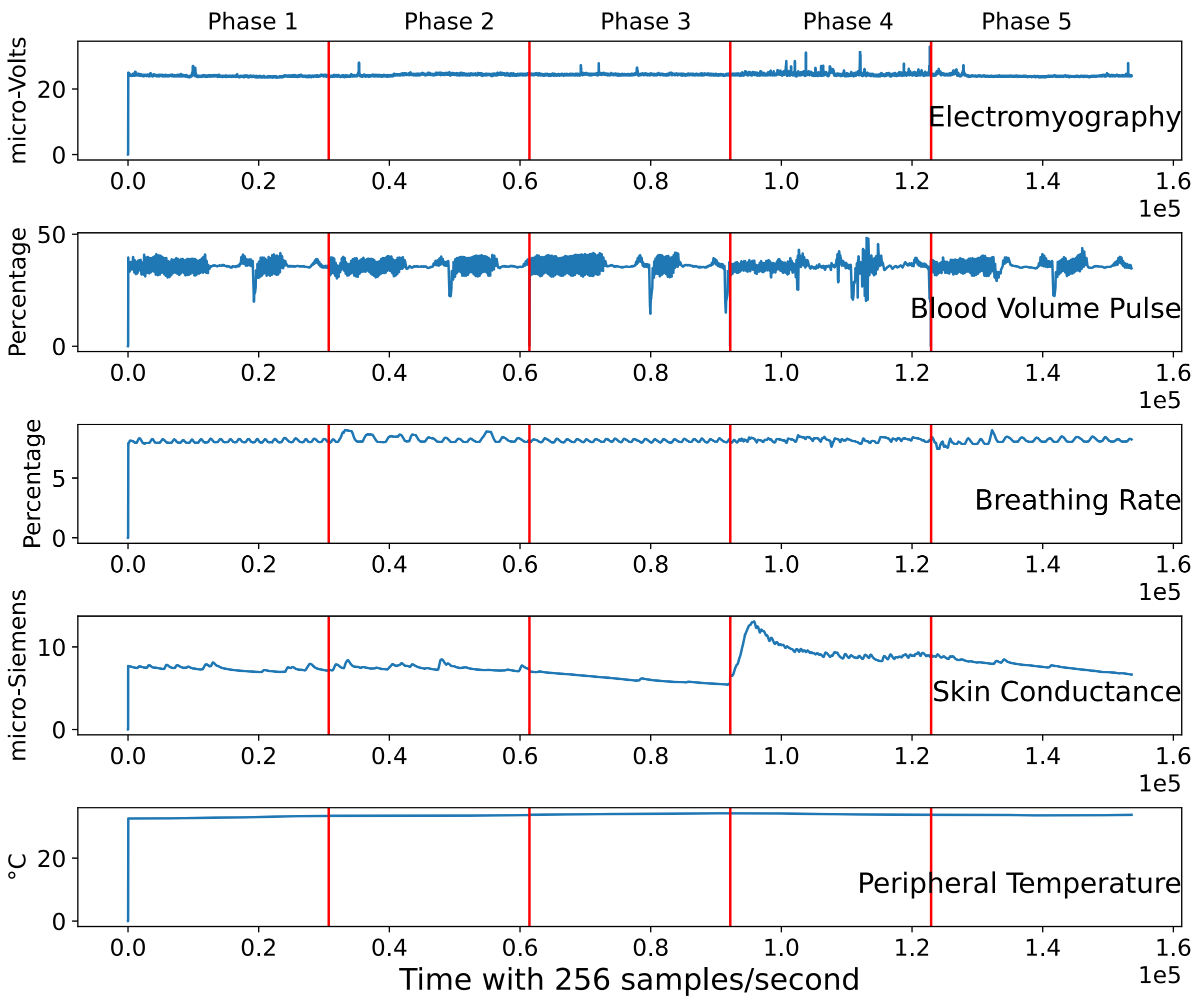


Figure 3.2: Raw signal of the five physiological features of one subject psychophysiological stress test.

[118]. The signal is encoded with the biofeedback device and captured in the computer. Figure 3.3 shows the signal in micro-Siemens against samples per second.

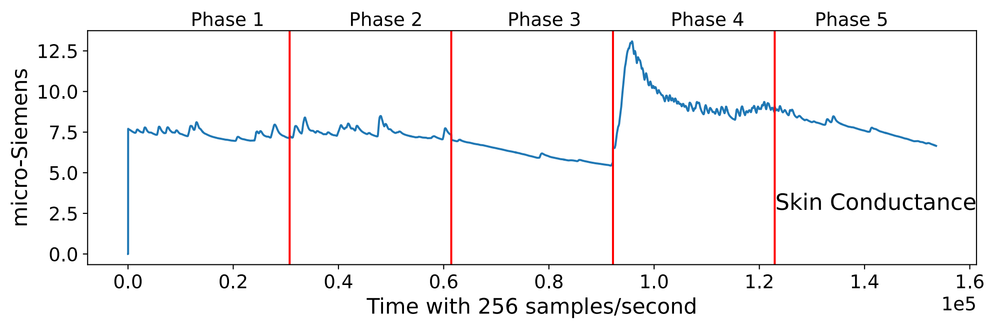


Figure 3.3: Skin conductance raw signal of one subject psychophysiological stress test.

A micro-Siemens unit is equivalent to a unit of the inverse of micro-ohms. The principle is to apply a small current through two electrodes, usually strapped to two fingers of one hand, and measure its conductance variation. As a person reacts to stress, the skin's conductance increases, as the individual tends to sweat.

Peripheral Temperature

The sensor converts changes in temperature to electrical current changes to be encoded by the biofeedback device module. The sensor is placed in the palm of any finger and varies according to the amount of blood flowing through the skin. As a person reacts to stress, their fingers tend to get colder. Usually, peripheral temperature changes are moderately slow. Figure 3.4 illustrates the change in temperature of the individual.

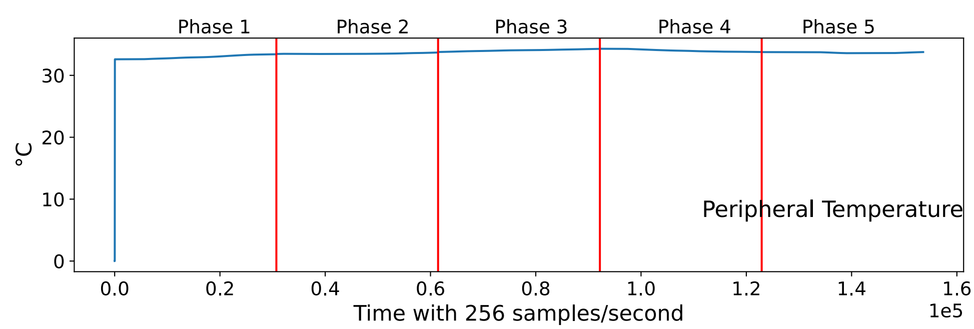


Figure 3.4: Peripheral temperature raw signal of one subject psychophysiological stress test.

Blood Volume Pulse

The sensor's principle is to bounce infrared light against the skin surface and measure the amount of reflected light. The amount of blood present in the skin varies accordingly with the heartbeat. When a heart pulse occurs, there is more blood present in the skin, blood reflects red light, and more light is reflected. Between pulses, the amount of blood in the veins decreases, and more red light is absorbed, then the amount of light returning to the sensor is lower [119]. Figure 3.5 illustrates the blood volume pulse signal's raw input in percentage.

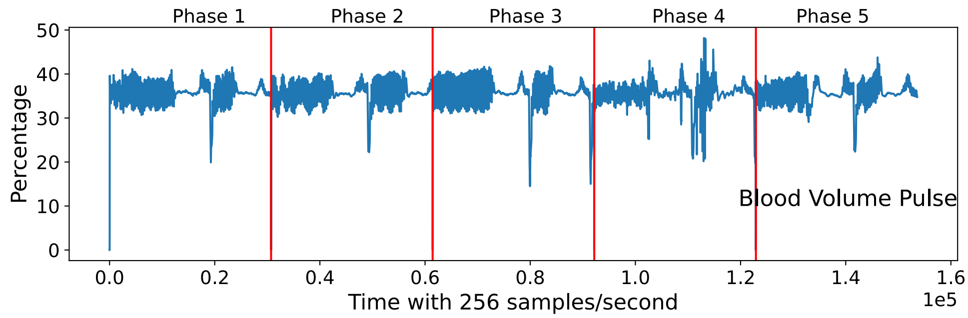


Figure 3.5: Blood volume pulse raw signal of one subject psychophysiological stress test.

The blood volume pulse is a relative measure; it does not have a standard unit. The typical signal shows a substantial rise with the systolic contraction, followed by a slower fall. Usually, with changes in sympathetic arousal, the peak-to-peak intensity of the signal can increase and decrease.

Breathing Rate

The respiration sensor is stretch-sensitive. It can translate the expansion and contraction of the rib cage or abdominal area to an increase and fall of the signal while strapped across a person's chest or abdomen. A relative indicator of chest expansion is the respiration signal. The biofeedback equipment does not produce standard measurement units for respiration.

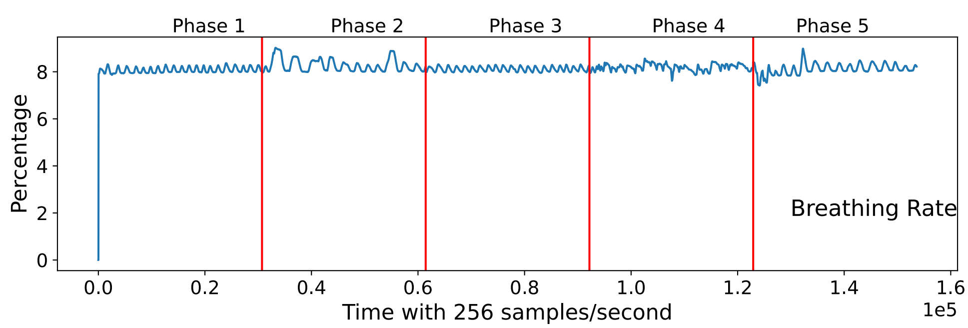


Figure 3.6: Breathing frequency raw signal of one subject psychophysiological stress test.

Several breathing patterns can be identified, but the respiration signal has no standard waveform. Usually, a quick rise slows at the top of the breath, followed by a rapid decline near the end of the breath. When the participant focuses on a task or speaks, the breathing rhythm will change itself.

Electromyography

The sensors measure muscle activity by detecting the electrical impulses generated by muscle fibers when they contract. Since all muscle fibers contract at various speeds within the sensor's recording area, the sensor's signal continuously changes the potential difference between its positive and negative electrodes. The amount of muscle fibers recruited during any given contraction depends on the force needed for the motion to be done. For this effect, the power (amplitude) of the resulting electrical signal is proportional to the contraction frequency [120]. Figure 3.7 illustrates the biofeedback device's raw signal that measures muscle contraction changes.

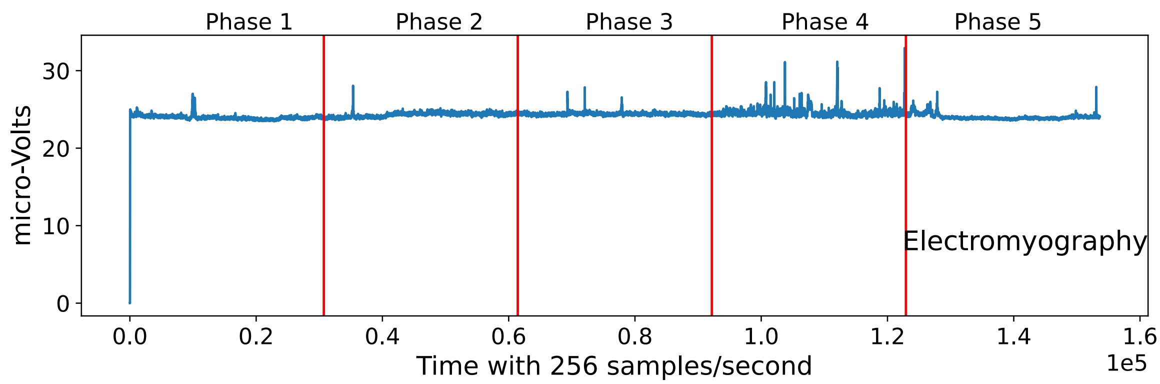


Figure 3.7: Electromyography raw signal of one subject psychophysiological stress test.

3.2.3 Psychophysiological Mental Stress Test

The psychophysiological mental stress test is a protocol to measure autonomic reactivity and recovery to psychological stress through a psychophysiological mental stress profile that includes the simultaneous recording of five physiological responses (peripheral temperature, blood volume pulse, respiratory rate, muscle contraction, and the galvanic response of the skin).

Autonomic Reactivity to Mental Stress

The conceptual definition of autonomic reactivity to stress is the magnitude of physiological activation and recovery responses controlled by the autonomous nervous system in the presence of stressful situations [121].

The operational definition is that the autonomic reactivity to stress will be measured from the sympathetic nervous system's magnitude of responses to discrete environmental stressors; in particular, the changes that are produced in the following physiological responses: galvanic response of the skin, systolic and diastolic blood pressure, heart rate, heart rate variability, and respiratory chest rate. To measure these variables, we employ the Computerized Biological Feedback equipment.

Autonomous Recovery from Mental Stress

The conceptual definition of the autonomous recovery from stress is the measurement of a set of responses inhibited by the autonomous nervous system in its parasympathetic branch through relaxation [121].

The operational definition is that the autonomous recovery from stress is measured from the degree of recovery of basal levels within a previously established period after applying a stressor. In particular, we focus on the changes produced in the following physiological responses: galvanic response of the skin, systolic and diastolic blood pressure, heart rate, heart rate variability, and respiratory chest rate.

3.2.4 Mental Stress Protocol

Stress is induced by challenging the subject with mathematical operations and language tasks. During the test, the evaluator needs to guarantee a constant pace and obtain quick and correct answers without any distractions such as laughing or moving too much.

The protocol consists of ten minutes covering five phases; each phase has a total duration of two minutes. There are two stages intended to induce external stressors, two recovery phases after the stressor and one to build a baseline.

The first phase consists of the baseline or calibration phase; the evaluator does not give instructions besides that the subject must remain seated and relaxed. This phase allows recording the physiological baseline of each subject. Before starting the second phase, the evaluator gives the instructions for the mathematical and language tasks. The induced stressor is applied in phases two and four; phases 3 and 5 are recovery periods. The protocol is illustrated in Figure 3.8, and the complete guide designed by the Psychology department of Tecnológico de Monterrey is shown in Appendix A.

