



**COMILLAS**  
UNIVERSIDAD PONTIFICIA

ICAI

# MÁSTER UNIVERSITARIO EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER

Detección de perfiles psicopáticos mediante el análisis  
automático de señales de encefalogramas

Autor: Erik Martin Vik Espárrago

Director: Dr. Mario Castro Ponce

Madrid

Julio de 2021

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título  
Detección de perfiles psicopáticos mediante el análisis automático de señales de  
electroencefalograma

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2020/2021 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido

tomada de otros documentos está debidamente referenciada.

*Erik M. Vik*

Fdo.: Erik Martin Vik Espárrago

Fecha: 20/ 07/ 2021

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO



Fdo.: Dr. Mario Castro Ponce

Fecha: 20/ 07/ 2021VI





**COMILLAS**  
UNIVERSIDAD PONTIFICIA

ICAI

# MÁSTER UNIVERSITARIO EN INGENIERÍA INDUSTRIAL

TRABAJO FIN DE MÁSTER

Detección de perfiles psicopáticos mediante el análisis  
automático de señales de encefalogramas

Autor: Erik Martin Vik Espárrago

Director: Dr. Mario Castro Ponce

Madrid

Julio de 2021

# Agradecimientos

Quiero agradecer a mi tutor, Mario la ayuda que me ha proporcionado a lo largo de este proyecto, tanto con su experiencia profesional, como con su asesoramiento personal; este proyecto habría sido inviable sin su apoyo y su paciencia.

También quiero agradecer a mi familia el soporte que me han brindado durante la preparación de este trabajo; y especialmente a mi madre, Bárbara, por enseñarme la importancia de hacer bien las cosas y del esfuerzo personal.

# DETECCIÓN DE PERFILES PSICOPÁTICOS MEDIANTE EL ANÁLISIS AUTOMÁTICO DE SEÑALES DE ENCEFALOGRAMAS.

**Autor: Vik Espárrago, Erik.**

Director: Castro Ponce, Mario

Entidad Colaboradora: ICAI – Universidad Pontificia Comillas

## RESUMEN DEL PROYECTO

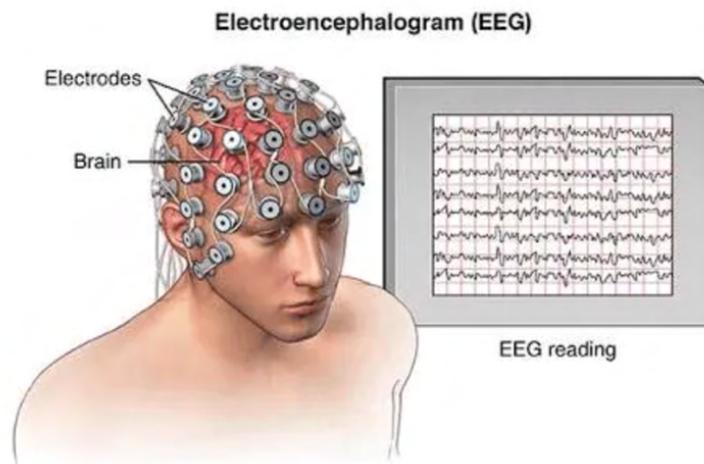
En este proyecto se han analizado las señales de la corteza cerebrales de un grupo de jóvenes, con la idea de crear un modelo que permita identificar aquellos individuos con rasgos altos de psicopatía. Para ello, se han estudiado las señales obtenidas con un electroencefalograma. Por un lado, se ha hecho un estudio de los potenciales relacionados con eventos, y por otro se ha usado el aprendizaje automático para intentar clasificar los individuos.

**Palabras clave:** EEG, aprendizaje automático, psicópata.

### 1. Introducción

El electroencefalograma es una prueba que se utiliza para estudiar el funcionamiento del sistema nervioso central, concretamente de la actividad de la corteza del cerebro. Esta prueba realiza un seguimiento de las ondas cerebrales, utilizando electrodos fijados sobre el cuero cabelludo que van registrando las corrientes eléctricas que se forman en las neuronas cerebrales [1].

Un electroencefalograma es capaz de detectar cambios en la actividad cerebral, es decir en el potencial eléctrico del cerebro. Estos cambios pueden ser útiles para diagnosticar trastornos cerebrales, especialmente epilepsia u otros trastornos convulsivos [2]. Un electroencefalograma también puede ser útil para diagnosticar o tratar otros trastornos, como tumores cerebrales, daños por lesiones en la cabeza, encefalopatía. En la Ilustración 1 se puede ver una representación de un electroencefalograma, y las señales que se obtienen de este.



*Ilustración 1 Representación de un electroencefalograma.*

En vistas del potencial del EEG para el estudio de la actividad cerebral, en este proyecto se propone la creación de una herramienta que permita el procesado de señales y el análisis exploratorio de señales tomadas a adolescentes con antecedentes delictivos. Finalmente, la utilización de herramientas de Aprendizaje Automático para la detección de perfiles psicopáticos mediante el análisis y clasificación de las señales de EEG.

## **2. Definición del Proyecto**

Los instrumentos que se utilizan actualmente para el diagnóstico de la psicopatía, presentan algunas dificultades. Principalmente, el grado de confianza de la información obtenida, puesto que las pruebas suelen ser cuestionarios autoadministrados. Es decir, su diagnóstico es costoso *“no hay marcadores o pruebas físicas o biológicas que definan los problemas de salud mental”* [3].

El principal inconveniente de los EEG, es la dificultad de procesamiento y análisis de las señales, que requieren mucha especialización.

Por todo esto, en este proyecto se propone crear una herramienta de análisis de encefalogramas. Además la utilización de herramientas de Aprendizaje Automático para la detección de perfiles psicopáticos de señales tomadas a adolescentes con antecedentes delictivos. Tener un instrumento que apoye el diagnóstico clínico basado en cuestionarios, permitiría adelantar y facilitar el diagnóstico de esta enfermedad, lo que ayudaría a tratar y dar la ayuda necesaria a estas personas, pues además este instrumento que se propone, no es invasivo y se puede hacer sin alterar a las personas, siendo por lo tanto de gran utilidad.

Para ello se tendrán que cumplir una serie de objetivos de este proyecto son:

- Crear una herramienta integrada de análisis de encefalogramas que permita visualizar y analizar dichos datos.
- Analizar y clasificar las señales tomadas de adolescentes delictivos con el uso del aprendizaje automático.
- Establecer una jerarquía de los algoritmos más eficaces para este tipo de análisis

## **3. Metodología**

En primer lugar, se investigó el concepto de psicopatía y sus formas de diagnóstico. Tras esto, se estudió cómo es la estructura, y las distintas componentes de las señales que se obtienen de los EEG. Una vez se comprendieron estos dos conceptos, se procedió a realizar un estudio de posibles soluciones que trataran el uso de EEG y el aprendizaje automático.

En segundo lugar, se procedió a realizar un análisis exploratorio de los datos que se iban a utilizar en el proyecto. Para comprender mejor, las pruebas a las que se sometieron a los sujetos y cómo fueron los resultados de estas. Tras esto, se intentó identificar alguna forma de clasificación de los sujetos en base a sus respuestas a las pruebas.

En tercer lugar, se procedió al pre-procesado de las señales, que incluye la detección de artificios, su filtrado, y la extracción de epochs. Tras esto, se usaron estas señales para 2 estudios diferentes: el estudio de los potenciales relacionados con eventos, y la clasificación de los individuos en base a estos; la extracción de características de las señales y la clasificación de estas.

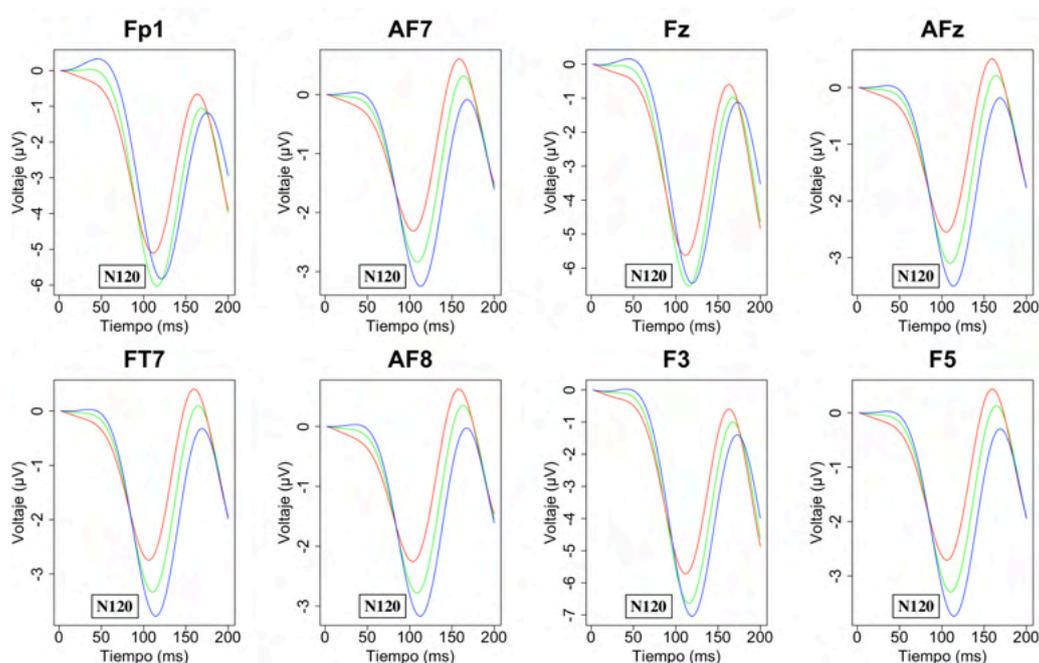
## 4. Resultados

En primer lugar, para comprender mejor las respuestas, las características y el desempeño de cada individuo en la prueba Ekman, se procedió a calcular la exactitud, sensibilidad y especificidad de cada sujeto. Estos se pueden ver en las tablas a continuación.