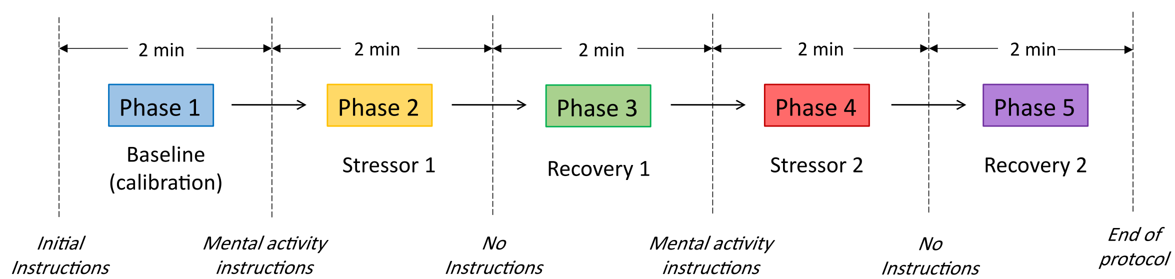


Figure 3.8: Flow diagram of the psychophysiological mental stress protocol.

3.2.5 Signal Preprocessing

Median Filter

This section describes the data preprocessing of the subject's data obtained using the biofeedback device. As shown in Figure 3.2, the signal must be preprocessed before using the data to avoid pseudo detection of peaks or noise. Hence, a median filtering technique is applied to remove peaks and noise from the signal. The kernel size of the filter was selected according to Equation 3.1 [122].

$$w = \frac{1}{4f_s} * length(v) \tag{3.1}$$

Where f_s is the sampling frequency and $length(v)$ is the total number of instances.

The kernel size must be odd as it selects its neighbors to get the median. Hence, a window size of 151 was selected. Additionally, the biofeedback device's raw signals have an offset at the beginning of the sensor measurements. Therefore, for this analysis, the first 0.5 seconds of data were removed.

Standard Scaler

The unit of measure of each sensor is on a different scale; therefore, a critical step to compare the subjects is standardization. For this, a standard scaler was applied after filtering the data. Equation 3.2 is used to standardize each point of all features. Standard scaling is a way of normalizing features by deleting their mean and scaling their variance to one. Because the normalized value is determined solely by the mean and variance, it has some advantages, including being linear, reversible, rapid, and highly scalable [123]:

$$z = \frac{x - a}{s} \tag{3.2}$$

where a is the mean of the feature, and s is the standard deviation of the variable.

Preprocessing Result

We can observe in Figure 3.9 that peak points in blood volume pulse and electromyography signals are reduced. Moreover, it illustrates the time division of the corresponding phase during the psychophysiological stress test. Phase one corresponds to the first 30,720 instances (first 2 minutes). At the beginning of phase two, an external stressor is applied to the subject. Each variable suffers alterations in the second phase in comparison to the signals in the first phase. These signal changes between phases are critical to capture the severity of stress that the individual suffers through the test. Therefore, we can quantify these variables' alteration using different techniques explained in Chapter 4.

It is important to notice that the signal offset is not an issue in calculating the RMSI or the AF because the distances are relative to each subject's first phase (baseline).

3.2.6 PCA Visualizations

To understand the five signals' behavior, we can transform the data with the principal components technique and visualize it in a bi-dimensional graph. Several steps were taken to ensure that the data comply with all the PCA assumptions. One of the subjects of study of this experiment is taken as an example through all the procedures.

Sample Rate Transformation

For visual representation purposes, the sample rate was reduced from 256 samples per second to one. The median of 256 samples per second was calculated to reduce the number of points. This allows us to avoid excessive points in the PCA plot.

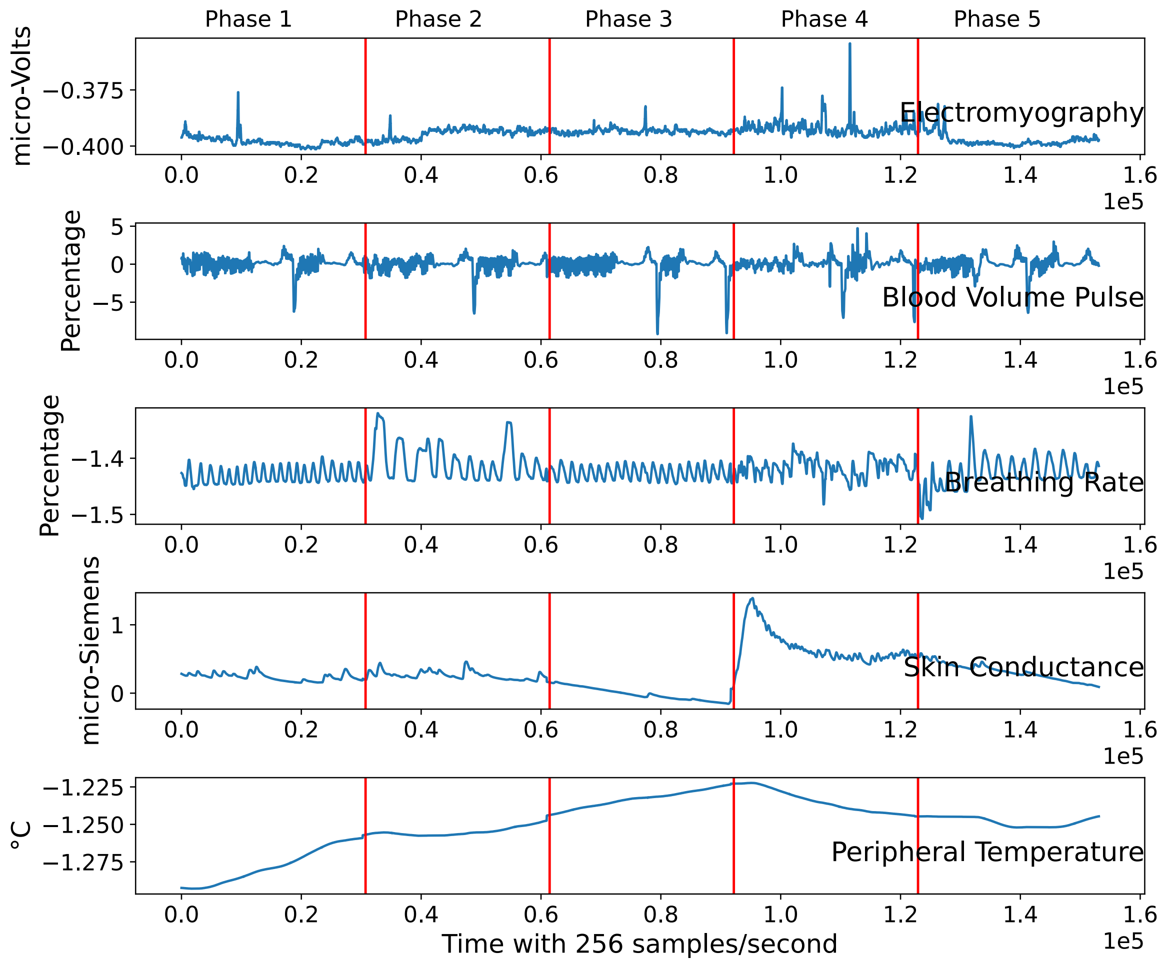


Figure 3.9: Filtered and standardized signals of the five physiological features divided by the psychophysiological stress test phases by vertical red line.

Exploration of the Data

After the sample rate transformation, to construct the PCA, we need to test for multicollinearity with the variance inflation factor (VIF) and correlation with Pearson, if the data has a normal distribution or Spearman, if the data does not have a normal distribution.

The Anderson-Darling (AD) test is applied to test if the data has a normal distribution. The AD test is used to test the null hypothesis that a sample is drawn from a population with a specific distribution [124], in this case, a normal distribution. The results of the AD test are shown in Table 3.4. From this point onwards, the following abbreviations are used to refer to the five physiological features: electromyography (El), blood volume pulse (BVP), breathing rate (BR), skin conductance (SC), and peripheral temperature (PT).

Table 3.4: Anderson-Darling test results of the five physiological features of one subject.

Feature	**El**	**BVP**	**BR**	**SC**	**PT**
AD Statistic	13.5217	53.4392	36.1838	13.6883	18.0058
P-value	> 0.0005	> 0.0005	> 0.0005	> 0.0005	> 0.0005
Conclusion	Reject H_0	Reject H_0	Reject H_0	Reject H_0	Reject H_0

We conclude that there is insufficient evidence to conclude that the population distribution is normal because we failed to accept the null hypothesis. As a result, we proceed to obtain the Spearman Rank correlation. The Pearson correlation between the rank values of two variables is equal to the Spearman correlation between those two variables. In contrast, Pearson's correlation evaluates linear relationships, Spearman's correlation evaluates monotonic relationships [125]. The matrix illustrated in Figure 3.10 shows the Spearman Rank coefficient between attributes.

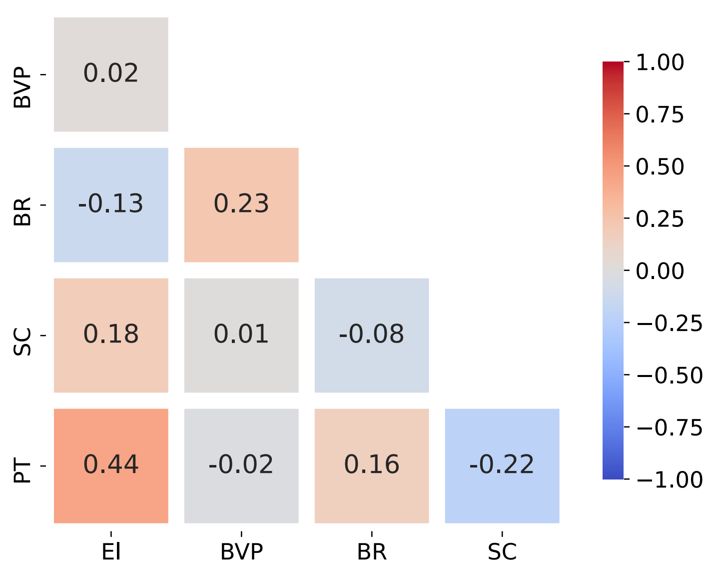


Figure 3.10: Spearman Rank coefficient results of one subjects of the five physiological features.

It can be observed that the highest coefficient has a positive correlation, which is between

PT and El (0.44). The following highest coefficients positively correlate between BVP and BR (0.23) and a negative correlation between SC and PT (-0.22). Furthermore, BVP did not have correlations bigger than $|0.02|$. This indicates that most of the variables are correlated with each other, but BVP has almost no correlation.

Another critical concept is multicollinearity. If there is multicollinearity between the features, the results of calculating Euclidean distances between vectors could not be confident [126]. As the principal components' procedure is to create a set of independent vectors that contain the linear combination of the original features, this unsupervised machine learning technique can be used to transform the original data to remove multicollinearity [127]. The variance inflation factor (VIF) is used to detect multicollinearity [128]. In this study, the Statsmodel[a] Python package tool is employed to obtain the VIF. The resulting VIF of the five physiological variables is shown in Table 3.5

Table 3.5: VIF results of a subject's physiological features with low multicollinearity.

Feature	**El**	**BVP**	**BR**	**SC**	**PT**
VIF	1.4637	1.0422	1.1771	1.1261	1.3128

In this example, the analyzed subject has low VIF's, which indicates low multicollinearity between variables. Nevertheless, this is not the case for all test subjects. For example, Table 3.6 shows the corresponding VIF of another example of data from a different subject.

Table 3.6: VIF results of another subject's physiological features with higher multicollinearity.

Feature	**El**	**BVP**	**BR**	**SC**	**PT**
VIF	1.3509	1.0262	370.3800	1.8049	387.1257

In Figure 3.11, it can be observed that the highest coefficients have a negative correlation, which is between SC-PT and SC-BR—followed by a positive correlation between BR-PT, El-SC, and El-BR. Furthermore, BVP does not have a correlation coefficient bigger than $|0.07|$. This indicates that most of the variables are correlated with each other, but BVP has almost no correlation. Additionally, BR and PT have multicollinearity according to the VIF. Hence, we need to transform all variables to avoid multicollinearity problems while applying PCA.

On the other hand, even when we have higher Spearman coefficient correlation values in some cases, it is recommended to transform all features to increase their linear relationship. Hence, before applying PCA, the data needed to be transformed for linearity with a power transformation to comply with the assumptions. For this, a power transformer with the Yeo-Johnson method of the ScikitLearn module[b] was applied.

Now, after scaling and transforming, PCA can be applied to the data. After fitting PCA for all subjects, we obtained a mean explained variance of 0.6384 for two components, which is enough to understand the data behavior visually. Figure 3.12 shows the explained variance ratio of all samples, and Figure 3.13 illustrates the resulting plot of bi-dimensional space. The five physiological features of these subjects do not have a normal distribution.

[a]https://www.statsmodels.org/stable/generated/statsmodels.stats.outliers_influence.variance_inflation_factor.html
[b]https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.PowerTransformer.html

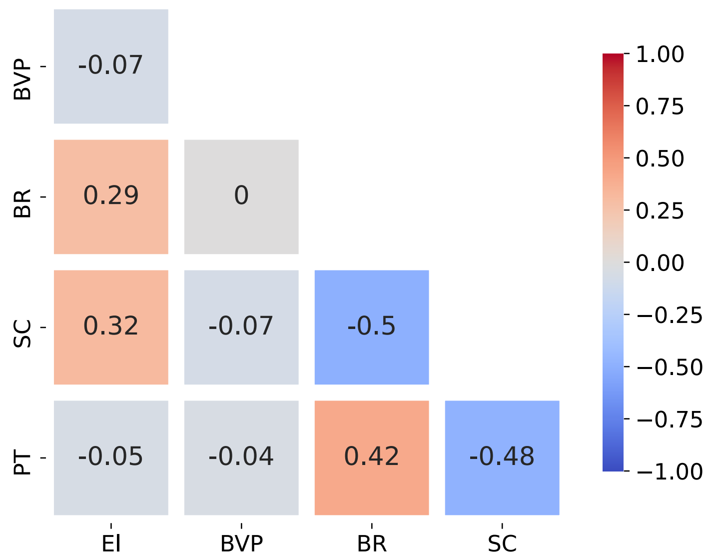


Figure 3.11: Spearman Rank coefficient results of another subject of the five physiological features with stronger correlations.

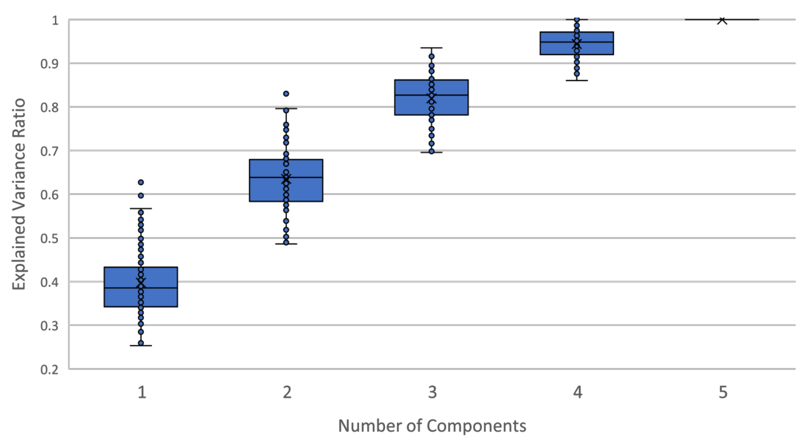
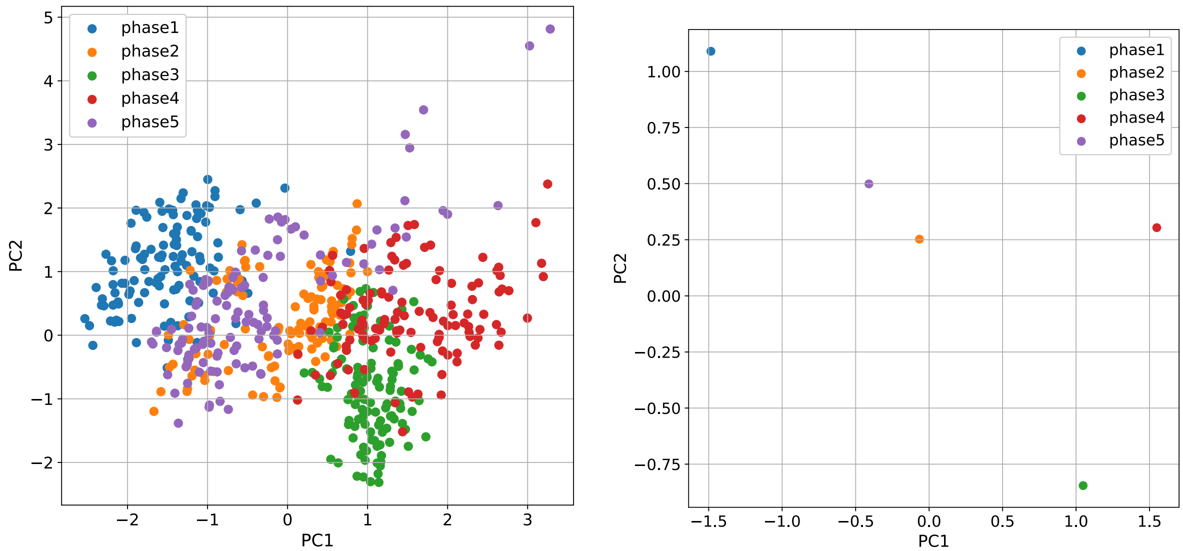


Figure 3.12: Explained variance ratio of constructed principal component analysis technique with the five physiological features for visualization.



(a) (b)

Figure 3.13: Two-dimensional space of PCA of the analyzed subject (a) and its corresponding centroid (b).

The PCA graph shows that each point consists of a second with two dimensions in the first principal component (PC1) and the second principal component (PC2). The color defines the corresponding phase of the stress test. There is an evident displacement be- tween phases, meaning that the PCA could capture all the measured variables' behavior in a two-dimensional space. This bahavior in two-dimensional spaces can be extrapolated to an n-dimensional space with the five measured variables, but a two-dimensional space was created for bahavior visualization. Hence, we can measure the relative distance between centroids points between phases using all measured variables.

The complete profile of the study subject's preprocessed signals can be observed in Figure 3.14. After 120 seconds (30,720 samples), the signals suffer a drastic change. The first phase consists of a baseline where the individual is calm. After 120 seconds, phase 2 begins inducing external mental stress with analytical problems. It is clear to notice how prominent the subject changes variables at the beginning of phase 2, meaning the subject is suffering a relative change in its variables. These variations are reflected in the PCA's two-dimensional space (Figure 2.2).

The relative displacement between phases can be compared to the baseline phase (phase1). It is essential to notice that an n-dimensional space with 256 samples per second, instead of one sample per second, allows us to avoid losing information when calculating relative distances between phases when calculating the index. Additionally, the series is time-dependent, and to capture this dependency; it is needed to add non-time dependencies to the dataset before calculating centroids.

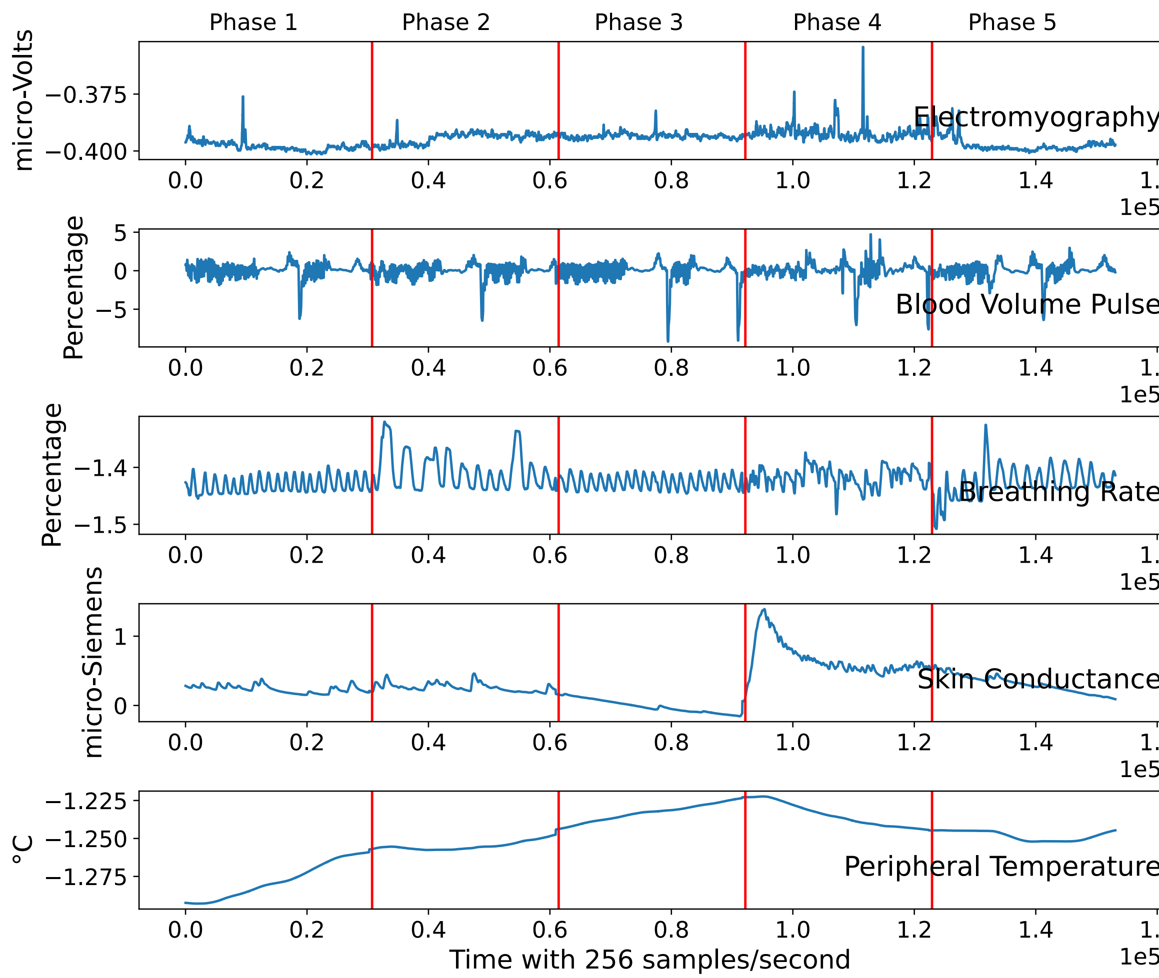


Figure 3.14: Line plot of the processed physiological signals of the study subject throughout the five mental stress phases.

3.2.7 Adding Features without Time-Dependency

A time series hold valuable information that varies continuously through time. A first-order difference can help to extract information in the temporal dependence of the series. Differencing is performed by subtracting the previous observation from the current instance. Equation 3.3 shows the basic principle to extract additional information by reducing trend and seasonality.

$$difference(t) = instance(t) - instance(t-1) \tag{3.3}$$

For this study, one order difference and an additional differencing in $t-2$, shown in Equation 3.4, was added to capture additional information for calculating distances between phases in n-dimensional space.

$$difference(t) = instance(t) - instance(t-2) \tag{3.4}$$

The resulting dataset consists of 15 attributes for each subject, two additional variables per physiological feature, as shown in Table 3.7, 3.8, and 3.9.

Table 3.7: Five instances of the 15 resulting features of one subject after preprocessing with the double differencing.

Time	El	BVP	BR	SC	PT
1.9922	-0.3900	0.4939	-1.4061	0.2153	-1.2483
1.9961	-0.3900	0.4939	-1.4062	0.2153	-1.2483
2.0000	-0.3900	0.4939	-1.4063	0.2153	-1.2483
2.0039	-0.3900	0.4939	-1.4063	0.2153	-1.2483
2.0078	-0.3900	0.4939	-1.4066	0.2153	-1.2483

Table 3.8: (continued)

Time	El_diff	BVP_diff	BR_diff	SC_diff	PT_diff
1.9922	0.0510	0.0000	-0.0010	0.0000	0.0010
1.9961	0.0000	0.0000	-0.0010	0.0000	0.0000
2.0000	0.0000	0.0000	-0.0010	0.0000	0.0000
2.0039	0.0000	0.0000	0.0000	0.0000	0.0000
2.0078	0.0000	0.0000	-0.0020	0.0000	0.0000

3.2.8 Resilience-to-Mental-Stress Index (RMSI) Proposal

Now that the information is complete and preprocessed, it can be used to quantify an alteration factor (AF). This work proposes three different approaches to obtain the RMSI and the AF values. The common ground of these approaches is that changes in the signals due to stress can be compared against a calibration stage.

Table 3.9: (continued)

Time	El_diff2	BVP_diff2	BR_diff2	SC_diff2	PT_diff2
1.9922	0.0510	0.0000	-0.0020	0.0000	0.0010
1.9961	0.0510	0.0000	-0.0020	0.0000	0.0010
2.0000	0.0000	0.0000	-0.0020	0.0000	0.0000
2.0039	0.0000	0.0000	-0.0010	0.0000	0.0000
2.0078	0.0000	0.0000	-0.0020	0.0000	0.0000

The calibration stage is unique to each subject; hence, all computed distances are relative to this starting point or baseline. The calibration stage is phase 1 of the stress test, where the subject is relaxed and resting for two minutes. We compute the distance between each phase to the calibration phase. In other words, we calculate the distance of phase one against all other phases; that is, $phase_1 \rightarrow phase_i$, where $i \in \{2, 3, 4, 5\}$.

Next, we explain our approach to compute distances between phases. The primary assumption is to consider a phase as a cluster of points (physiological measurements) close to each other. We then use three well-known techniques to calculate the distance among clusters: Euclidean distance, Mahalanobis distance, and Cluster Validity Index distance.

Euclidean Distance of PCA to Calculate Inter-phase Distances (ED-PCA)

PCA allows reducing high-dimensional data to a smaller number of dimensions while maintaining the information [129]. Because a covariance matrix is used to perform PCA, it eliminates multicollinearity between the attributes [43]. Therefore, euclidean distance can be used to calculate the distance between clusters.

The PCA was constructed for all subjects using the same preprocess and assumptions explained in Section 3.2.5. The explained variance ratio of all subjects is used to select the number of components to minimize information loss. Figure 3.15 illustrates the explained variance ratio of the components for all subjects. The mean average of the explained variance ratio with 15 components was 100 percent with a standard deviation of zero. With 14 components, the explained variance ratio slightly decreases to a mean value of 99.87 percent with a standard deviation of 0.11. In this study, to minimize the loss of information, we decided to use 15 principal components to compute the distances between clusters.

The relative displacement between phases is now compared to the calibration phase (phase1). It is crucial to notice that an n-dimensional (n = 15 in our case) space with 256 samples per second minimizes information loss when calculating relative distances between phases. Hence, the centroids are calculated using 30,720 observations per phase (a total of 153,600 observations per subject). This step's results include the distances between the centroid of the calibration phase and the subsequent psychophysiological mental stress test phases. A summary of the methodology of the method is shown in Figure 3.16.

Mahalanobis Distance to Calculate Inter-phase Distances (MD)

The second technique to measure cluster distance is the Mahalanobis distance [50]. Compared to the euclidean distance, the Mahalanobis distance can be calculated without standardizing

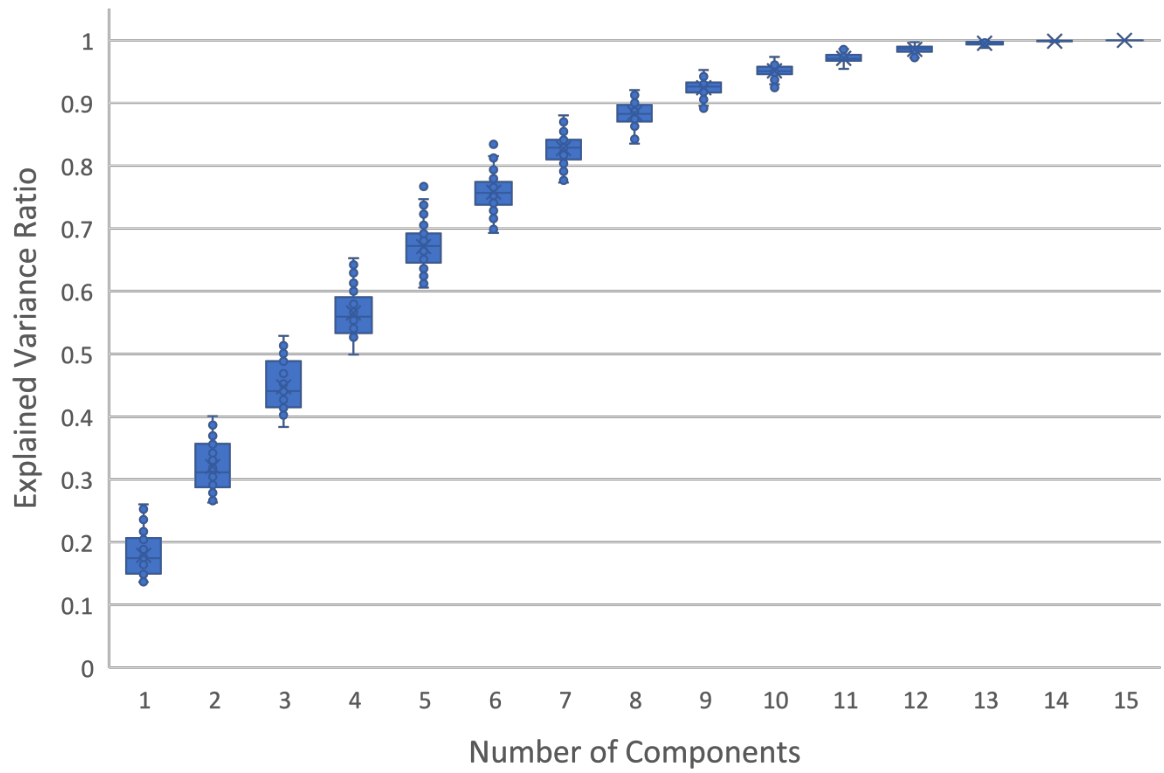


Figure 3.15: Boxplot of the explained variance ratio of the number of the principal components for the 15 features.