|                | Exactitud | Sensibilidad | Especificidad |
|----------------|-----------|--------------|---------------|
| <b>Sergio</b>  | 87%       | 88%          | 86%           |
| <b>Samuel</b>  | 34%       | 35%          | 32%           |
| <b>Luis</b>    | 71%       | 64%          | 80%           |
| <b>Wender</b>  | 70%       | 68%          | 72%           |
| <b>Héctor</b>  | 51%       | -            | 51%           |
| <b>Claudio</b> | 75%       | 85%          | 70%           |
| <b>Mikel</b>   | 87%       | 88%          | 86%           |

|                  | Exactitud | Sensibilidad | Especificidad |
|------------------|-----------|--------------|---------------|
| <b>Alejandro</b> | 80%       | 81%          | 79%           |
| <b>Andrei</b>    | 62%       | 60%          | 64%           |
| <b>David</b>     | 73%       | 90%          | 67%           |
| <b>Diego</b>     | 21%       | 17%          | 24%           |
| <b>Ivan</b>      | 73%       | 74%          | 72%           |
| <b>Jeison</b>    | 77%       | 72%          | 83%           |
| <b>Gabriel</b>   | 36%       | 36%          | 34%           |

Más allá, se realizó un análisis de los potenciales relacionados con eventos. En la Ilustración 2 se pueden ver los resultados que se obtienen para el lóbulo frontal.



*Ilustración 2 ERP para el lóbulo frontal.*

Tras realizar la media de las señales, se pueden apreciar diferencias entre los dos grupos. Se puede ver que la componente N120 para el grupo con HCU (rojo) tiene una mayor amplitud que el grupo LCU (azul), esto concuerda con la hipótesis de que los sujetos con HCU tienen dificultad para procesar las emociones, ya sean de miedo o neutras.

Más allá, se observa mayor diferencia en los grupos para las zonas temporales que están relacionadas con la afectividad, y el reconocimiento de rostros, y para las zona frontal que procesa las emociones.

Finalmente, se realizó una transformación discreta wavelet a las señales, y se extrajeron algunas características como la entropía y la energía en cada nivel. Con estas variables se

construyó un clasificador SVM. En la Ilustración 3 se puede ver las distintas curvas AUC-ROC que se obtienen.

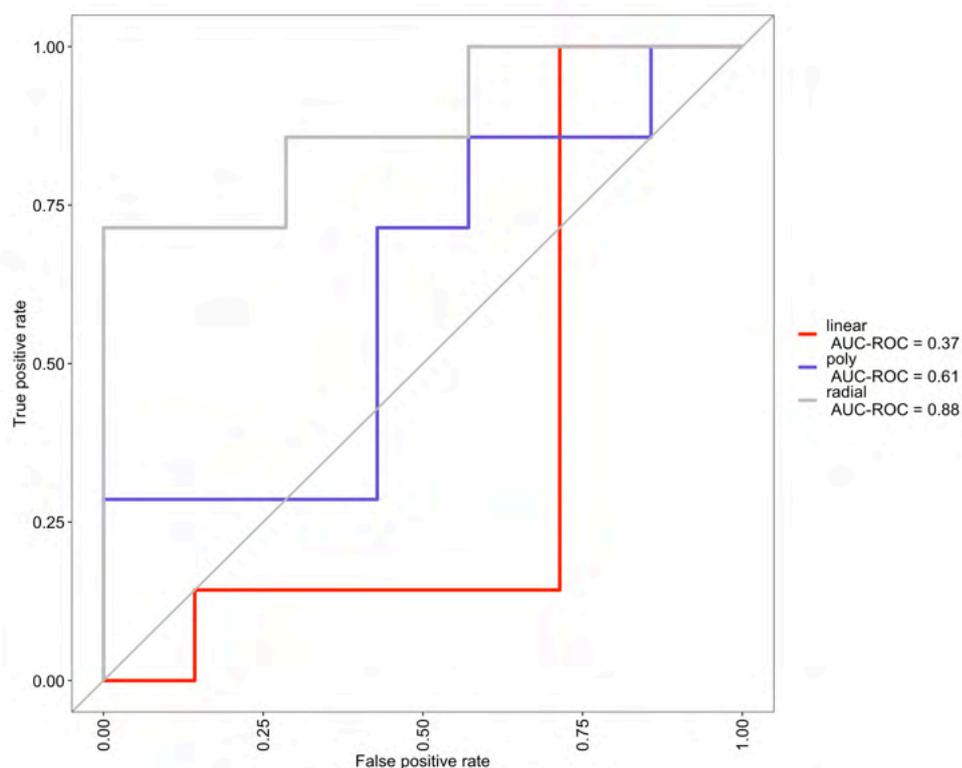


Ilustración 3 Curva AUC-ROC de los clasificadores SVM.

Viendo la curva AUC-ROC, se observa que el mejor modelo de clasificación SVM es el radial. Obteniendo un área bajo la curva del 88%, unos resultados bastante buenos teniendo en cuenta la naturaleza de las señales. Asimismo, se observa que puesto que el SVM radial proporciona mejores resultados que el lineal. Esto muestra claramente que no hay una relación lineal entre el valor de la entropía de cada canal y el nivel de psicopatía de los sujetos.

## 5. Conclusiones

En este proyecto, se realizó un estudio sobre las señales cerebrales obtenidas de sujetos en un reformatorio, y se intentó realizar un modelo de clasificación que permitiese distinguir aquellos sujetos con rasgos HCU, y aquellos con rasgos LCU. Para ello se realizaron 3 estudios diferentes, primero se estudiaron las respuestas a la prueba ekman y se intentó ver si esto permitía distinguir a los individuos. En segundo lugar, se realizó un estudio de los potenciales relacionados con eventos y se compararon estas señales entre los distintos sujetos de estudio. Finalmente, se extrajeron características de las señales y se intentó crear un modelo de clasificación automático.

Del estudio de las respuestas de la prueba ekman se concluyó, que estas no permitían distinguir a los individuos. Esto se debió a que los resultados de las pruebas mostraron muchos fallos por parte de todos los sujetos, lo que supone que la prueba podría ser complicada. Además, se pudo ver que algunos individuos no entendieron bien la prueba, y sus respuestas parecían estar al revés, es decir, que presionaban el botón que no debían para cada imagen. Además, algunos sujetos realizaron la prueba sin seriedad, respondiendo todo

el rato lo mismo o aparentemente aleatoriamente. También la dificultad para identificar las imágenes puede venir de la poca socialización a la que se enfrentan ambos grupos, y por ello la tasa de error en las respuestas es tan alta [4].

Con el estudio de los ERP de cada sujeto, se llegó a la conclusión de que aquellos con rasgos HCU tenían mayor dificultad para procesar las imágenes. Y las señales del EEG para la componente N120 presentaban menor activación para el grupo HCU frente al grupo LCU. Esta menor activación se puede ver a lo largo de toda la corteza cerebral, pero especialmente en la zona frontal, que es la encargada de las emociones.

Con la extracción de características, se llegó a la conclusión de que estas si permitían la diferenciación de sujetos, es decir, que las señales de cada sujeto son bastante diferentes. Pero aun así existe dificultad a la hora de construir un clasificador, debido a la naturaleza de las señales, donde los cambios son muy sutiles, y a la pequeña muestra con la que se ha trabajado. Con todo ello se consiguió crear un clasificador SVM, con forma radial que permite cierta clasificación de los individuos. Pero este modelo no permitiría sustituir las pruebas que existen actualmente, aunque un mayor desarrollo podría permitir su uso como herramienta complementaria.

## 6. Referencias

- [1] R. C. H. S. Diego, «<https://www.rchsd.org/>,» [En línea]. Available: <https://www.rchsd.org/health-articles/electroencefalograma-eeeg/>.
- [2] Mayo Clinic, [En línea]. Available: <https://www.mayoclinic.org/es-es/tests-procedures/eeg/about/pac-20393875>.
- [3] Consejo general de psicología de España, [En línea]. Available: [http://www.infocop.es/view\\_article.asp?id=8176](http://www.infocop.es/view_article.asp?id=8176).
- [4] L. Halty, «Impairment in the processing of fear gaze in adolescents with callous–unemotional traits,» Madrid, 2019.

# DETECTION OF PSICOPATH PROFILES BY USING MACHINE LEARNING ON EEG SIGNALS

**Author:** Vik Espárrago, Erik.

Supervisor: Castro Ponce, Mario.