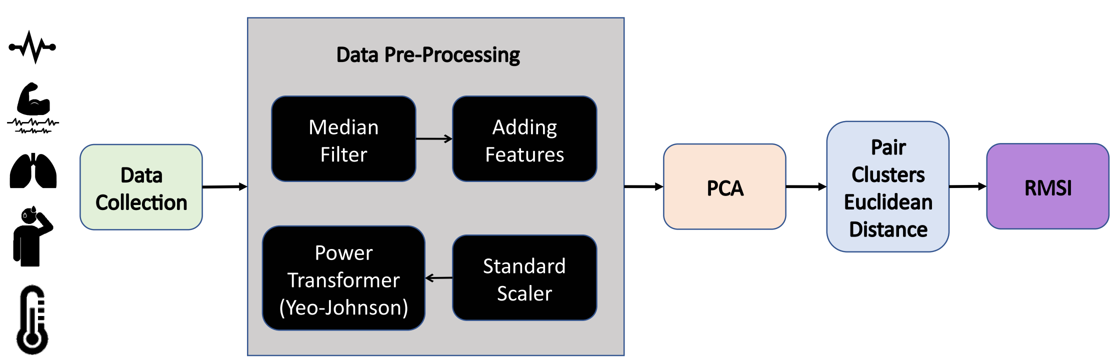


Figure 3.16: Flow diagram of the steps to calculate the Euclidean distances of principal components of the phases.

the data [130]. It can measure distances between points in multivariate space, even if their variables are correlated [50]. For this study, the distances are calculated using the Scipy python package.

The method's objective is to calculate the distance of the pair of phases: $phase_1 \rightarrow phase_i$, where $i \in \{2, 3, 4, 5\}$. The process to calculate the Mahalanobis distance is the following (illustrated in Figure 3.17):

1. Calculate the inverse of the covariance matrix with the instances that correspond to the pair of phases to calculate the distance (e.g., phase 1 and phase 2 instances).

2. The centroid of the baseline (phase 1) is calculated by the mean of each feature. Therefore, the resulting centroid will have a vector of 15 attributes.

3. Compute the Mahalanobis distance of each instance vector against the centroid vector of baseline.

4. Compute the mean of all calculated distances of the previous step to obtain the result.

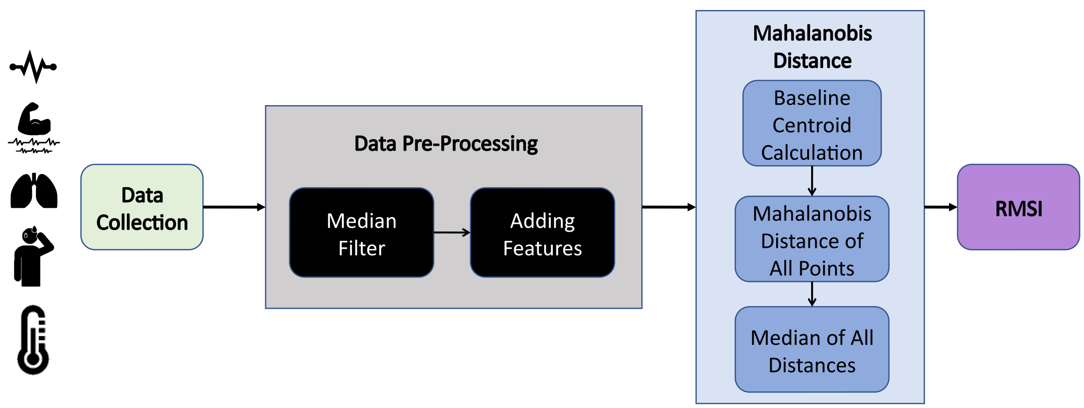


Figure 3.17: Flow diagram of the steps to calculate the Mahalanobis distances of the phases.

Cluster Validity Index Distance (CVID)

The last approach to measure cluster distance is employing the Silhouette index [44]. The separation metric is based on the nearest neighbor distance, and the Mahalanobis metric is used as the distance metric [131]. The Silhouette score is calculated with the Sklearn python package. The function returns the coefficient of the mean silhouette for all samples. It calculates the mean of the nearest cluster distance and the mean of the intra-cluster distance.

A Silhouette value of zero indicates that the clusters are overlapping. A negative value usually means that a sample was wrongly assigned and could belong to a different cluster [47]. A value close to one indicates that the clusters are well matched. Since the Silhouette coefficient measures cohesion and separation between clusters [44], all pairs of clusters' bahavior can be calculated.

The resulting dataset consists of the silhouette coefficient of each combination of pairs of the data. For this application, a value closer to one means that the compared pair of phases has fair cohesion and separation between each other, meaning that the subject has phases that do not overlap and the biofeedback signals are different. A summary of the steps is illustrated in Figure 3.18.

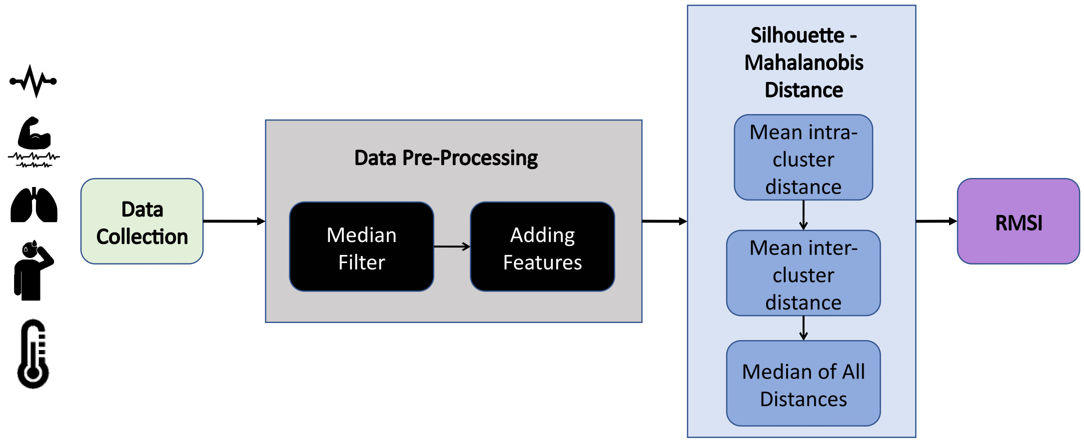


Figure 3.18: Flow diagram of the steps to calculate the cluster validity index distance of the phases.

Summary

Overall, the proposals to obtain the cluster distances are: The first approach captures the distance between clusters with the euclidean distance based on the Principal Components Analysis results. The second approach is to capture the distance between clusters with the Mahalanobis distance, including all features. Finally, the third approach involves capturing the subject's bahavior with a cluster validity index. The cluster validity index captures the cohesion and separation between clusters, allowing capturing the bahavior compared against phase one and the bahavior of all combinations of phase pairs between relaxation and recovery phases against stress phases.

This section explained the three proposed methodologies as a previous step before calculating the RMSI. The following chapter will explain how the RMSI and AF are obtained and the recommended RMSI methodology.

Chapter 4

Results and Discussion

This chapter discusses the proposal of an index that could measure the severity of a subject's resilience to mental stress in the face of analytical challenges and an alteration factor that could reflect the strength of mental stress. The principal objective of developing the RMSI is that the severity of an individual's mental stress resilience has essential information that could lead to an accurate prediction of depression. However, depression is not tested in this thesis but concentrates on creating the resilience index. The proposed index can capture the physiological data of an individual. After computing the cluster distances or the silhouette coefficient between pairs, we calculated the RMSI and the maximum AF values. The process to obtain these values is explained below.

4.1 Resilience-to-Mental-Stress-Index (RMSI) Calculation

The RMSI calculation based on cluster distances consists of measuring the cluster distance between the individual's phase four (the last stressor phase) and phase five (the last recovery phase). The RMSI is defined by the following equation (Equation 4.1).

$$RMSI = \frac{\Delta R}{\Delta S} \tag{4.1}$$

Where ΔR is the difference between the centroid of phase four and the centroid of phase five of the subject and ΔS is the maximum ΔR of the sample.

An RMSI value close to one implies that the subject's physiological responses could quickly recover the mental stress challenge. On the other hand, negative values indicate that the individual could not recover its physiological response after the last induced stressor. After computing the RMSI values for all subjects using the three distance metrics discussed previously, we compared the variance between them and proved no significant statistical difference.

To test for significance, we first analyzed the normality of the residuals with the Anderson-Darling test. The statistical test was conducted for the three methods (Euclidean with PCA, Mahalanobis, and Silhouette with Mahalanobis) and a dataset of 71 paired samples. For this analysis, we used the python package Autorank [132] and a significance level of 0.05. We rejected the null hypothesis (stating that the population is normally distributed) for the Mahalanobis ($p < 0.0005$) and Silhouette with Mahalanobis ($p < 0.0005$) methods as the p-values obtained were smaller than the significance level. Also, the Anderson-Darling test of the three methods' residuals, with a statistic of 3.391 ($p < 0.005$), reached the same conclusion. Given these results, we used the non-parametric Friedman test to assess the significance between the methods [133].

We failed to reject the null hypothesis (p=0.110) of the Friedman test, concluding that there was no difference in the populations' central tendency, Mahalanobis, Euclidean of principal components, and Silhouette with Mahalanobis. Hence, we concluded that there was no statistically significant difference between the median values of the populations. The results of the median absolute deviation (MAD), the median (MD), and the mean rank (MR) are shown in Table 4.1.

Table 4.1: Median, median absolute deviation, and mean rank results of the Friedman test of the three resilience to mental stress index methodologies results.

Metric	MD	MAD	MR
Euclidean of Principal Components	0.077 ± 0.083	0.183	2.183
Mahalanobis	0.174 ± 0.119	0.278	1.831
Silhouette with Mahalanobis	0.030 ± 0.083	0.183	2.183

4.2 Maximum Alteration Factor

We measured their maximum distance from phase 1 to the rest to determine the subjects' maximum stress level. We propose a new metric named the alteration factor (AF) to measure the subject's mental stress level. This metric is obtained by comparing the maximum stretch of the subject to the sample's maximum stretch. Equation 4.2 shows the calculation of this metric.

$$AF = \frac{SubMaxS - SamMinS}{\Delta SamST} \tag{4.2}$$

$SubMaxS$ is the subject's maximum distance compared to the subject's baseline. $SamMinS$ is the sample's minimum distance, and $\Delta SamST$ is the difference between the sample's maximum distance and the sample's minimum distance of the dataset.

An AF value close to one suggests that the subject's physiological response has suffered a higher alteration than most individuals in the sample. On the other hand, a value close to zero indicates that the variables did not suffer such a noticeable change concerning the sample.

4.3 Subject Example Analysis

This section illustrates the results of the metrics created by two subjects with opposite RMSI values. The example illustrates the RMSI based on the Euclidean distance of the PCA because the RMSI with this method allows a visual inspection of a subject's bahavior through the psychophysiological mental stress test and evaluates its corresponding RMSI.

Figure 4.1 depicts the five standardized physiological features of two subjects through the five phases of the psychophysiological mental stress test. The vertical red line separates each phase through time. Figure 4.1a shows a positive RMSI value for Subject 1, whereas Figure 4.1b Subject 2, with a negative RMSI.

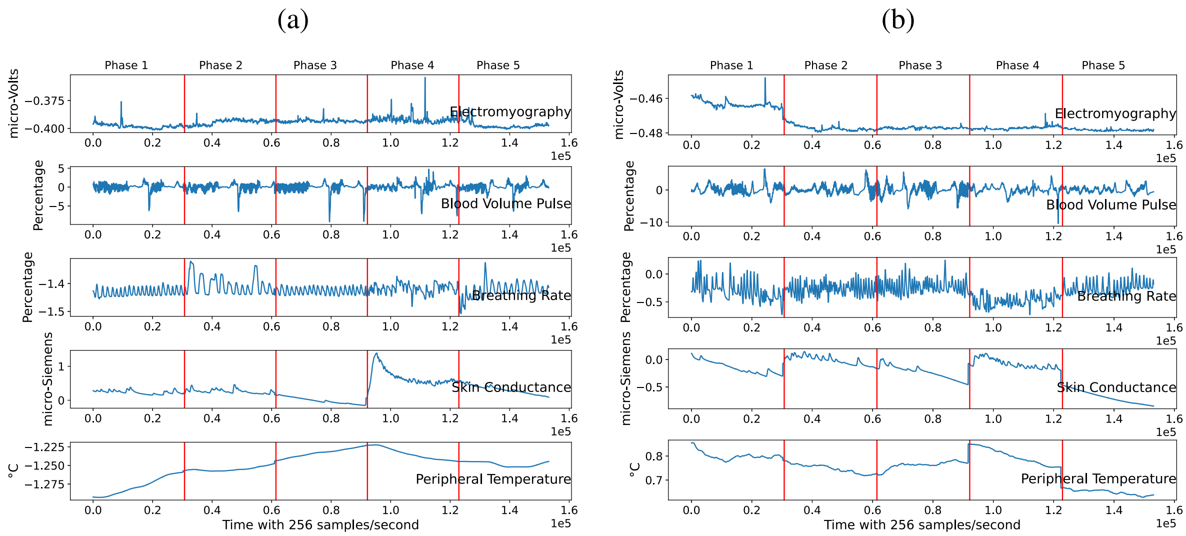


Figure 4.1: Standardized physiological features of two different subjects. Subject one depicted in (a) represents an individual with positive RMSI, and subject two, shown in (b), represents an individual with negative RMSI.

From the visual inspection of subject one (shown in Figure 4.1a), we can observe that his/her features suffered an alteration in phase two and phase four. These alterations indicate that the individual's physiological variables suffered an alteration in the face of a mental challenge. These variabilities show that the baseline is different from the phases that the mental stressor is induced. Compared to the previous phase (phase four), the last phase's signals are partially restored. We can conclude that the features' general bahavior is similar to the baseline, specifically, the electromyography, skin conductance, breathing rate, and blood volume pulse.

On the contrary, the subject analyzed in Figure 4.1b, representing an individual with a negative RMSI, has physiological features in the last phase (recovery phase) similar to the stress phases (phases two and four). This indicates that the subject's physiological features' were not restored.

As shown in sub-section 3.2.6, we can visually inspect individuals' physiological variables in a two-dimensional space using PCA. Variables from individuals analyzed in Figure 4.1 are depicted in two dimensions in Figure 4.2. Figure 4.2a shows that the cluster encompassing phase five is closer to the baseline for the first subject than it is to phase four, which translates into a positive RMSI. On the contrary, Figure 4.2b shows that phase five is farther from the baseline for the second subject than it is to phase four, which results in a negative RMSI.

Subject 1 has an RMSI = 1.0, and an AF = 0.64, whereas subject 2 has an RMSI = -0.77, and an AF = 0.41. RMSI equal to 1.0 means that subject 1 has the highest value

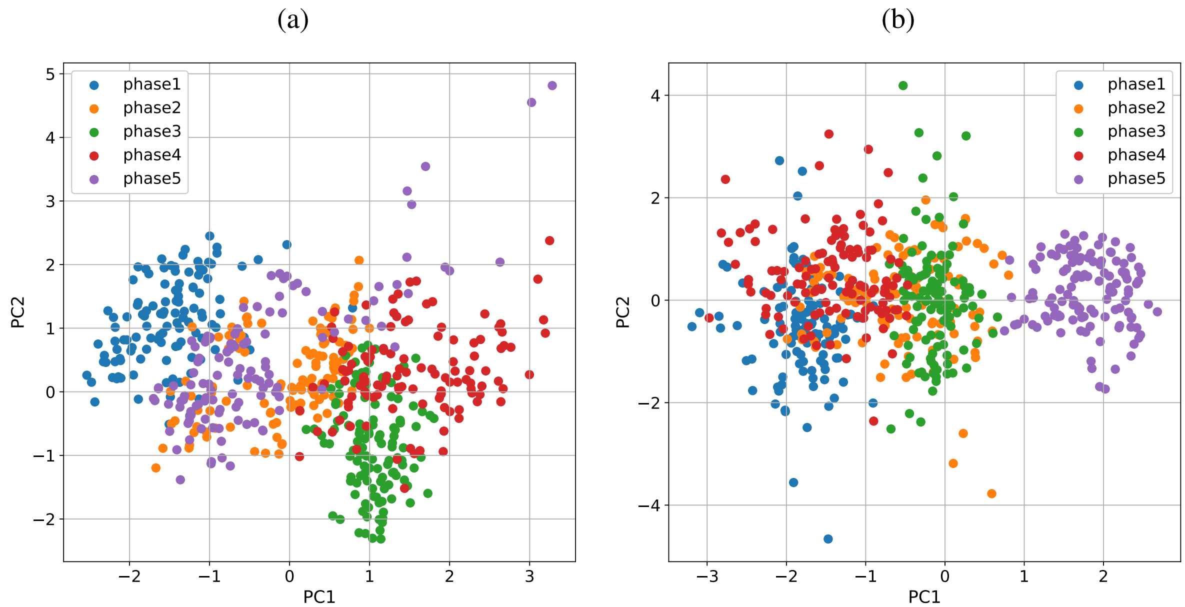


Figure 4.2: 2D-PCA representation of two different subjects. Subject one depicted in (a) is an individual with a positive RMSI, and subject two shown in (b) is an individual with negative RMSI.

in our population. In subject 2, phase five is far from the baseline, resulting in a negative RMSI. When analyzing the AF values, subject 1's phase four is farther from the baseline than the same phase for subject 2, meaning that, in the first case, the variables suffered a bigger change in their values; that is, they showed a higher alteration factor ($0.64 > 0.41$).

Our results are consistent with Healey and Picard's work [17]. The observed physiological features reflect the alteration that people suffer when induced by stressful situations; in this case, stress is induced by analytical challenges. Moreover, the study of Lu, Wang, and You [114] suggest that there are traits of mental stress and resilience traits to mental stress that can be measured based on physiological responses. Our findings suggest that the five measured physiological features have enough information to capture an individual response to mental stress and that the severity is defined by how much these variables are altered.

4.4 Results Validation

This section explores the validation of the RMSI. This study compares the results of the RMSI with two procedures. The first procedure is to evaluate the RMSI with the RESI-M tool. Even though there is no direct measure of resilience to mental stress, we could get some insights from comparing the index's ranking results against the RESI-M ranking. The second procedure evaluates each phase's raw physiological responses correlation to assess if the variables' response has positive or negative resilience to mental stress and its quantification.

4.4.1 RESI-M Tool

The RESI-M scores range between 43 and 172, where a higher score means higher traits of resilience. By calculating the correlation between the average rank of the RESI-M scores and the RMSI results, we could compare the degree of correlation between the proposed RMSI and the evaluation of the RESI-M scores. Table 4.2 shows the calculation of the RMSI and RESI-M scores of seven individuals with the three proposed methods. Additionally, it shows the average ranking for the seven individuals. From this point onwards, we refer to the RMSI methods with the following abbreviations: Euclidean distance of principal components as ED-PCA, Mahalanobis distance as MD, and cluster validity index distance as CVID.

Table 4.2: RMSI, RESI-M results, and an average ranking of seven individuals of the three proposed methods.

ID	**RESI-M**	**RMSI based on ED-PCA**	**RMSI based on MD**	**RMSI based on CVID**	**RESI-M Rank**	**ED-PCA Rank**	**MD Rank**	**CVID Rank**
1	142	-0.0133	0.2024	0.1203	3	1	3	5
2	159	0.0998	0.2200	-0.0230	6	2	4	2
3	162	0.1633	-0.1718	0.0081	7	3	2	3
4	106	0.2175	0.2420	0.0137	1	4	6	4
5	148	0.2853	0.2249	1.0000	4	5	5	7
6	152	0.4129	-0.5933	-0.1771	5	6	1	1
7	138	0.5728	0.2562	0.6317	2	7	7	6

To obtain the correlation coefficient, we first need to check for the variables' distribution. If the variables are normally distributed, then the Pearson coefficient is calculated; otherwise, we obtain the Spearman coefficient. The Anderson-Darling test for normality is used to prove if the variables have a normal distribution. From the results of the Anderson-Darling test with a significance level of 0.05, we can assume that the RESI-M (AD=0.455, p-value=0.182), ED-PCA (AD=0.155, p-value=0.919) and CVID (AD=0.658, p-value=0.049) populations are normal. However, we can assume that the MD (AD=1.057, p-value < 0.005) is not. Hence, we obtained the Pearson coefficient for ED-PCA and CVID and the Spearman coefficient for MD.

The resulting correlation coefficient for the average ranking of the methods with a significance level of 0.10 was ED-PCA of -0.1149 (p-value = 0.801), MD of -0.714 (p-value = 0.071), and CVID of -0.04 (p-value = 0.932). These results demonstrate a high negative linear relationship between the results of the RESI-M and the MD method. Moreover, a low negative linear relationship between RESI-M and ED-PCA and almost a null negative linear relationship with CVID and RESI-M.

A lower correlation value could indicate that the RMSI measures a subset of the resilience set's whole spectrum. It also follows our previous statement that identifies resilience to mental stress as a subset of the resilience set. This is the case of the ED-PCA method, where it was expected to have lower correlation values because there is no direct method to measure resilience to mental stress.

The limitation of this validation is the number of analyzed subjects of the RESI-M. Due to contact restrictions of the COVID-19 pandemic, the number of participants for the RESI-M test was low. Another limitation is the validation of the RESI-M tool. Even though it is validated in a Mexican sample, more subjects are needed to obtain statistically significant results. As we do not have a ground truth for this problem, a low correlation level was expected. That was the case with the ED-PCA and CVID methods. However, the MD showed higher traits of association to the RESI-M, which suggests that this method captures a little bit more of the RESI-M test.

4.4.2 Correlation between phase one and phase five

The second validation method involves capturing through the Spearman correlation if there is a monotonic relationship with its degree between phases one and five and capture if the behavior is resilient or non-resilient.

We calculated the Spearman coefficient of the seven test subjects for the five physiological features. Table 4.3 shows these results along with the mean of each subjects' physiological variables and its corresponding result of the RESI-M. The correlation coefficient was calculated with the preprocessed data of phase one and phase five. A sampling frequency of one every second was calculated with the average data of one second. In this case, we wanted to compare the baseline with the recovery phase to capture the relationship between both phases instead of the last stressor phase.

Table 4.3: Spearman correlation coefficient of the seven test subjects physiological variables with its corresponding Resilience in Mexicans test result.

ID	El	HR	RR	SC	PT	Mean	RESI-M
1	-0.45	0.25	-0.27	0.92	-0.99	-0.11	159
2	0.00	0.27	0.11	0.99	0.87	0.45	148
3	-0.03	-0.09	-0.02	0.88	0.94	0.34	138
4	-0.31	0.08	0.16	0.96	-0.35	0.11	152
5	-0.09	0.24	0.14	0.99	-0.82	0.09	106
6	-0.10	0.15	0.02	0.99	-1.00	0.01	142
7	-0.11	0.04	0.03	0.99	-1.00	-0.01	162

Next, we applied a Spearman correlation with the mean results of the seven subjects with its corresponding score of the RESI-M. We obtained a correlation of -0.50 (p-value = 0.253). These results concluded that we could not reject the null hypothesis of the Spearman coefficient, which states that there is no monotonic association between the mean results of the five physiological features and the RESI-M.

In conclusion, even though there is no monotonic relationship between the mean results of the five physiological variables and the RESI-M, we have a significant relationship between the MD method and the RESI-M, proving a direct association of the first validation method of this analysis and the RESI-M.

4.5 Discussion

Machine learning provides us tools to interpret and analyze various types of data. In this study, the non-supervised technique of PCA allowed the visualization of five clusters reflecting the behavior of the physiological features in the different mental stress phases. The visual inspection of the data behavior allowed us to measure cluster distances between the psychophysiological mental stress test phases.

Measuring stress has been widely studied. However, few studies consider the evaluation of resilience to stress in individuals using machine learning techniques. Recent studies have demonstrated the importance of resilience to stress for mental health. They are based on the idea that depending on the degree that individuals improve their resilience to stress, they will be more capable of coping with difficult situations and, as a result, be able to reduce their risk for depression. This study focuses on evaluating students' mental stress to develop a quantifiable metric that can be measured and managed. The benefits of having a metric for measuring resilience to mental stress are multiple. For instance, to the extent that the students can track their resilience to mental stress, they can improve their academic life. These findings contribute to further analysis of depression detection. With this study, we aim to give academic institutions the tools to identify and support students with higher risks of depression and, most importantly, prevent any risk to students' physical integrity.

In the face of analytical challenges, the development of a Resilience to Mental Stress Index was done using modern techniques of machine learning. Based on the Friedman results, we can infer that the most appropriate method compared to the proposed three to calculate the RMSI is the Mahalanobis distance as there is no significant difference between the other two methods and its high monotonic association with the RESI-M. Even though the method requires high computational expense, it is recommended for its relationship with the RESI-M.

Physiological features such as electromyography, breathing rate, blood volume, skin conductivity, and peripheral temperature can reflect stress conditions. These alterations can be measured with the Mahalanobis distance or Euclidean distance in a space without multicollinearity. These n-dimensional space clusters reflect a subject's general behavior to stress compared to a baseline.

This study's limitations are based on the assumption that the population has healthy biological parameters, that an individual cannot relax more than he was at the beginning of the psychophysiological stress test, and that the relative comparison between samples depends on the sample parameters.

In addition to the experiments presented in this thesis, we conducted a different approach to construct the RMSI with supervised models. We tried to capture the alterations of the physiological variables with evaluation metrics. Although they give insight for further analysis, these experiments were not representative because the severity of the alteration of the physiological variable could not be correctly quantified.

The study can be replicated using a similar biofeedback device and psychophysiological stress test to collect the specified variables. In comparison to related work, this study proposed three robust RMSI's independent of socio-demographic variables, as they are calculated based on relative distances of each subject and independent of written tests' bias.

Chapter 5

Conclusion

5.1 Conclusion

The features extracted in this work are recommended when evaluating psychological stress and resilience to mental stress. As a result of the proposed techniques, resilience to mental stress can be quantified relative to a sample.

The research questions that guided this thesis are answered below:

1. Which methodology is the most appropriate to build resilience to the mental stress index and an alteration factor?

MD is the recommended methodology to calculate the RMSI. This answer is based on the results that the three methods' central tendency does not significantly differ. Therefore, we selected the MD due to its high monotonic association with the RESI-M score. Additionally, we recommended the construction of the 2-dimension principal components for visual inspection, as it provides a resourceful insight into the general behavior of the five physiological features through the psychophysiological mental stress protocol.

2. Can the (dynamic) time-series data be transformed into static data to include in a classification problem without vectorizing or training ensemble models?

Dynamic time-series data can be transformed into static information to train conventional machine learning models (e.g., Random Forest, Logistic Regression) that could help predict depression. The physiological features have essential information that suffers changes when exposed to mental stress. These alterations that change through time can be quantified, and the calculations converge in a static metric that ranges from -1 to 1 for each individual. This data represents an index that is not dependent on time. However, this method is specific to our study and requires background theoretical knowledge of the topic.

3. Could the resilience to mental stress index and the alteration factor be compared among individuals?

The RMSI and AF can be compared among individuals because they compare the relative distance of clusters between their baseline values. Hence, scaling all subjects' physiological features will help us compare the distances between the psychophysiological mental stress test phases. It does not require prior knowledge of average population parameters or equipment calibration because the methodology only compares alterations of the variables.

4. Can the resilience to mental stress index and the alteration factor be independent of socio-demographic variables?

The RMSI and AF are independent of socio-demographic variables. Because each subject's baseline is different, the method calculates all distances relative to the baseline. The only parameter needed to compare a group is the maximum distance of the sample to assess an individual's severity among the group. Nevertheless, the RMSI will indicate if the subject has positive or negative resilience independent of all other group information.

5. Does the resilience to mental stress index is correlated with the resilience in Mexicans (RESI-M) psychological tool?

The RMSI is highly correlated to the RESI-M. There is no direct tool that measures resilience to mental stress. However, we consider RESI-M to measure the whole spectrum of resilience for its validation in a Mexican sample. We expected either a low or medium relationship between the RESI-M and the RMSI. The ED-PCA and CVID achieve a low correlation with the RESI-M with no statistical significance, and on the other hand, the MD method got a higher correlation. A higher correlation value between the RESI-M and RMSI could mean that the RMSI explains more of the RESI-M test. However, as we do not have ground-truth results to compare the RMSI, we can not achieve definitive conclusions. Nevertheless, we can conclude that the measured physiological features react to mental stress. Furthermore, in a controlled environment is possible to measure the severity of these variables' alteration and compare them among individuals.

5.2 Final Comments and Future Work

An RMSI and AF proposal serves as a foundation for continuing research within resilience to mental stress and psychiatric studies. Many applications and work could derive from this study. Some of the identified opportunities and future research work is:

- The proposed RMSI could be measured after and before clinical treatments to measure its performance and further validate our results.
- Embedded sensors in smartwatches could obtain additional features that capture different behaviors to obtain the RMSI. The proposed methodology could be replicated with data from these non-invasive devices instead of utilizing a biofeedback device such as the ProComp5 Infiniti.
- Construct the RMSI with data of different days from one individual. This will help to construct the metric relative to itself instead of a sample. This type of analysis could be helpful while monitoring the progress of resilience to the mental stress of psychological treatment.
- Various tests could be applied to measure the severity of various physiological responses, such as depression. Moreover, the RMSI and AF could be crucial information to predict depression, as stress is directly related to the onset of depression symptoms.
- Different scenarios can be prepared to test additional responses other than mental stress in the face of analytical challenges. These tests need to be designed to measure the whole spectrum of resilience and not just mental stress.