Collaborating Entity: ICAI – Universidad Pontificia Comillas

## ABSTRACT

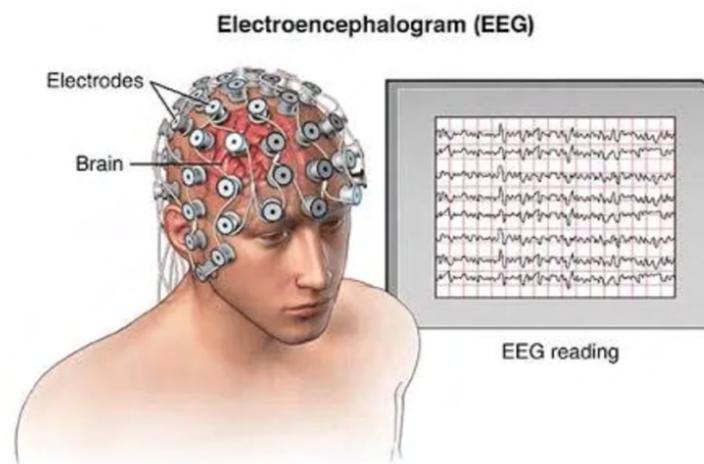
In this project, the signals of the cerebral cortex of a group of young people have been analyzed, with the idea of creating a model that would allow to identify those with high psychopathy traits. For this, the signals obtained with an EEG have been studied. First, the ERP have been studied, and then we have created a SVM classifier.

**Keywords:** EEG, machine learning, psicopath

## 1. Introduction

The electroencephalogram is a test used to study the central nervous system, specifically the activity of the cerebral cortex. This test tracks brain waves by using electrodes attached to the scalp, these electrodes record the electrical currents that are formed in brain neurons [1].

An EEG is capable of detecting changes in brain activity, specifically, the changes in the brains electrical potential. These changes can be helpful in diagnosing brain disorders, such as epilepsy or other seizure disorders [2]. An electroencephalogram can also be useful to diagnose or treat other disorders, such as brain tumors, concussions... In Figure 1 you can see a representation of an electroencephalogram, and the signals that are obtained from it.



*Figure 1 Representation of an EEG.*

In view of the potential of the EEG in the study of brain activity, this project proposes the development of a tool for EEG analysis and the exploratory analysis of signals taken from teenagers with criminal records. Finally, the use of Machine Learning tools for the detection of psychopathic profiles.

## **2. Project definition**

The instruments currently used for the diagnosis of psychopathy present some difficulties. Mainly, the degree of confidence of the information obtained, since the tests are usually self-administered. In other words, its diagnosis is difficult "there are no physical or biological tests that define mental health problems" [3].

The main drawback of EEGs is the difficulty in processing and analyzing the signals, which require a lot of specialization and computer power.

For all this, in this project it is proposed to develop an EEG analysis tool. In addition, the use of Machine Learning tools for the detection of psychopathic profiles of signals taken from teenagers with a criminal record. Since having an instrument that supports the diagnosis based on questionnaires, would allow to advance and facilitate the diagnosis of this disease, which would help treat these people.

For this, a series of objectives will have to be met:

- Develop an integrated EEG analysis tool that allows visualization and analysis of said data.
- Analyze and classify the signals taken from teenagers with the use of machine learning.
- Establish a hierarchy of the most efficient algorithms for this type of analysis

## **3. Methodology**

First, we investigated the concept of psychopathy and its forms of diagnosis. After this, we study the structure, and the different components of the signals. Once these two concepts were understood, we began to carry out a study of possible solutions that had been developed in this area.

Secondly, an exploratory analysis was carried out to better understand the tests that were done to the subjects. After this, we tried to identify some form of classification of the subjects based on their responses to the tests.

Thirdly, the signals were pre-processed, which included the detection of artifacts, the filtering, and the extraction of epochs. After this, these signals were used for 2 different studies: the study of event-related potentials, and the classification of individuals based on these; the extraction of characteristics of the signals and the classification of these.

#### 4. Results

First, to better understand the responses, characteristics, and performance of each individual in the Ekman test, the accuracy, sensitivity, and specificity of each subject were calculated.

|                | Exactitud | Sensibilidad | Especificidad |
|----------------|-----------|--------------|---------------|
| <b>Sergio</b>  | 87%       | 88%          | 86%           |
| <b>Samuel</b>  | 34%       | 35%          | 32%           |
| <b>Luis</b>    | 71%       | 64%          | 80%           |
| <b>Wender</b>  | 70%       | 68%          | 72%           |
| <b>Héctor</b>  | 51%       | -            | 51%           |
| <b>Claudio</b> | 75%       | 85%          | 70%           |
| <b>Mikel</b>   | 87%       | 88%          | 86%           |

|                  | Exactitud | Sensibilidad | Especificidad |
|------------------|-----------|--------------|---------------|
| <b>Alejandro</b> | 80%       | 81%          | 79%           |
| <b>Andrei</b>    | 62%       | 60%          | 64%           |
| <b>David</b>     | 73%       | 90%          | 67%           |
| <b>Diego</b>     | 21%       | 17%          | 24%           |
| <b>Ivan</b>      | 73%       | 74%          | 72%           |
| <b>Jeison</b>    | 77%       | 72%          | 83%           |
| <b>Gabriel</b>   | 36%       | 36%          | 34%           |

Furthermore, an analysis of the potentials related to events was carried out, in the Figure 2 you can see the results obtained for the frontal lobe.

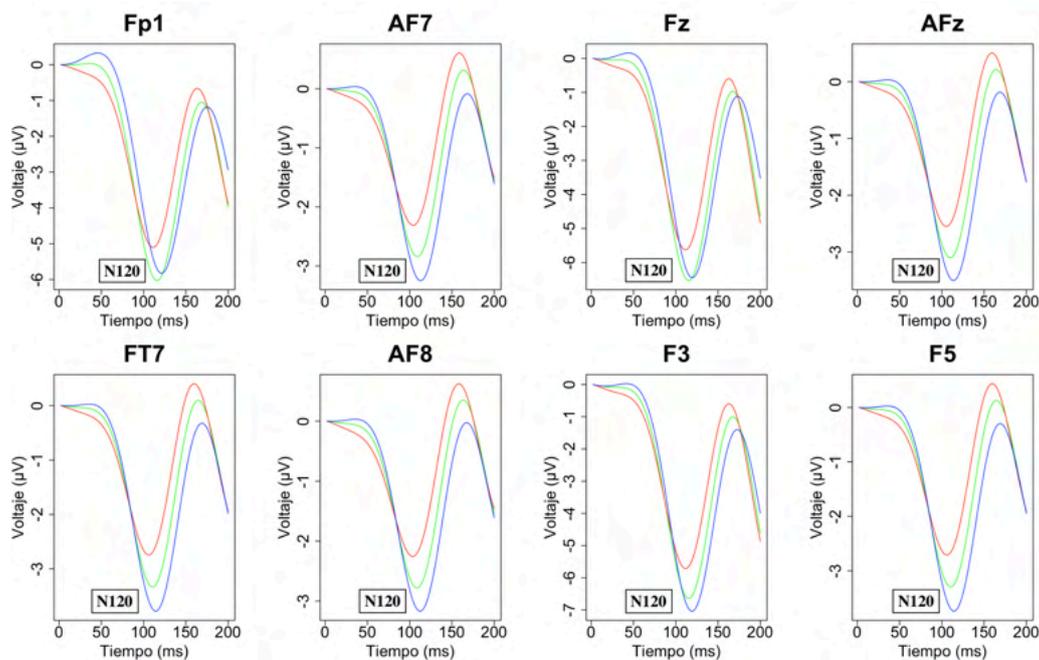


Figure 2 ERP of the frontal lobe.

After averaging the signals, differences between the two groups can be seen. It can be seen that the N120 component for the group with HCU (red) traits has a greater amplitude than the group LCU (blue), this agrees with the hypothesis that subjects with HCU have difficulty processing emotions, whether of fear or neutral.

Furthermore, a greater difference is observed in the groups for the temporal lobe that are related to affectivity, and face recognition, and for the frontal lobe that process emotions.

Finally, a discrete wavelet transformation was performed to the signals, and some characteristics such as entropy and energy were extracted at each level. With these

variables an SVM classifier was built. In the Figure 3 you can see the different AUC-ROC curves that are obtained.

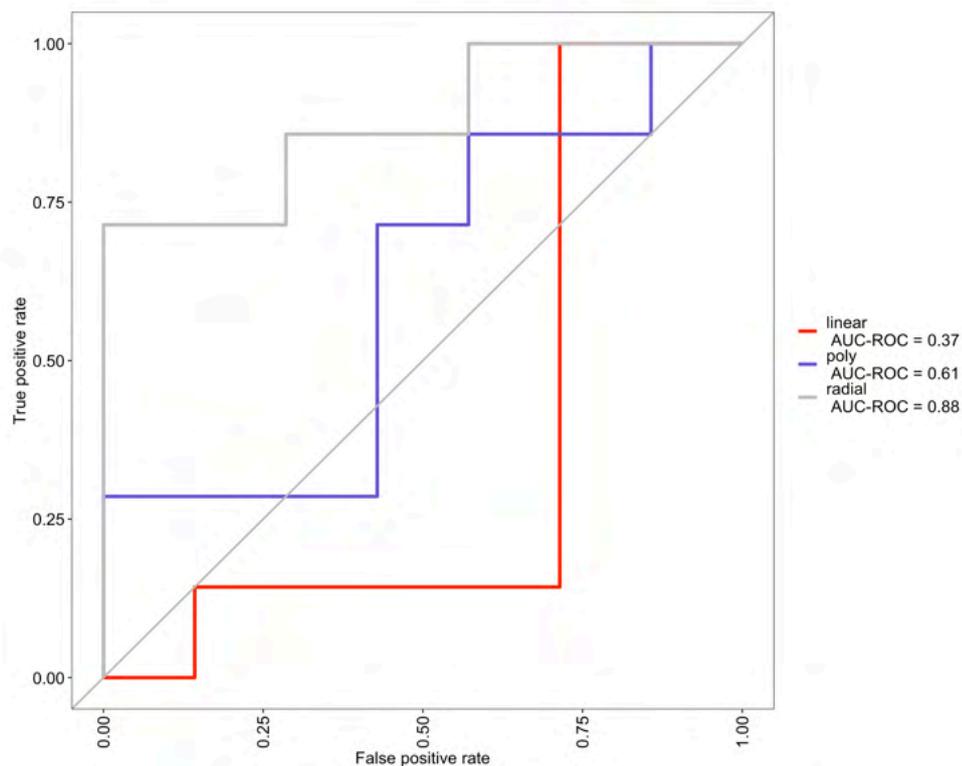


Figure 3 AUC-ROC curve for SVM classifier.