- A neural network could be trained to predict the RESI-M scores based on the five physiological features.

We hope that this work can be replicated to study resilience to stress profoundly and that the proposed methodology could be a solid baseline that can contribute to the study of an individual's resilience.

Appendix A

App Images

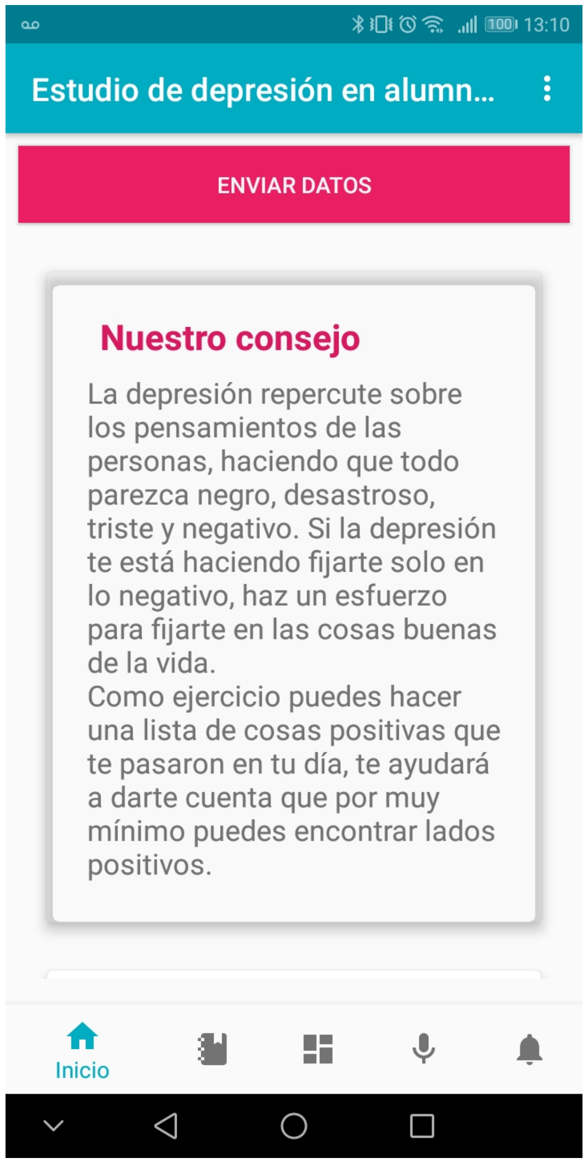


Figure A.1: Home page of the application.

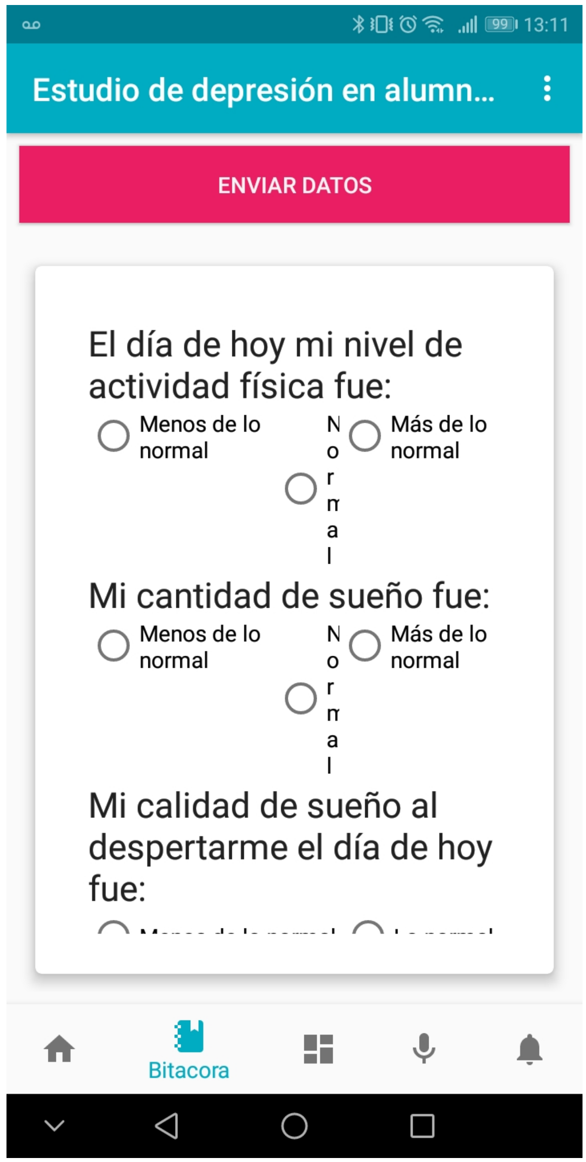


Figure A.2: Daily log window of application.

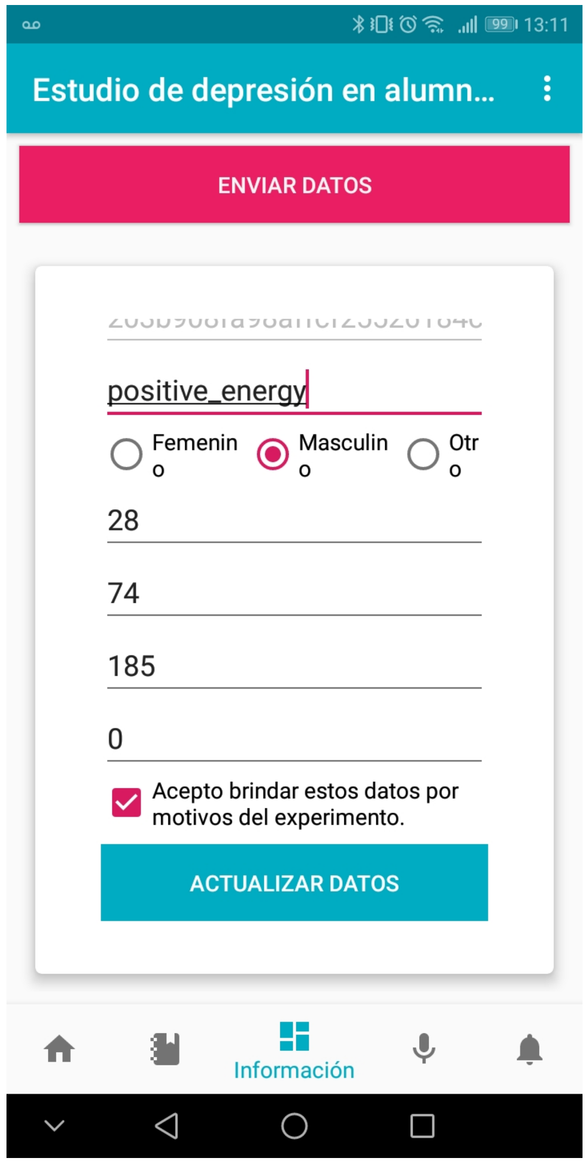


Figure A.3: New user registry window of the application.

Figure A.4: Voice recording menu of the application.

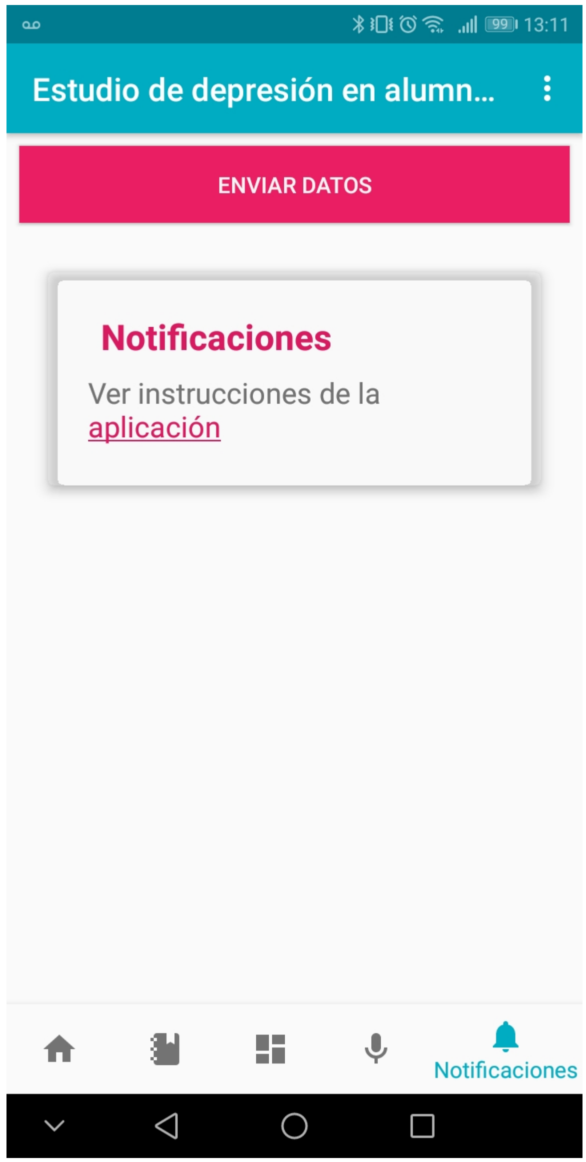


Figure A.5: Latest updates and notification menu of the application.

Appendix B

Daily Log

INSTRUCCIONES: Favor de llenar el cuestionario antes de terminar su día. En los siguientes incisos por favor marque uno de cada grupo.

1. El día de hoy mi tiempo de ejercicio fue:
   (a) Menos de lo normal
   (b) Normal
   (c) Más de lo normal
2. Mi calidad de sueño al despertarme el día de hoy fue:
   (a) Menos de lo normal
   (b) Lo normal
3. Mi cantidad de sueño fue:
   (a) Menos de lo normal
   (b) Lo normal
   (c) Más de lo normal
4. El día de hoy me sentí:
   (a) Deprimido
   (b) Tranquilo
   (c) Feliz

   En caso de que la respuesta al inciso anterior haya sido "deprimido" continúe con los siguientes incisos, en caso contrario finalice y guarde sus respuestas.
5. Favor de indicar su nivel máximo de depresión durante el día:
   (a) Leve
   (b) Moderado

    (c) Severo

6. Favor de indicar el momento del día en que se sintió más deprimido:
    (a) En la mañana
    (b) En la tarde
    (c) En la noche
7. Favor de indicar la duración del episodio de su mayor depresión:
    (a) Menos de una hora
    (b) De una a 2 horas
    (c) Más de 2 horas
8. ¿Cuál fue el tipo de respuesta emocional que tuvo en su nivel de máxima depresión? Puede seleccionar más de uno.
    (a) Ansiedad
    (b) Estrés
    (c) Ira
    (d) Tristeza
    (e) Miedo

Appendix C

Psychophysiological Mental Stress Test Protocol

Motivos para suspender una prueba antes de iniciar:

- Que la persona evaluada esté enferma.
- Que la mujer evaluada esté en su periodo menstrual.
- Que la persona haya consumida café, cigarro o bebida energizante dos horas previas a la prueba.
- Que la persona haya dormido menos de lo que regularmente duerme (Si normalmente duerme 6, el día anterior deberá haber dormido 6).

Motivos para suspender una prueba durante la evaluación:

- Que la persona se niegue a seguir con la actividad.
- Que la persona sufra una crisis de ansiedad.
- Que la persona abandone la tarea estresante por más de 20 segundos.

Protocolo

- Fases: 5
- Duración por fase: 2 minutos
- Duración total: 10 minutos
- Protocolo de medición psicofisiológica 5 fases - 5 canales - Dr. Figueroa

Sensores por utilizar

- Volymen del Pulso Sanguíneo
  - Colocar en el pulgar

- Frecuencia Respiratoria
  - Colocar a la altura del pecho
- Conductancia de la piel (requiere gel conductor)
  - Colocar en índice y anular-segunda falange
- Temperatura Periférica
  - Colocar en la punta del dedo índice
  - Cubrir con el velcro la punta del sensor
- Electromiografia
  - Colocar en el trapecio

C.1 Instrucciones iniciales

(Pasa la persona y se sienta en una silla con la postura derecha, la espalda derecha, los dos pies tocando el suelo y las manos sobre las piernas).

Evaluador: Hola, el día de hoy haremos tu perfil psicofisiológico, en el cual esteremos midiendo tus respuestas fisiológicas ante ciertas actividades.

Evaluador: Para iniciar, vamos a medir tus niveles basales para tener un parámetro para iniciar con el registro.

C.2 Registro Linea Base 1

(Se conecta a la persona a los sensores y se comienza a registrar en el software la línea base de 2 minutos; la persona no se debe mover ni hablar, debe estar con los ojos cerrados)

Evaluador: Ahora vamos a medir tu estado basal de actividad fisiológica, te pido que una vez que te conecte, deberás cerrar los ojos, evitar moverte esta primera parte será sin hablar, hasta que se te indique, nosotros te diremos la siguiente actividad.

Evaluador: Permanece en silencio con los ojos cerrados, en un momento te daré más instrucciones.

(Comenzar registro)

C.3 Registro Descanso 1

(Se da inicio a la segunda etapa de 2 minutos sin dar ninguna indicación extra, sin responder preguntas del sujeto)

(Comenzar registro)

C.4 Registro del Estresor 1

Evaluador: Ahora vamos a realizar unas actividades, vamos a estar trabajando con calculo mental y lenguaje. Vamos a realizar operaciones mentales, por ejemplo, si yo te dijera 2x2 tú me deberías responder 4. Recuerda que solo me debes de dar el resultado lo más rápido posible. Por ejemplo, si yo te digo 3x3, tú me dices…

Evaluador: También vamos a trabajar con el abecedario, cada vez que yo te diga A, tu vas a tener que comenzar a decirme una palabra en el orden del abecedario, por ejemplo, A… Árbol, becerro, casa, dentista, etc etc. Si te equivocas o tardas mucho, deberás de comenzar de nuevo, pero sin repetir las palabras. Por ejemplo, si yo te digo A, tú me dices…

Evaluador: También vamos a trabajar con un ejercicio de silaba final, en este te daremos una palabra y deberás de decir otra con la última silaba de la palabra, por ejemplo, lupa ... palo o brújula… Latinoamérica, deberás de seguir hasta que te digamos otra palabra, recuerda que no se puede repetir una palabra que hayas dicho. Por ejemplo, si yo te digo comida…

Evaluador: En cualquier momento podemos movernos entre actividades, cada vez que yo te diga cambio.