Looking at the AUC-ROC curve, it is observed that the best SVM classification model is the radial. Getting an area under the curve of 88%, which are good results taking into account the nature of the signals. Likewise, it is observed that since the radial SVM provides better results than the linear, this shows that there is no linear relationship between the value of the entropy of each channel and the level of psychopathy of the subjects.

## 5. Conclusions

In this project, a study was carried out on the brain signals obtained from subjects in a detention center, and an attempt was made to develop a classification model that would allow us to distinguish those subjects with HCU traits, and those with LCU traits. For this, 3 different studies were carried out. First the responses to the ekman test were studied and an attempt was made to see if this allowed individuals to be distinguished. Second, a study of event-related potentials was conducted and these signals were compared between the different study subjects. Finally, characteristics of the signals were extracted and an attempt was made to create an automatic classification model.

From the study of the responses of the ekman test, it was concluded that they did not allow to distinguish individuals. This was because the test results showed many failures on the part of all the subjects, which means that the test could be complicated. In addition, it could be seen that some individuals did not understand the test well, and their answers seemed to be backwards, that is, that they pressed the wrong button for each image. Furthermore, some subjects performed the test without seriousness, answering the same or apparently randomly all the time. Also, the difficulty to identify the images can come

from the little socialization that both groups face, and for this reason the error rate in the answers is so high [4].

With the study of the ERP of each subject, it was concluded that those with HCU traits had greater difficulty in processing the images. And the EEG signals for the N120 component showed less activation for the HCU group compared to the LCU group. This lower activation can be seen throughout the entire cerebral cortex, but especially in the frontal area, which is in charge of emotions.

With the extraction of characteristics, it was concluded that these did allow the differentiation of subjects, that is, that the signals of each subject are quite different. But even so there is difficulty when building a classifier, due to the nature of the signals where the changes are very subtle, and the small sample that has been worked with. With all this, it was possible to create an SVM classifier, with a radial shape that allows a certain classification of individuals. But this model would not allow to replace the tests that currently exist, although further development could allow its use as a complementary tool.

## 6. References

- [1] R. C. H. S. Diego, «<https://www.rchsd.org/>,» [En línea]. Available: <https://www.rchsd.org/health-articles/electroencefalograma-eeeg/>.
- [2] Mayo Clinic, [En línea]. Available: <https://www.mayoclinic.org/es-es/tests-procedures/eeeg/about/pac-20393875>.
- [3] Consejo general de psicología de España, [En línea]. Available: [http://www.infocop.es/view\\_article.asp?id=8176](http://www.infocop.es/view_article.asp?id=8176).
- [4] L. Halty, «Impairment in the processing of fear gaze in adolescents with callous–unemotional traits,» Madrid, 2019.



## *Índice de la memoria*

|  |           |
|--|-----------|
| <b>Capítulo 1. Introducción.....</b>   | <b>6</b>  |
| <b>Capítulo 2. Estado de la Cuestión.....</b>  | <b>12</b> |
| 2.1 Aplicaciones del aprendizaje automático para la clasificación de señales de eeg..... | 12        |
| 2.1.1 Clasificación según la visión .....  | 12        |
| 2.1.2 Clasificación de movimientos de la mano d/i a través de eeg .....                  | 12        |
| 2.1.3 Diagnóstico de patologías detectables con eeg.....                                 | 13        |
| 2.1.4 Diagnóstico de la epilepsia.....   | 13        |
| 2.1.5 Clasificación del estado emocional .....   | 14        |
| 2.1.6 Predicción de la edad usando EEG .....   | 14        |
| 2.2 EEG en la psicología y criminología .....  | 14        |
| <b>Capítulo 3. Definición del Trabajo.....</b>   | <b>16</b> |
| 3.1 Justificación .....  | 16        |
| 3.2 Objetivos .....  | 17        |
| 3.3 Metodología .....  | 17        |
| 3.4 Recursos a emplear .....   | 18        |
| 3.5 Estudio económico.....   | 19        |
| 3.5.1 Mediciones .....   | 19        |
| 3.5.2 Precio unitario .....  | 20        |
| 3.5.3 Suma total .....   | 20        |
| <b>Capítulo 4. Análisis y estructura de los datos .....</b>                              | <b>21</b> |
| 4.1 Origen de los datos.....   | 21        |
| 4.2 Estructura de los datos .....  | 24        |
| 4.3 Análisis exploratorio de los datos .....   | 26        |
| 4.3.1 Análisis de las distintas pruebas.....   | 26        |
| 4.3.2 Análisis de sujetos.....   | 27        |
| 4.3.3 Reducción de dimensiones con PCA.....  | 29        |
| 4.4 Análisis de los resultados de la prueba Ekman.....                                   | 30        |
| <b>Capítulo 5. Procesado de las señales.....</b>   | <b>36</b> |

|  |   |           |
|--|---|-----------|
| 5.1  | Evento .....  | 36        |
| 5.2  | Señales sin procesar .....                              | 37        |
| 5.3  | Filtrado de artificios de las señales .....             | 38        |
| 5.4  | Filtrado de los datos .....                             | 40        |
| 5.5  | Epoching/Segmentación de los datos.....                 | 41        |
| <b>Capítulo 6. Potencial relacionado con eventos .....</b>             |   | <b>44</b> |
| 6.1  | Potencial relacionado con eventos .....                 | 44        |
| 6.2  | Media de los sujetos.....                               | 44        |
| 6.3  | Media de las pruebas.....                               | 46        |
| 6.4  | Análisis de potenciales relacionados con eventos.....   | 48        |
| 6.4.1  | Potenciales relacionados con eventos por sujeto .....   | 48        |
| 6.4.2  | Potenciales relacionados con eventos de la prueba ..... | 51        |
| <b>Capítulo 7. Extracción de características y clasificación .....</b> |   | <b>56</b> |
| 7.1  | Extracción de características.....                      | 56        |
| 7.1.1  | Transformada discreta wavelet (DWT).....                | 56        |
| 7.2  | Extracción de características.....                      | 59        |
| 7.3  | Clasificador support vector machine.....                | 61        |
| <b>Capítulo 8. Conclusiones y Trabajos Futuros.....</b>                |   | <b>64</b> |
| 8.1  | Conclusiones .....                                      | 64        |
| 8.2  | Trabajos futuros .....                                  | 65        |
| <b>Capítulo 9. Bibliografía.....</b>                                   |   | <b>66</b> |
| <b>ANEXO I: Alineación del proyecto con los ODS.....</b>               |   | <b>72</b> |

## *Índice de figuras*

|           |   |    |
|-----------|---|----|
| Figura 1  | Representación de un electroencefalograma.....  | 7  |
| Figura 2  | Configuración de los canales 10-20 extendida y partes de la corteza cerebral .                                      | 23 |
| Figura 3  | Imágenes mostradas con expresiones de miedo y neutrales [32].   | 24 |
| Figura 4  | Canales del electroencefalograma. ....  | 25 |
| Figura 5  | Gráfico de variables de PCA de un individuo. ....   | 27 |
| Figura 6  | Componentes 2 y 3 del PCA para 3 sujetos de la prueba Ekman.....  | 28 |
| Figura 7  | PCA de 3 sujetos de La prueba Ekman.....  | 29 |
| Figura 8  | Gráfico de correlaciones de canales del EEG respecto a las componentes principales del análisis. ....               | 30 |
| Figura 9  | Histogramas, tiempos medios y desviación típica de los tiempos de respuesta de cada sujeto con altos rasgos CU..... | 33 |
| Figura 10 | Histogramas, tiempos medios y desviación típica de los tiempos de respuesta de cada sujeto con bajos rasgos CU..... | 34 |
| Figura 11 | Evento de una parte del experimento Aman.....   | 36 |
| Figura 12 | Canal AF8 del experimento Ekman.....  | 37 |
| Figura 13 | Canales Fp1, F7, AF4 y VEOG_1 del experimento Ekman.....  | 38 |
| Figura 14 | Proceso ICA para la eliminación de artificios. ....   | 40 |
| Figura 15 | Representación del proceso de filtrado de señales.....  | 41 |
| Figura 16 | Canal Fz, con tiempos de un epoch. ....   | 42 |
| Figura 17 | Epoch nº 3 y nº 10 del canal Fz. ....   | 43 |
| Figura 18 | Promedio de todos los epochs para algunos canales.....  | 45 |
| Figura 19 | Media canal Fz total, frente al estímulo S3 y S4. ....  | 45 |
| Figura 20 | Promedio de todos los sujetos para distintos canales. ....  | 46 |
| Figura 21 | Media de epoch de cada sujeto vs Media de epoch de todo el grupo para el canal Fp1.....                             | 47 |
| Figura 22 | Media de epoch de cada sujeto vs Media de epoch de todo el grupo para el canal F5.....                              | 47 |