(Comenzar registro)

Actividades: Calculo mental:

- 3 x 4
- 5 x 7
- 8 x 7
- 30 - 8
  - - 9
  - + 5
  - - 3
- 2 x 2
  - x 2
  - x 2
  - - 9
  - + 3
- 57 - 9
  - - 3
  - + 5
  - - 7
- 5 x 4
  - - 7

  - + 9
  - - 4

Abecedario:

- A...
- A...
- A...
- Las veces que sea necesario

Silaba Final:

- Corona
- Billete
- Pájaro
- Mochila
- Ganado
- Tienda
- Jardín
- Raqueta
- Oveja
- Fuego

Notas al evaluador:

- Ritmo constante
- Pedir respuestas rápidas
- Respuestas correctas
- Ojos cerrados
- Observar que no mueva partes del cuerpo, no se ría, en caso de hacerlo, pedirle firmemente que no lo haga y solo se limite a contestar lo solicitado

C.5 Registro Descanso 2

Evaluador: De igual forma vamos a estar trabajando con calculo mental y lenguaje. Vamos a realizar operaciones mentales,

Evaluador: También vamos a trabajar con un ejercicio de deletreo, yo te voy a decir una palabra y lo más rápido posible deberás de deletrearla sin repetirme la palabra. Por ejemplo, si yo te digo CASA, tu me deberías de decir … (C – A- S- A)

Evaluador: De igual forma, vamos a trabajar con series de números, yo te voy a decir números y tú deberás de responderme con los mismos números, pero ordenados de menor a mayor. Por ejemplo, si te digo 2-1-3 tú deberías de decirme … (1,2,3)

Evaluador: Recuerda que podemos cambiar entre actividades en cualquier momento, debes de dar respuestas rápidas y correctas, si no deberemos empezar de nuevo.

(Comenzar registro)

Actividades: Calculo mental:

- 9 x 6
  - - 5
  - + 8
- 7 x 4
  - - 6
  - + 12
- 7 x 8
  - - 7
  - - 9
- 70 x 3
- 9 x 6
  - + 13
  - - 7
  - + 9
- 64 - 27

Deletrear:

- Senil
- Conciencia
- Trasplante
- Avestruz

- Maguey
- Decepcionado
- Avergonzado
- Capturar
- Onomatopeya
- Imprescindible

Memoria reciente:

- 3 - 5 - 2
- 4 - 7 - 1
- 8 - 9 - 2
- 4 - 3 - 5
- 9 - 7 - 8
- 6 - 9 - 2
- 2 - 5 - 3
- 8 - 7 - 9 - 4
- 5 - 4 - 6 - 1
- 8 - 2 - 7 - 9

Notas al evaluador:

- Ritmo constante
- Pedir respuestas rápidas
- Respuestas correctas
- Ojos cerrados
- Observar que no mueva partes del cuerpo, no se ría, en caso de hacerlo, pedirle firmemente que no lo haga y solo se limite a contestar lo solicitado

C.6 Linea Base 2

Evaluador: Esta fase ha terminado, continua con los ojos cerrados hasta que te lo indique.

(Comenzar registro)

(Termina protocolo)

Evaluador: Hemos terminado el protocolo, puedes abrir los ojos.

Evaluador: Agradecemos tu participación, estos datos serán de gran valor para la investigación. Estaremos analizando los resultados para poder ver tus resultados de tu perfil psicofisiológico.

Evaluador: Te pedimos qué si conoces a alguien que asistirá a su evaluación, no comentes que estuvimos trabajando, esto con el objetivo que puedan tener un perfil sin sesgo. Si te llegan a preguntar solo contesta cosas generales y no comentes ningún aspecto especifico o de las actividades realizadas.

Evaluador: ¿Del uno al diez que tanto estrés sientes?

(En caso de decir seis o más, hacer un ejercicio de respiración para llevar a la persona reducir su estrés)

Bibliography

[1] WHO. Depression and other common mental disorders: global health estimates. *World Health Organization*, pages 1–24, 2018.

[2] Maria Medina-Mora, Gabriela Borges, Corina Benjet, Carmen Lara, E Rojas, Clara Fleiz, J Zambrano, Jorge Villatoro Velazquez, J Blanco, and Sergio Aguilar-Gaxiola. Estudio de los trastornos mentales en México: resultados de la Encuesta Mundial de Salud Mental. *Epidemiología de la salud mental en Latinoamérica y el Caribe*, pages 79–89, jan 2009.

[3] Jane M. Gunn, Victoria J. Palmer, Christopher F. Dowrick, Helen E. Herrman, Frances E. Griffiths, Renata Kokanovic, Grant A. Blashki, Kelsey L. Hegarty, Caroline L. Johnson, Maria Potiriadis, and Carl R. May. Embedding effective depression care: Using theory for primary care organisational and systems change. *Implementation Science*, 5(1):1–16, 2010.

[4] Oliver Oyama PhD and Nancy A Piotrowski PhD. Depression, 2019.

[5] Viviane Brachet-Márquez. Estudios Sociológicos. *Estudios Sociológicos*, 13(39):645–653, 1995.

[6] Corina Benjet, Guilherme Borges, Ma Elena Medina-Mora, Clara Fleiz-Bautista, and Joaquín Zambrano-Ruiz. La depresión con inicio temprano: prevalencia, curso natural y latencia para buscar tratamiento. *Salud pública de México*, 46:417–424, 2004.

[7] Holger Strulik. An economic theory of depression and its impact on health behavior and longevity. *Journal of Economic Behavior and Organization*, 158:269–287, 2019.

[8] Lynne M Jackson, S Mark Pancer, Michael W Pratt, and Bruce E Hunsberger. Great Expectations: The Relation Between Expectancies and Adjustment During the Transition to University 1. *Journal of Applied Social Psychology*, 30(10):2100–2125, 2000.

[9] Kristina L. Bradley, Darcy A. Santor, and Rylee Oram. A Feasibility Trial of a Novel Approach to Depression Prevention: Targeting Proximal Risk Factors and Application of a Model of Health-Behaviour Change. *Canadian Journal of Community Mental Health*, 34(January):47–61, 2016.

[10] Ranna Parekh. What is Depression?, 2017.

[11] World Health Organization. *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelins.* WHO, 1992.

[12] Norifumi Tsuno, Alain Besset, and Karen Ritchie. Sleep and Depression., 2005.

[13] Marina Marcus, M Taghi Yasamy, Mark van van Ommeren, Dan Chisholm, and Shekhar Saxena. Depression: A global public health concern. 2012.

[14] Statista. Global smartwatch unit sales 2014-2018, 2017.

[15] Sheldon Cohen, Denise Janicki-Deverts, and Gregory E Miller. Psychological stress and disease. *Jama*, 298(14):1685–1687, 2007.

[16] Kenneth S Kendler, Laura M Karkowski, and Carol A Prescott. Causal relationship between stressful life events and the onset of major depression. *American Journal of Psychiatry*, 156(6):837–841, 1999.

[17] Jennifer A Healey and Rosalind W Picard. Detecting stress during real-world driving tasks using physiological sensors. *IEEE Transactions on intelligent transportation systems*, 6(2):156–166, 2005.

[18] Hong Lu, Denise Frauendorfer, Mashfiqui Rabbi, Marianne Schmid Mast, Gokul T Chittaranjan, Andrew T Campbell, Daniel Gatica-Perez, and Tanzeem Choudhury. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones. In *Proceedings of the 2012 ACM conference on ubiquitous computing*, pages 351–360, 2012.

[19] David Sun, Pablo Paredes, and John Canny. MouStress: detecting stress from mouse motion. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 61–70, 2014.

[20] Martijn Haak, Steven Bos, Sacha Panic, and L J M Rothkrantz. Detecting stress using eye blinks and brain activity from EEG signals. *Proceeding of the 1st driver car interaction and interface (DCII 2008)*, pages 35–60, 2009.

[21] Nigar Khawaja and Helen Stallman. Understanding the coping strategies of international students: A qualitative approach. *Journal of Psychologists and Counsellors in Schools*, 21(2):203–224, 2011.

[22] Simon Cassidy. Resilience building in students: the role of academic self-efficacy. *Frontiers in psychology*, 6:1781, 2015.

[23] Sam Goldstein and Robert B Brooks. Why study resilience? In *Handbook of resilience in children*, pages 3–14. Springer, 2013.

[24] Gail Kinman and Louise Grant. Exploring stress resilience in trainee social workers: The role of emotional and social competencies. *The British Journal of Social Work*, 41(2):261–275, 2010.

[25] Catherine Neubauer, Joshua Woolley, Peter Khooshabeh, and Stefan Scherer. Getting to know you: a multimodal investigation of team behavior and resilience to stress. In *Proceedings of the 18th ACM International Conference on Multimodal Interaction*, pages 193–200, 2016.

[26] Kathryn M Connor and Jonathan R T Davidson. Development of a new resilience scale: The Connor-Davidson resilience scale (CD-RISC). *Depression and anxiety*, 18(2):76–82, 2003.

[27] Odin Hjemdal, Oddgeir Friborg, Stéphanie Braun, Chantal Kempenaers, Paul Linkowski, and Pierre Fossion. The Resilience Scale for Adults: Construct validity and measurement in a Belgian sample. *International Journal of Testing*, 11(1):53–70, 2011.

[28] Joaquina Palomar Lever and Norma E Gómez Valdez. Desarrollo de una escala de medición de la resiliencia con mexicanos (RESI-M). *Interdisciplinaria*, 27(1):7–22, 2010.

[29] Jonghwa Kim and Elisabeth André. Emotion recognition based on physiological changes in music listening. *IEEE transactions on pattern analysis and machine intelligence*, 30(12):2067–2083, 2008.

[30] C Godin, F Prost-Boucle, A Campagne, S Charbonnier, S Bonnet, and A Vidal. Selection of the most relevant physiological features for classifying emotion. *Emotion*, 40:20, 2015.

[31] Lin Shu, Jinyan Xie, Mingyue Yang, Ziyi Li, Zhenqi Li, Dan Liao, Xiangmin Xu, and Xinyi Yang. A review of emotion recognition using physiological signals. *Sensors*, 18(7):2074, 2018.

[32] Steven M Southwick, Meena Vythilingam, and Dennis S Charney. The psychobiology of depression and resilience to stress: implications for prevention and treatment. *Annu. Rev. Clin. Psychol.*, 1:255–291, 2005.

[33] Steven M Southwick and Dennis S Charney. The science of resilience: implications for the prevention and treatment of depression. *Science*, 338(6103):79–82, 2012.

[34] Anna Leontjeva and Ilya Kuzovkin. Combining static and dynamic features for multivariate sequence classification. In *2016 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, pages 21–30. IEEE, 2016.

[35] Lior Rokach. Ensemble-based classifiers. *Artificial intelligence review*, 33(1):1–39, 2010.

[36] Sinan Ozdemir. *Principles of Data Science*. Packt Publishing, Birmingham, 2016.

[37] Randa Tantawi PhD. Machine learning, 2018.

[38] Vijay K Vemuri. The Hundred-Page Machine Learning Book: by Andriy Burkov, Quebec City, Canada, 2019, 160 pp., 49.99(Hardcover); 29.00 (paperback); 25.43(KindleEdition),(Alternatively,canpurchaseatleanpub.comataminimumpriceof 20.00), ISBN 978-1999579517, 2020.

[39] John C Munson and Taghi M Khoshgoftaar. The dimensionality of program complexity. In *Proceedings of the 11th international conference on Software engineering*, pages 245–253, 1989.

[40] Hervé Abdi and Lynne J Williams. Principal component analysis. *Wiley interdisciplinary reviews: computational statistics*, 2(4):433–459, 2010.

[41] Ian T Jolliffe and Jorge Cadima. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 374(2065):20150202, 2016.

[42] Goeffrey J McLachlan. Mahalanobis distance. *Resonance*, 4(6):20–26, 1999.

[43] Maizah Hura Ahmad, Robiah Adnan, and Norliza Adnan. A comparative study on some methods for handling multicollinearity problems. *MATEMATIKA: Malaysian Journal of Industrial and Applied Mathematics*, 22:109–119, 2006.

[44] Olatz Arbelaitz, Ibai Gurrutxaga, Javier Muguerza, Jesús M Pérez, and Inigo Perona. An extensive comparative study of cluster validity indices. *Pattern Recognition*, 46(1):243–256, 2013.

[45] Deepak Sinwar and Rahul Kaushik. Study of Euclidean and Manhattan distance metrics using simple k-means clustering. *Int. J. Res. Appl. Sci. Eng. Technol*, 2(5):270–274, 2014.

[46] Minho Kim and R S Ramakrishna. New indices for cluster validity assessment. *Pattern Recognition Letters*, 26(15):2353–2363, 2005.

[47] Peter J Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[48] Jon R Kettenring. The practice of cluster analysis. *Journal of classification*, 23(1):3–30, 2006.

[49] M V Khachumov. Distances, metrics and cluster analysis. *Scientific and Technical Information Processing*, 39(6):310–316, 2012.

[50] Roy De Maesschalck, Delphine Jouan-Rimbaud, and Désiré L Massart. The mahalanobis distance. *Chemometrics and intelligent laboratory systems*, 50(1):1–18, 2000.

[51] Leo Liberti, Carlile Lavor, Nelson Maculan, and Antonio Mucherino. Euclidean distance geometry and applications. *SIAM review*, 56(1):3–69, 2014.

[52] Brooke E Muller and Bradley T Erford. Choosing assessment instruments for depression outcome research with school-age youth. *Journal of Counseling & Development*, 90(2):208–220, 2012.

[53] Max Hamilton. Rating depressive patients. *The Journal of clinical psychiatry*, 1980.

[54] Ivo Aben, Frans Verhey, Richel Lousberg, Jan Lodder, and Adriaan Honig. Validity of the Beck Depression Inventory, Hospital Anxiety and Depression Scale, SCL-90, and Hamilton Depression Rating Scale as screening instruments for depression in stroke patients. *Psychosomatics*, 43(5):386–393, sep 2002.

[55] Paul Richter, Joachim Werner, Andrés Heerlein, Alfred Kraus, and Heinrich Sauer. On the validity of the Beck Depression Inventory. A review. *Psychopathology*, 31(3):160–168, 1998.

[56] Aaron T. Beck, Robert A. Steer, and Margery G. Carbin. Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. *Clinical Psychology Review*, 8(1):77–100, jan 1988.

[57] Yuan Pang Wang and Clarice Gorenstein. Psychometric properties of the Beck Depression Inventory-II: A comprehensive review. *Revista Brasileira de Psiquiatria*, 35(4):416–431, 2013.

[58] Javaid I Sheikh and Jerome A Yesavage. Geriatric Depression Scale (GDS): recent evidence and development of a shorter version. *Clinical Gerontologist: The Journal of Aging and Mental Health*, 1986.

[59] Tyler Biscontini. Geriatric Depression Scale (GDS), 2017.

[60] William W K Zung. Zung self-rating depression scale and depression status inventory. In *Assessment of depression*, pages 221–231. Springer, 1986.

[61] Shinji Sakamoto, Nobuhiko Kijima, Atsuko Tomoda, and Masahiko Kambara. Factor structures of the Zung Self-Rating Depression Scale (SDS) for undergraduates, 1998.

[62] Debra L Murphy. Personality rating scales, 2018.

[63] Raymond B Cattell. The description of personality: Basic traits resolved into clusters. *The journal of abnormal and social psychology*, 38(4):476, 1943.