|   |    |
|---|----|
| Figura 23 ERP para un sujeto. (Verde: media de toda la prueba; Azul: promedio de eventos S3; Rojo: promedio de eventos S4).....             | 48 |
| Figura 24 ERP del canal Fz para sujetos LCU y HCU. ....   | 49 |
| Figura 25 ERP del canal Cz para sujetos LCU y HCU.....  | 50 |
| Figura 26 ERP para varios canales de individuos HCU (rojo) y LCU (azul).....  | 51 |
| Figura 27 ERP asociado al lóbulo frontal. ....  | 52 |
| Figura 28 ERP asociado al lóbulo parietal.....  | 52 |
| Figura 29 ERP asociado al lóbulo temporal.....  | 53 |
| Figura 30 ERP asociado al lóbulo occipital. ....  | 53 |
| Figura 31 ERP de varios canales restándole el promedio del experimento. ....  | 54 |
| Figura 32 Resoluciones en tiempo y frecuencia de distintas transformaciones en comparación con una señal en el dominio del tiempo [41]..... | 57 |
| Figura 33 Descomposición de DWT, donde $\downarrow 2$ significa muestreo reducido a la mitad  | 58 |
| Figura 34 Mother wavelete para daubechies 4, y la función de escalado [45]. ....  | 59 |
| Figura 35 Ejemplo de clasificador SVM en 2-D.....   | 61 |
| Figura 36 Curva AUC-ROC de los distintos modelos de clasificación. ....   | 62 |
| Figura 37 Objetivos de desarrollo sostenible .....  | 72 |

## *Índice de tablas*

|  |    |
|--|----|
| Tabla 1 Elementos de la prueba PLC-R [6] y [11].....                             | 9  |
| Tabla 2 Tareas y recursos utilizados. ....                                       | 19 |
| Tabla 3 Mediciones para el estudio económico. ....                               | 19 |
| Tabla 4 Precio por unidad de los recursos del proyecto.....                      | 20 |
| Tabla 5 Precio por actividad y precio total del proyecto.....                    | 20 |
| Tabla 6 Tasas de respuestas de la prueba Ekman. (Naranja HCU; Azul LCU).....     | 31 |
| Tabla 7 Exactitud, sensibilidad y especificidad. ....                            | 32 |
| Tabla 8 Resumen de la tendencia de identificación errónea de cada individuo..... | 35 |
| Tabla 9 Niveles de la descomposición DWT .....                                   | 60 |

## **Capítulo 1. INTRODUCCIÓN**

El interés por el estudio del cerebro se remonta a tiempos inmemoriales, pues ya en el antiguo Egipto (año 1600 a.C) se hacían estudios anatómicos y fisiológicos del cerebro; posteriormente, en el renacimiento italiano, aparece el método científico para la exploración del sistema nervioso. Las nuevas tecnologías, como la fotografía o los rayos X tuvieron gran impacto en la medicina ya que permitían la exploración del cuerpo; sin embargo, su impacto en neurología fue menor, pues éstas no eran capaces de visualizar el cerebro.

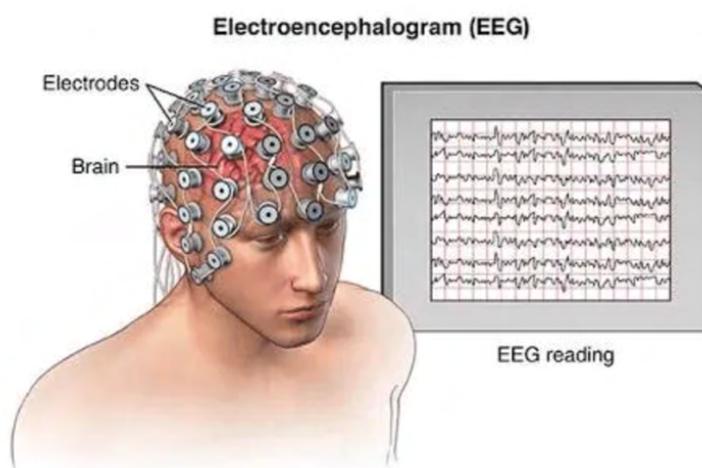
Con todo, el gran salto conceptual se produce en el último cuarto del siglo XIX, con el descubrimiento de la actividad eléctrica en el sistema nervioso, debido al estudio del sistema nervioso de animales, el cuál abrió las puertas al desarrollo de la electroencefalografía. Con esta técnica, podemos mapear la actividad eléctrica en la corteza cerebral tanto en reposo como en respuesta a un estímulo [1].

En 1920 el neurólogo alemán Hans Berger construyó un dispositivo capaz de registrar los impulsos eléctricos del cerebro. Su obra fue un claro ejemplo de la búsqueda de datos objetivos, para estudiar las bases científicas del funcionamiento cerebral. Pero no fue hasta 1930, cuando gracias a Edgar Adrián (Premio Nobel de 1932.), se popularizó el uso de electroencefalogramas como herramienta para la psiquiatría y neurología, abriendo nuevos campos de investigación en el estudio de la epilepsia y otras lesiones del cerebro. [2]

Poco a poco, los expertos empezaron a valorar el EEG como un indicador importante del funcionamiento del cerebro, y se empezó a utilizar para explorar todo tipo de fenómenos psiquiátricos, neurológicos y psicológicos [3].

El electroencefalograma es una prueba que se utiliza para estudiar el funcionamiento del sistema nervioso central, concretamente, de la actividad de la corteza del cerebro. Esta prueba realiza un seguimiento de las ondas cerebrales, utilizando electrodos fijados sobre el cuero cabelludo, que van registrando las corrientes eléctricas que se forman en las neuronas cerebrales [4].

Un electroencefalograma es capaz de detectar cambios en la actividad cerebral, es decir en el potencial eléctrico del cerebro. Estos cambios pueden ser útiles para diagnosticar trastornos cerebrales, especialmente epilepsia u otros trastornos convulsivos [5]. Un electroencefalograma también puede ser útil para diagnosticar o tratar otros trastornos, como tumores cerebrales, daños por lesiones en la cabeza, encefalopatía. En la Figura 1 se puede ver una representación de un electroencefalograma, y las señales que se obtienen de este.



*Figura 1 Representación de un electroencefalograma.*

En vistas del potencial del EEG para el estudio de la actividad cerebral, en este proyecto se propone la creación de una herramienta que permita el procesado de señales. Además, del análisis exploratorio de señales tomadas a adolescentes con antecedentes delictivos. Finalmente, la utilización de herramientas de Aprendizaje Automático para la detección de perfiles psicopáticos mediante el análisis y clasificación de las señales.

Los perfiles psicopáticos han llamado la atención ya desde 250 a.C, donde Teofrasto, alumno de Aristóteles, describe hombres sin escrúpulos que realizan acciones de mucha violencia [6]. Destacaba de estas personas su inteligencia, y capacidad de comprensión frente a las acciones que cometían. Esta idea se clasificaba como una enfermedad debida a causas sobrenaturales.

El concepto se mantiene hasta que en el siglo XIX, Phillipe Pinel utiliza el término de “*manie sans délire*”, o manía sin delirio. Describe la psicopatía como sujetos con las capacidades intelectuales intactas, pero dominados por la violencia y con las funciones afectivas alteradas [7]. Más tarde, James Cowel Prichard, usa el término “*Moral Insanity*”, destacando la incapacidad de los individuos de seguir las normas.

El término de psicopatía es acuñado por primera vez por Julius Koch cuando propone el término “*inferioridad psicopática constitucional*”. Con este término se refiere a la base biológica de esta anomalía psicológica [8]. Emil Kraepelin siguió con esta idea, y fue ampliando y matizando su concepción. Describe el término de “*Personalidades psicopáticas*”, con la que hace referencia a las diferentes formas en las que se desarrolla la personalidad consideradas degenerativas, es decir, se distinguen distintos tipos de psicópatas. Más tarde, Kurt Schneider avanza sobre la idea de que estos individuos tienen un trastorno de personalidad, por el cual sufren o hacen sufrir a los demás. Destaca diez tipos diferentes [8].

Posteriormente, se realizan importantes avances gracias a Hervey Cleckley con su libro *The mask of Sanity*. En este se realiza una descripción fenomenológica interna y externa de los psicópatas, indicando que los psicópatas tienen las capacidades intelectuales intactas, pero afecciones en el desarrollo de la personalidad y conducta. Es decir, dicen unas cosas pero realizan otras; además indica que son capaces de imitar la moral y sentimientos, pero no tienen emociones asociadas a ella. También destaca la diferenciación que realiza entre criminales y psicópatas. Establece una serie de criterios para su diagnóstico, describe a los psicópatas como personas con “*encanto superficial y buena inteligencia, ausencia de delirios, poco fiable, falta de sinceridad, falta de remordimiento, conducta antisocial sin motivos, dificultad para aprender de las experiencias, incapacidad para amar y insensibilidad en las relaciones interpersonales*” [9]. Tras esto, Robert Hare desarrolla, usando estas características, un instrumento para el diagnóstico de la psicopatía llamado Psychopathy Checklist (PCL).