[64] Ernest C Tupes and Raymond E Christal. Recurrent personality factors based on trait ratings. *Journal of personality*, 60(2):225–251, 1992.

[65] Warren T Norman. Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, 66(6):574, 1963.

[66] Murray R. Barrick and Michael K. Mount. The Big Five Personality Dimensionad and Job Performance: A Meta-Analysis. *Personnel Psychology*, 44(1):1–26, 1991.

[67] Eric H Kessler. *Encyclopedia of Management Theory*. SAGE Publications, Inc, Thousand Oaks, California, 2013.

[68] Krzysztof Pancerz, Vitaly Levashenko, Elena Zaitseva, and Jerzy Gomulla. Experiments with classification of MMPI profiles using fuzzy decision trees. In *Proceedings of the 2018 Federated Conference on Computer Science and Information Systems, FedCSIS 2018*, volume 15, pages 125–128, 2018.

[69] James N Butcher and Carolyn L Williams. Personality assessment with the MMPI-2: Historical roots, international adaptations, and current challenges. *Applied Psychology: Health and Well-Being*, 1(1):105–135, 2009.

[70] David W Brokaw. Minnesota Multiphasic Personality Inventory (MMPI), 2018.

[71] Yossef S Ben-Porath. *Interpreting the MMPI-2-RF*. U of Minnesota Press, 2012.

[72] Harrison G Gough. The California psychological inventory. *Testing and assessment in counseling practice*, 45:71, 2000.

[73] Thomas E DeWolfe. California Psychological Inventory (CPI), 2018.

[74] Richard Rogers, Johnny Flores, Karen Ustad, and Kenneth W Sewell. Initial validation of the personality assessment inventory–Spanish version with clients from Mexican American communities. *Journal of Personality Assessment*, 64(2):340–348, 1995.

[75] Leslie C Morey. The Personality Assessment Inventory. *The Encyclopedia of Clinical Psychology*, 2014.

[76] Oddgeir Friborg, Odin Hjemdal, Jan H Rosenvinge, and Monica Martinussen. A new rating scale for adult resilience: what are the central protective resources behind healthy adjustment? *International journal of methods in psychiatric research*, 12(2):65–76, 2003.

[77] Gail M Wagnild and Heather M Young. Development and psychometric. *Journal of nursing measurement*, 1(2):165–17847, 1993.

[78] Joaquina Palomar Lever and Norma E Gómez Valdez. Desarrollo de una escala de medición de la resiliencia con mexicanos (RESI-M). *Interdisciplinaria*, 27(1):7–22, 2010.

[79] Nancy R Ahern, Ermalynn M Kiehl, Mary Lou Sole, and Jacqueline Byers. A review of instruments measuring resilience. *Issues in comprehensive Pediatric nursing*, 29(2):103–125, 2006.

[80] Gail Wagnild. A review of the Resilience Scale. *Journal of nursing measurement*, 17(2):105–113, 2009.

[81] Marcus W. Agelink, Cavit Boz, Heiko Ullrich, and Jürgen Andrich. Relationship between major depression and heart rate variability. Clinical consequences and implications for antidepressive treatment. *Psychiatry Research*, 113(1-2):139–149, 2002.

[82] Andrew H. Kemp, Daniel S. Quintana, Marcus A. Gray, Kim L. Felmingham, Kerri Brown, and Justine M. Gatt. Impact of Depression and Antidepressant Treatment on Heart Rate Variability: A Review and Meta-Analysis. *Biological Psychiatry*, 67(11):1067–1074, 2010.

[83] Julian Koenig, Andrew H Kemp, Theodore P Beauchaine, Julian F Thayer, and Michael Kaess. Depression and resting state heart rate variability in children and adolescents — A systematic review and meta-analysis. *Clinical Psychology Review*, 46:136–150, 2016.

[84] Daniel N. Klein, Roman Kotov, and Sara J. Bufferd. Personality and Depression: Explanatory Models and Review of the Evidence. *Annual Review of Clinical Psychology*, 7(1):269–295, 2011.

[85] R. Michael Bagby, Lena C. Quilty, and Andrew C. Ryder. Personality and depression. *Canadian Journal of Psychiatry*, 53(1):14–25, 2008.

[86] Darcy A. Santor, R. Michael Bagby, and Russell T. Joffe. Evaluating stability and change in personality and depression. *Journal of Personality and Social Psychology*, 73(6):1354–1362, 1997.

[87] Terry C Camacho, Robert E Roberts, Nancy B Lazarus, George A Kaplan, and Richard D Cohen. Physical activity and depression: evidence from the Alameda County Study. *American journal of epidemiology*, 134(2):220–231, 1991.

[88] Charlotte M. McKercher, Michael D. Schmidt, Kristy A. Sanderson, George C. Patton, Terence Dwyer, and Alison J. Venn. Physical Activity and Depression in Young Adults. *American Journal of Preventive Medicine*, 36(2):161–164, 2009.

[89] Andreas Ströhle. Physical activity, exercise, depression and anxiety disorders. *Journal of Neural Transmission*, 116(6):777–784, 2009.

[90] Antonina Luca, Maria Luca, and Carmela Calandra. Sleep disorders and depression : brief review of the literature , case report , and nonpharmacologic interventions for depression. *Clinical interventions in aging*, 8:1033–1039, 2013.

[91] Melissa Vandeputte and Al de Weerd. Sleep disorders and depressive feelings: A global survey with the Beck depression scale. *Sleep Medicine*, 4(4):343–345, 2003.

[92] Paul E. Peppard, Mariana Szklo-Coxe, K. Mae Hla, and Terry Young. Longitudinal association of sleep-related breathing disorder and depression. *Archives of Internal Medicine*, 166(16):1709–1715, 2006.

[93] H. M. Van Praag. Can stress cause depression? *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 28(5):891–907, 2004.

[94] Avshalom Caspi, Karen Sugden, Terrie E. Moffitt, Alan Taylor, Ian W. Craig, Hona Lee Harrington, Joseph McClay, Jonathan Mill, Judy Martin, Antony Braithwaite, and Richie Poulton. Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. *Science*, 301(5631):386–389, 2003.

[95] Allison T. Knoll and William A. Carlezon. Dynorphin, stress, and depression. *Brain Research*, 1314:56–73, 2010.

[96] Christiaan H. Vinkers, Renske Penning, Juliane Hellhammer, Joris C. Verster, John H.G.M. Klaessens, Berend Olivier, and Cor J. Kalkman. The effect of stress on core and peripheral body temperature in humans. *Stress*, 16(5):520–530, 2013.

[97] Donatella Marazziti, Angela Di Muro, and Paolo Castrogiovanni. Psychological stress and body temperature changes in humans. *Physiology and Behavior*, 52(2):393–395, 1992.

[98] Clemens Kirschbaum, Karl-Martin Pirke, and Dirk H Hellhammer. The 'Trier Social Stress Test'–a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology*, 28(1-2):76–81, 1993.

[99] Ulf Lundberg, Roland Kadefors, Bo Melin, Gunnar Palmerud, Peter Hassmén, Margareta Engström, and Ingela Elfsberg Dohns. Psychophysiological stress and EMG activity of the trapezius muscle. *International journal of behavioral medicine*, 1(4):354–370, 1994.

[100] Rosan Luijcks, Hermie J Hermens, Lonneke Bodar, Catherine J Vossen, Jim Van Os, and Richel Lousberg. Experimentally induced stress validated by EMG activity. *PloS one*, 9(4):e95215, 2014.

[101] Arpit Christian. Analysis of physiological signals to measure physiological stress, 2018.

[102] William M Suess, A Barney Alexander, Deborah D Smith, Helga W Sweeney, and Richard J Marion. The effects of psychological stress on respiration: a preliminary study of anxiety and hyperventilation. *Psychophysiology*, 17(6):535–540, 1980.

[103] Yuri Masaoka and Ikuo Homma. Anxiety and respiratory patterns: their relationship during mental stress and physical load. *International Journal of Psychophysiology*, 27(2):153–159, 1997.

[104] Birgitta Berglund, Ulf Berglund, and Thomas Lindvall. Scaling loudness, noisiness, and annoyance of aircraft noise. *The Journal of the Acoustical Society of America*, 57(4):930–934, 1975.

[105] M. Harker. Psychological sweating: A systematic review focused on aetiology and cutaneous response. *Skin Pharmacology and Physiology*, 26(2):92–100, 2013.

[106] Hsiao-Pei Lin, Hung-Yu Lin, Wei-Lun Lin, and Andrew Chih-Wei Huang. Effects of stress, depression, and their interaction on heart rate, skin conductance, finger temperature, and respiratory rate: sympathetic-parasympathetic hypothesis of stress and depression. *Journal of clinical psychology*, 67(10):1080–1091, 2011.

[107] Nis Hjortskov, Dag Rissén, Anne Katrine Blangsted, Nils Fallentin, Ulf Lundberg, and Karen Søgaard. The effect of mental stress on heart rate variability and blood pressure during computer work. *European journal of applied physiology*, 92(1):84–89, 2004.

[108] Kaviraja Udupa, T. N. Sathyaprabha, Jagadisha Thirthalli, K. R. Kishore, G. S. Lavekar, T. R. Raju, and B. N. Gangadhar. Alteration of cardiac autonomic functions in patients with major depression: A study using heart rate variability measures. *Journal of Affective Disorders*, 100(1-3):137–141, 2007.

[109] Lydia Brown, Chandan Karmakar, Richard Gray, Ripu Jindal, Terrence Lim, and Christina Bryant. Heart rate variability alterations in late life depression: A meta-analysis. *Journal of Affective Disorders*, 235(March):456–466, 2018.

[110] Arkaprabha Sau and Ishita Bhakta. Screening of anxiety and depression among seafarers using machine learning technology. *Informatics in Medicine Unlocked*, 16(December):100228, 2019.

[111] Christopher M. Hatton, Lewis W. Paton, Dean McMillan, James Cussens, Simon Gilbody, and Paul A. Tiffin. Predicting persistent depressive symptoms in older adults: A machine learning approach to personalised mental healthcare. *Journal of Affective Disorders*, 246(December 2018):857–860, 2019.

[112] Le Yang, Dongmei Jiang, Lang He, Ercheng Pei, Meshia Cédric Oveneke, and Hichem Sahli. Decision tree based depression classification from audio video and language information. In *Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge*, pages 89–96. ACM, 2016.

[113] Krešimir Ćosić, Marko Šarlija, Vladimir Ivkovic, Quan Zhang, Gary Strangman, and Siniša Popović. Stress resilience assessment based on physiological features in selection of air traffic controllers. *IEEE access*, 7:41989–42005, 2019.

[114] Wei Lü, Zhenhong Wang, and Xuqun You. Physiological responses to repeated stress in individuals with high and low trait resilience. *Biological Psychology*, 120:46–52, 2016.

[115] Rüdiger Wirth and Jochen Hipp. CRISP-DM: Towards a standard process model for data mining. In *Proceedings of the 4th international conference on the practical applications of knowledge discovery and data mining*, volume 1. Springer-Verlag London, UK, 2000.

[116] A T Beck. Depression: causes and treatment. Philadelphia, PA: Univer, 1967.

[117] Linda C Hatzenbuehler, Mary Parpal, and Leroy Matthews. Classifying college students as depressed or nondepressed using the Beck Depression Inventory: An empirical analysis. *Journal of Consulting and Clinical Psychology*, 51(3):360, 1983.

[118] Arturo Nakasone, Helmut Prendinger, and Mitsuru Ishizuka. Emotion recognition from electromyography and skin conductance. In *Proc. of the 5th international workshop on biosignal interpretation*, pages 219–222. Citeseer, 2005.

[119] Azadeh Kushki, Jillian Fairley, Satyam Merja, Gillian King, and Tom Chau. Comparison of blood volume pulse and skin conductance responses to mental and affective stimuli at different anatomical sites. *Physiological measurement*, 32(10):1529, 2011.

[120] Carlo J De Luca. Electromyography. Encyclopedia of medical devices and instrumentation. *John Wiley Publisher*, pages 98–109, 2006.

[121] Jesse C Stewart and Christopher R France. Cardiovascular recovery from stress predicts longitudinal changes in blood pressure. *Biological psychology*, 58(2):105–120, 2001.

[122] B V P Prasad and Velusamy Parthasarathy. Detection and classification of cardiovascular abnormalities using FFT based multi-objective genetic algorithm. *Biotechnology & Biotechnological Equipment*, 32(1):183–193, 2018.

[123] Pedro Ferreira, Duc C Le, and Nur Zincir-Heywood. Exploring feature normalization and temporal information for machine learning based insider threat detection. In *2019 15th International Conference on Network and Service Management (CNSM)*, pages 1–7. IEEE, 2019.

[124] Nornadiah Mohd Razali and Yap Bee Wah. Power comparisons of shapiro-wilk, kolmogorov-smirnov, lilliefors and anderson-darling tests. *Journal of statistical modeling and analytics*, 2(1):21–33, 2011.

[125] Jerrold H Zar. Significance testing of the Spearman rank correlation coefficient. *Journal of the American Statistical Association*, 67(339):578–580, 1972.

[126] Eric Heikkila. Multicollinearity in regression models with multiple distance measures. *Journal of Regional Science*, 28(3):345–362, 1988.

[127] Abubakari S Gwelo. Principal components to overcome multicollinearity problem. *Oradea Journal of Business and Economics*, 4(1):79–91, 2019.

[128] Edward R Mansfield and Billy P Helms. Detecting multicollinearity. *The American Statistician*, 36(3a):158–160, 1982.

[129] Svante Wold, Kim Esbensen, and Paul Geladi. Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3):37–52, 1987.

[130] Shiming Xiang, Feiping Nie, and Changshui Zhang. Learning a Mahalanobis distance metric for data clustering and classification. *Pattern recognition*, 41(12):3600–3612, 2008.

[131] Ujjwal Maulik and Sanghamitra Bandyopadhyay. Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on pattern analysis and machine intelligence*, 24(12):1650–1654, 2002.

[132] Steffen Herbold. Autorank: A python package for automated ranking of classifiers. *Journal of Open Source Software*, 5(48):2173, 2020.

[133] Dulce G Pereira, Anabela Afonso, and Fátima Melo Medeiros. Overview of Friedman's test and post-hoc analysis. *Communications in Statistics-Simulation and Computation*, 44(10):2636–2653, 2015.

Curriculum Vitae

Ramon E. Diaz-Ramos was born in Guadalajara, Jalisco, Mexico, in 1992. He received a B.S. degree in mechatronics engineering from Tecnologico de Monterrey University in 2015. He was accepted to study his Master in Computer Science with a Machine Learning specialization from Tecnologico de Monterrey University with a full scholarship. His research interests include machine learning applied to mental health detection and data science applications in technology.

This document was typed in using $\LaTeX\,2_{\varepsilon}$[a] by Ramón Eduardo Díaz Ramos.

[a]The style file `phdThesisFormat.sty` used to set up this thesis was prepared by the Center of Intelligent Systems of the Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey Campus