Actualmente, se utiliza el instrumento de Hare revisado (PCL-R) para identificar el riesgo de conducta criminal de los sujetos, con un gran capacidad predictiva [10]. Estos miden distintas dimensiones del sujeto, la dimensión interpersonal, la dimensión afectiva, el estilo de vida y la dimensión antisocial. En la Tabla 1, podemos ver qué elementos se encuentran en cada dimensión:

|  |  |
|--|--|
| Factor 1: mide aspectos de la personalidad clásica del psicópata.  | Factor 2: mide la conducta antisocial y la impulsividad.   |
| Dimensión 1: Interpersonal.  | Dimensión 3: Estilo de vida.   |
| <ol style="list-style-type: none"> <li>1. Encanto superficial</li> <li>2. Sentimiento de grandiosidad / egocentrismo</li> <li>3. Mentira patológica</li> <li>4. Manipulación</li> </ol>  | <ol style="list-style-type: none"> <li>1. Necesidad de estimulación/ tendencia al aburrimiento</li> <li>2. Estilo de vida parasitario</li> <li>3. Falta de metas a largo plazo</li> <li>4. Impulsividad</li> <li>5. Irresponsabilidad</li> </ol> |
| Dimensión 2: Afectiva  | Dimensión 4: Antisocial  |
| <ol style="list-style-type: none"> <li>1. Falta de sentimiento de culpa y de arrepentimiento</li> <li>2. Emociones superficiales</li> <li>3. Insensibilidad/ Falta de empatía</li> <li>4. Incapaz de reconocer la personalidad de sus actos</li> </ol> | <ol style="list-style-type: none"> <li>1. Falta de control conductual</li> <li>2. Delincuencia juvenil</li> <li>3. Revocación de la condicional</li> <li>4. Versatilidad criminal</li> </ol>   |
| Otros elementos que no pertenecen a ningún factor  |  |
| <ol style="list-style-type: none"> <li>1. Conducta sexual promiscua</li> <li>2. Varias relaciones maritales breves</li> </ol>  |  |

*Tabla 1 Elementos de la prueba PLC-R [6] y [11].*

Mirando estos rasgos, y las distintas puntuaciones que se pueden obtener, se distinguen 3 grupos distintos [6]:

- Clásico: altos resultados en todas las dimensiones.
- Manipulador: altos resultados en las dimensiones interpersonales y afectivas.
- “Macho”: bajos resultados en la dimensión interpersonal, y alta en resto.

También se utiliza el DSM-V, que es un instrumento que se usa para distinguir distintos trastornos mentales, donde la psicopatía sería un trastorno antisocial de la personalidad. El inconveniente, es que este instrumento excluye los rasgos interpersonales y afectivos, por lo que la mayoría de los delincuentes podrían ser clasificados con este trastorno. Por ello, existe controversia en la idea que la psicopatía y el trastorno antisocial de personalidad sean sinónimos [11]

Como se puede ver, existen distintos instrumentos que sirven para el diagnóstico de la psicopatía, pero estas presentan algunas dificultades. Estas se deben al grado de confianza sobre la información obtenida, puesto que las pruebas se tratan de cuestionarios autoadministrados y con escalas de evaluación clínico conductual [12].

Por ello, la principal motivación de este proyecto, es crear modelos que puedan ayudar a la ciencia médica en el diagnóstico de trastornos mentales, en concreto, de las psicopatías, mediante el análisis y clasificación de las señales obtenidas de adolescentes delictivos.

Como Ingeniero, considero de gran valor que la tecnología pueda ayudar a la ciencia biomédica a la detección y diagnóstico de enfermedades que resultan muy perjudiciales a la sociedad. Considero que todo lo que se pueda hacer para investigar, adelantar el diagnóstico e identificar y tratar esas conductas tan dañinas para la sociedad, es muy beneficioso y supondría un gran paso adelante, en el que la ingeniería puede estar presente y tomar parte.

La memoria de este proyecto seguirá la siguiente estructura: Primero se describirán las tecnologías utilizadas. Tras esto, se expondrá el estado de la cuestión donde se discuten algunas soluciones y aplicaciones del aprendizaje automático junto con EEG. Después, se

explicará en análisis exploratorio que se ha realizado, y el procesado de las señales. Tras esto se realizara un análisis de los ERP, y un modelo de clasificación para clasificar los individuos.

## **Capítulo 2. ESTADO DE LA CUESTIÓN**

En este capítulo, se revisarán proyectos que tratan el uso del aprendizaje automático para la clasificación de señales de electroencefalogramas. Estos proyectos tratan temas como la detección de enfermedades patológicas tales como la epilepsia, la detección de un estímulo visual, el uso de EEG para el control de máquinas o la clasificación del estado emocional. Y finalmente, se verá la historia del uso de electroencefalogramas en la clasificación de perfiles psicopáticos.

### ***2.1 APLICACIONES DEL APRENDIZAJE AUTOMÁTICO PARA LA CLASIFICACIÓN DE SEÑALES DE EEG***

#### **2.1.1 CLASIFICACIÓN SEGÚN LA VISIÓN**

Una de las posibles aplicaciones que hay que destacar es el uso del aprendizaje automático y de EEG en la neurociencia cognitiva, por ejemplo, en la velocidad de procesamiento de una imagen en los humanos, en la cual normalmente se detecta mirando la diferencia de potencial relacionado con un evento (ERP) en un EEG. El inconveniente de usar el ERP es que no se muestra la totalidad de la información que se obtiene en un EEG. Por ello la literatura [13], propone utilizar el aprendizaje automático para poder detectar cuándo un individuo está recibiendo un estímulo visual y utilizar más información del EEG. Para ello, realizan un experimento en el cual van clasificando los distintos EEG (“objeto presente” o “objeto ausente”). Tras esto, entrenan un clasificador SVM con esa información, y estudian la precisión del modelo. Finalmente, obtienen un clasificador capaz de clasificar los EEG con una precisión del 87%.

#### **2.1.2 CLASIFICACIÓN DE MOVIMIENTOS DE LA MANO D/I A TRAVÉS DE EEG**

En la literatura [14], se propone el uso del aprendizaje automático y EEG para poder clasificar los movimientos de la mano derecha o izquierda. La idea es crear un modelo que

permita su uso con un *Brain Computer Interface* (BCI), que es un dispositivo que permite el control de máquinas a través de la actividad cerebral, por ejemplo, para robots o extremidades artificiales. Para ello desarrollan distintos métodos de extracción de características, y clasificadores: redes neuronales y SVM.

### **2.1.3 DIAGNÓSTICO DE PATOLOGÍAS DETECTABLES CON EEG**

En el estudio de las enfermedades cerebrales se debe diferenciar entre patologías físicas y patologías mentales. El EEG se utiliza para detectar gran variedad de patologías. Por un lado, enfermedades físicas; la epilepsia, daños cerebrales, infartos cerebrales etc. y por otro lado, puede ayudar en el diagnóstico de enfermedades mentales: demencia, depresión, Alzheimer, que tienen un diagnóstico clínico. El uso del aprendizaje automático para la detección de estas patologías permitiría acelerar y automatizar el proceso de diagnóstico. En este estudio [15], se comparan distintos modelos de clasificación para dividir los pacientes en dos grupos, aquellos con patologías y sin patologías; con estos modelos se consiguen buenos resultados. El artículo concluye que hay bastante futuro en el uso del aprendizaje automático en el análisis de EEG para la detección de distintas enfermedades.

### **2.1.4 DIAGNÓSTICO DE LA EPILEPSIA**

La epilepsia no es una enfermedad psiquiátrica ni mental, se trata de un problema físico causado por un funcionamiento anormal esporádico de un grupo de neuronas. En el cual los pacientes sufren convulsiones, que se deben a descargas eléctricas excesivas de algunos grupos de células cerebrales en diferentes partes del cerebro. Las características de las convulsiones varían y dependen de en qué parte del cerebro comienza la alteración y cómo se propaga [16].

Como se comentó en el apartado anterior, el uso de EEG es muy importante en el diagnóstico de la epilepsia. Las convulsiones que causan la enfermedad pueden provocar que los pacientes sufran heridas. Por ello es muy importante poder detectar cuándo un paciente va a sufrir un ataque epiléptico y adelantarse. En el artículo [17], se estudia un dispositivo capaz de detectar cuando se va a sufrir un ataque epiléptico utilizando el aprendizaje automático

para procesar la información obtenida de un EEG. Para ello, deciden resolver un problema de clasificación binario, en el cual los EEG serán clasificados en periodos, donde no hay ataque epiléptico y en los que sí. En este estudio utilizan un clasificador SVM, con el cual son capaces de clasificar los EEG con una latencia de 4.2 segundos, y una alta precisión.

### **2.1.5 CLASIFICACIÓN DEL ESTADO EMOCIONAL**

Las emociones son un factor muy importante en la interacción entre personas y máquinas, por ello la literatura [18], propone un método de clasificación del estado emocional. Entre los distintos métodos que existen actualmente, están: la clasificación analizando las expresiones faciales, la clasificación de las señales fisiológicas periférica y la clasificación de las señales de los EEG. La literatura destaca las principales dificultades de la clasificación de las señales del EEG. Por un lado, la necesidad de grandes ventanas de tiempo para identificar el estado emocional, y por otro lado la poca actividad cerebral que se captura. Esto se debe a que no tienen en cuenta la naturaleza de los EEG frente a las emociones, es decir, que las características que se extraen de las señales son cambios instantáneos, y sin embargo las emociones cambian de forma gradual. Por esto, la literatura propone un método que permita seguir la trayectoria de las emociones utilizando un método de suavizado de las características y un clasificador SVM.

### **2.1.6 PREDICCIÓN DE LA EDAD USANDO EEG**

En la literatura [19], se busca poder predecir la edad de los distintos sujetos en función de sus señales cerebrales. La idea era combinar el aprendizaje automático junto con los EEG para poder predecir la edad de los sujetos, consiguiendo al final unos resultados que proporcionaban unas predicciones fiables.

## **2.2 EEG EN LA PSICOLOGÍA Y CRIMINOLOGÍA**

A diferencia de la epilepsia, que es una enfermedad física, las psicópatas son enfermedades psiquiátricas y mentales. La psicopatía es difícil de diagnosticar, puesto que es un diagnóstico clínico que cuesta hacer y se trata de ver a través de un patrón de

comportamiento, de ofensas delictivas, “no hay marcadores o pruebas físicas o biológicas que definan los problemas de salud mental” [20]; el tener un instrumento que no sean meras apreciaciones, que apoye el diagnóstico sería muy útil, de ahí la gran importancia de los EEG, que podrían facilitar e incluso adelantar el diagnóstico, siendo además una prueba nada invasiva. .

Desde 1938 se pueden ver los primeros estudios sobre análisis de EEG y el comportamiento de los niños [21], en este el autor indica que las personas desviadas tienen ondas cerebrales desviadas. Todo esto propició al uso de EEG como una herramienta prometedora, que facilitaría la clasificación de los criminales, individuos antisociales y personas violentas de una forma más objetiva, como se puede ver en [22] y [23]. También se propuso que en un futuro se pudiese usar como herramienta en la criminología. Pero uno de los principales problemas que tuvo los EEG en estos campos, era la dificultad para interpretar los datos, y procesar la gran cantidad de datos. Esto se ha ido solucionando con la mejora de los estudios y la aparición de nuevas tecnologías que facilitan el procesamiento automático de señales.

En el artículo [24], se discute el uso del aprendizaje automático para el diagnóstico de personalidad antisocial. Para ello usaron las señales de EEG de 84 sujetos en reposo. Como clasificador utilizan un *Random forest* con la que obtienen una exactitud del 77.5%

Asimismo, se han empezado a implementar las nuevas tecnologías junto con el aprendizaje automático, para distintos proyectos en estos campos. Por ejemplo, para la detección de pacientes con depresión [25], para clasificar individuos con trastorno de apetito desenfrenado [26], para detectar las emociones a partir de un EEG [27] y para clasificar individuos con psicopatologías.

## **Capítulo 3. DEFINICIÓN DEL TRABAJO**

En este capítulo, se justificará la razón por la que se realiza este proyecto, y su utilidad. Tras esto, se establecerán cuales son los objetivos. Además, se explicará la metodología que se ha seguido, y finalmente los recursos que se han empleado.

### **3.1 JUSTIFICACIÓN**

Como se puede ver en el Capítulo 1. , los instrumentos que se utilizan actualmente para el diagnóstico de la psicopatía, presentan algunas dificultades. Principalmente, el grado de confianza de la información obtenida, puesto que las pruebas suelen ser cuestionarios auto administrados. Es decir, como se ve en [20], su diagnóstico es costoso *“no hay marcadores o pruebas físicas o biológicas que definan los problemas de salud mental”*.

El electroencefalograma es una prueba que se utiliza para estudiar el funcionamiento del sistema nervioso central, concretamente, de la actividad de la corteza del cerebro de forma no invasiva. El principal inconveniente de esta tecnología, es la dificultad de procesamiento y análisis de las señales, pues requieren mucha especialización.

Por todo esto, lo anterior, en este proyecto se propone crear una herramienta de análisis de encefalogramas. Además la utilización de herramientas de Aprendizaje Automático para la detección de perfiles psicopáticos de señales tomadas a adolescentes con antecedentes delictivos. El tener un instrumento que apoye el diagnóstico basado en cuestionarios, permitiría adelantar y facilitar el diagnóstico de esta enfermedad, lo que ayudaría a tratar y dar la ayuda necesaria a estas personas. También se plantea su uso como herramienta con potencial forense y su posible aplicación en investigaciones policiales.

## **3.2 OBJETIVOS**

El principal objetivo de este proyecto es analizar y clasificar las señales tomadas de adolescentes delictivos con el uso del aprendizaje automático. Para cumplir con este objetivo se desarrollarán una serie de objetivos secundarios:

- Crear una herramienta integrada de análisis de encefalogramas que permita visualizar y analizar dichos datos.
- Analizar y clasificar las señales tomadas de adolescentes delictivos con el uso del aprendizaje automático.
- Establecer una jerarquía de los algoritmos más eficaces para este tipo de análisis

## **3.3 METODOLOGÍA**

Para la realización de este proyecto, se han establecido una serie de objetivos, y para poder completarlos se han ido realizando una serie de tareas menores:

En primer lugar, se investigó el concepto de psicopatía y sus formas de diagnóstico. Tras esto, se estudió cómo es la estructura, y las distintas componentes de las señales que se obtienen de los EEG. Una vez se comprendieron estos dos conceptos, se procedió a realizar un estudio de posibles soluciones que trataran el uso de EEG y el aprendizaje automático.

En segundo lugar, se procedió a realizar un análisis exploratorio de los datos que se iban a utilizar en el proyecto con el fin de comprender mejor las pruebas a las que se sometieron a los sujetos y cómo fueron los resultados de estas. Tras esto, se intentó identificar alguna forma de clasificación de los sujetos en base a sus respuestas a las pruebas.

En tercer lugar, se procedió al pre-procesados de las señales, que incluye la detección de artificios, su filtrado, y la extracción de epochs. Tras esto, se usaron estas señales para 2 estudios diferentes: el estudio de los potenciales relacionados con eventos, y la clasificación de los individuos en base a estos; la extracción de características de las señales y la clasificación de estas.

Finalmente, se procedió a la elaboración de la memoria y las conclusiones del proyecto.

### **3.4 RECURSOS A EMPLEAR**

Para la elaboración de este proyecto se han usado diversos recursos. En el estudio de literatura se ha usado Google Scholar, ScienceDirect y IEEExplore. En el desarrollo del software para el análisis, procesado y clasificación de las señales se ha usado el entorno de programación Rstudio. Este es un entorno de programación de código abierto basado en R para la computación estadística [28]. Además, se han utilizado diversas librerías:

- Eeguana: para la visualización y la extracción de datos recogidos con el encefalograma [29].
- Eegkit: para el procesado de datos de EEG, como el filtrado de las señales [30].
- fastICA: para realizar distintos análisis PCA,FAMD,ICA...
- Wavelets: para la obtención de la transformada wavelet de las señales.
- Officer: usado en la creación de informes.
- ggplot: usado en la visualización de gráficas.
- tidyr, dplyr: para la organización y uso de los datos.
- caret: modelos de clasificación.

En la Tabla 2 podemos ver las tareas que se han ido realizando y los recursos empleados.

| <b>TAREAS Y RECURSOS UTILIZADOS</b> |                                    |                   |
|-------------------------------------|------------------------------------|-------------------|
| <b>NUMERO</b>                       | <b>TAREA</b>                       | <b>RECURSO</b>    |
| <b>1</b>                            | <b>Estudio de concepto</b>         |                   |
| 1.1                                 | Estudio de la psicopatía           | Google Scholar    |
| 1.2                                 | Estudio de los EEG                 | Google Scholar    |
| <b>2.</b>                           | <b>Análisis exploratorio</b>       |                   |
| 2.1                                 | Análisis de las señales de EEG     | RStudio (eeguana) |
| 2.2                                 | Análisis de las pruebas realizadas | Rstudio           |
| 2.3                                 | Análisis de los resultados         | Rstudio (ggplot)  |
| <b>3.</b>                           | <b>Procesado de las señales</b>    |                   |
| 3.1                                 | Detección de artificios            | Rstudio (fastICA) |
| 3.2                                 | Filtrado de las señales            | Rstudio (eegkit)  |
| 3.3                                 | Segmentación/Epoching              | Rstudio           |

|           |   |                         |
|-----------|---|-------------------------|
| <b>4.</b> | <b>ERP</b>                                  |                         |
| 4.1       | Obtención de ERP                            | Rstudio                 |
| 4.2       | Estudio de ERP                              | Rstudio (ggplot)        |
| <b>5.</b> | <b>Extracción de características</b>        |                         |
| 5.1       | Uso de transformada wavelet                 | Rstudio (wavelets)      |
| 5.2       | Estudio de características obtenidas        | Rstudio (SciView)       |
| 5.3       | Clasificación en función de características | Rstudio (caret)         |
| <b>6.</b> | <b>Redacción de la memoria</b>              | Microsoft office (Word) |

*Tabla 2 Tareas y recursos utilizados.*

### **3.5 ESTUDIO ECONÓMICO**

En este apartado se realiza un estudio económico del proyecto, teniendo en cuenta el software utilizado y la mano de obra.

#### **3.5.1 MEDICIONES**

En la Tabla 3 se pueden ver la frecuencia de cada actividad o recurso utilizado durante el proyecto.

| <b>Componente</b>             | <b>Frecuencia</b> | <b>Unidades</b> |
|-------------------------------|-------------------|-----------------|
| <b>Licencia Rstudio</b>       | 1                 | u.d             |
| <b>Investigación</b>          | 100               | horas           |
| <b>Análisis de datos</b>      | 200               | horas           |
| <b>Elaboración de informe</b> | 60                | horas           |

*Tabla 3 Mediciones para el estudio económico.*

### 3.5.2 PRECIO UNITARIO

En la Tabla 4 se puede ver el precio por unidad de cada actividad o recurso utilizado en el proyecto. El precio de mano de obra se ha sacado como el precio de un analista de datos sin experiencia [31].

| <b>Componente</b>             | <b>Precio/ud</b> |
|-------------------------------|------------------|
| <b>Licencia Rstudio</b>       | 0 €/ud           |
| <b>Investigación</b>          | 15 €/h           |
| <b>Análisis de datos</b>      | 20 €/h           |
| <b>Elaboración de informe</b> | 15 €/h           |

*Tabla 4 Precio por unidad de los recursos del proyecto.*

### 3.5.3 SUMA TOTAL

En la Tabla 5 se puede ver el precio de cada actividad o recurso, y el precio total del proyecto.

| <b>Componente</b>             | <b>Precio</b> |
|-------------------------------|---------------|
| <b>Licencia Rstudio</b>       | - €           |
| <b>Investigación</b>          | 1.500 €       |
| <b>Análisis de datos</b>      | 4.000 €       |
| <b>Elaboración de informe</b> | 900 €         |
| <b>Total</b>                  | 6.400 €       |

*Tabla 5 Precio por actividad y precio total del proyecto.*

## **Capítulo 4. ANÁLISIS Y ESTRUCTURA DE LOS DATOS**

En este capítulo se explicará en primer lugar la estructura de los datos con los que se ha trabajado. Tras esto, se comentará el análisis exploratorio que se ha realizado para obtener un mayor entendimiento de los datos, y de cómo proceder en el resto del proyecto.

### **4.1 ORIGEN DE LOS DATOS**

Los datos que se utilizan en el proyecto [32] fueron recogidos por la Dra. Lucía Halty, la Facultad de Ciencias Humanas y Sociales de la Universidad Pontificia Comillas. El estudio recogió datos de 48 jóvenes que estaban en reformatorios en la Comunidad de Madrid. Los jóvenes son todos varones, tienen entre 15 y 20 años, y provienen de España, Sudamérica, el Norte de África y Europa del Este. Se encontraban en los reformatorios debido a crímenes relacionados con robo agravado, pertenencia a bandas, homicidio, agresión sexual y maltrato a familiares, entre otros.

Para este proyecto se utilizan concretamente los datos de 14 jóvenes del total. Los de aquellos cuyos resultados en la prueba *“Inventory of Callous–Unemotional traits”* han sido más altos. Esta prueba consiste en un cuestionario de 24 ítems cuyo objetivo es evaluar los rasgos de insensibilidad y falta de emociones (CU) en niños y adolescente. Estos rasgos definen unos sujetos con falta de empatía y remordimiento, que provoca que tengan problemas graves de conducta. Por un lado, se estudian 7 sujetos cuyos resultados mostraban la presencia más alta de rasgos *“Callous-Unemotional”*, a este grupo se referirá como HCU, y por otro lado otros 7 sujetos que tenían baja presencia de estas características, a los que se refiere como LCU.

Para la toma de datos se utilizó un electroencefalograma de 64 canales en una configuración extendida 10-20, que se puede ver en Figura 2. La actividad se grabó utilizando el programa Brainvision. La actividad electrocular se midió en el borde derecho de la órbita del ojo,

detectando el movimiento vertical y horizontal de estos. Los datos del EEG se tomaron con una frecuencia de muestreo de 500Hz.

La colocación de los electrodos se realiza de manera que se obtengan las variaciones de potencial alrededor de toda la corteza cerebral. Como se ve en la Figura 2 cada canal está ubicado de forma que identifica los cambios de potencial de distintas zonas del cerebro. Por ello la nomenclatura de los electrodos variará en función de su ubicación. Las letras indican el lóbulo en el que se ubicará el electrodo, y el número el hemisferio. Siendo los números impares (1,3,5,7) referencias al hemisferio izquierdo y los pares (2,4,6,8) al hemisferio derecho. Cada lóbulo de la corteza cerebral se encarga de distintas tareas [33]:

- Lóbulo frontal: es la parte más grande del cerebro. Está relacionada con el lenguaje y movimiento, con el control de las emociones y el razonamiento. Es detectada por los canales que contiene o empiezan la letra Fp (zona frontopolar) o F(frontal).
- Lóbulo parietal: se encarga de procesar la información sensorial del cerebro. Es detectada por los canales que contienen o empiezan por C(central) o P(parietal).
- Lóbulo occipital: se encarga de la visión. Es detectada por los canales que empiezan o contienen la O (occipital).
- Lóbulo temporal: se relaciona con la memoria, identificación de rostros y procesamiento de estímulos auditivos. Es detectada por los canales que contienen o empiezan por T(temporal) [34].

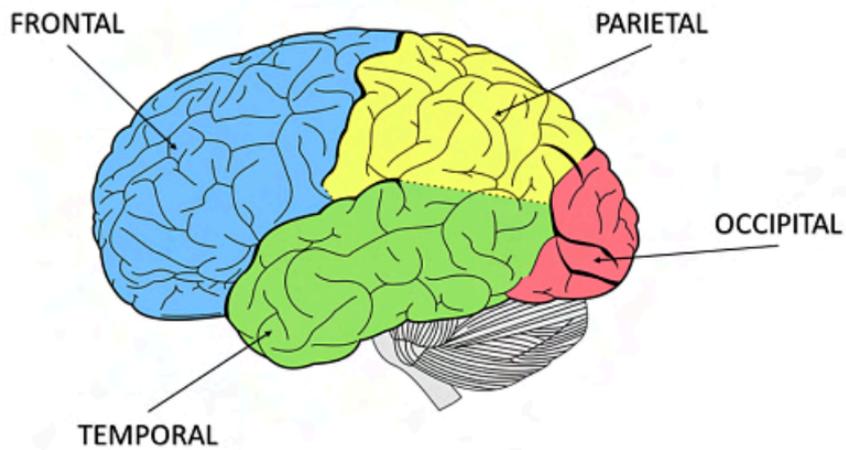
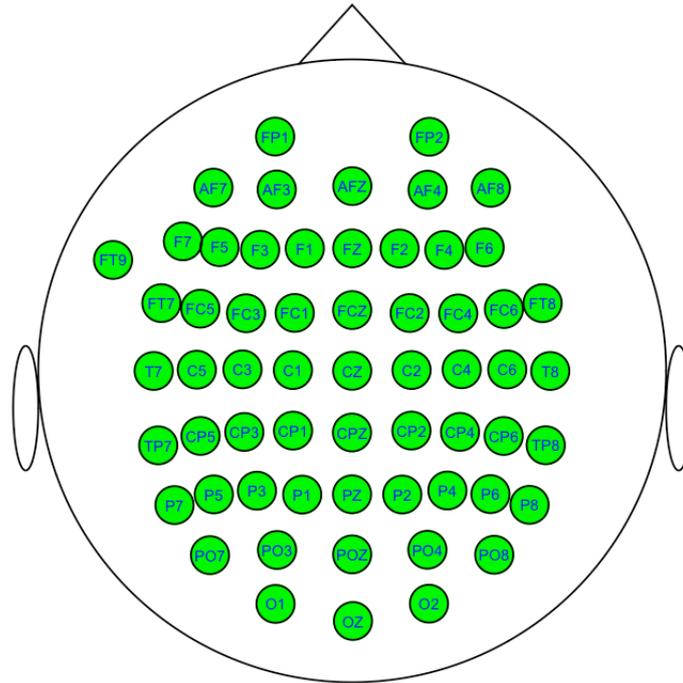
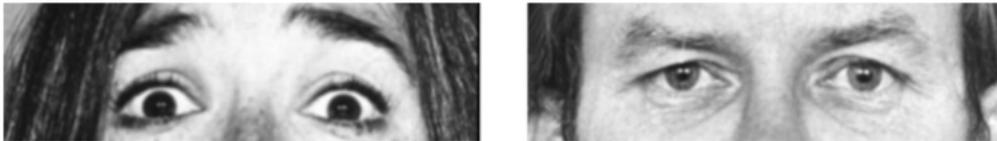


Figura 2 Configuración de los canales 10-20 extendida y partes de la corteza cerebral .

A los sujetos se le realizaron 3 tipos de pruebas diferentes, siendo la primera de ellas la que se estudia más en profundidad en este proyecto:

- Prueba EKMAN: la prueba consistió en visualizar dos tipos de expresiones. Por un lado, 10 imágenes que mostraban una expresión en los ojos neutral, y por otro lado, 10 imágenes que mostraban miedo. En la Figura 3 podemos ver un ejemplo del tipo de imágenes que se utilizaron. En total se presentaron unas 80 imágenes, donde 40 expresaban emociones neutras y las otras 40 miedo. Cada uno de estos estímulos se presentaba durante 500ms, y con un intervalo entre estímulos, de 1500 ms, siendo la duración total de la prueba de 2.25 min. Los sujetos debían presionar “M” cuando identificaban una expresión de miedo y “Z” cuando identificaban una expresión neutral.



*Figura 3 Imágenes mostradas con expresiones de miedo y neutras [32].*

- Go/No Go: a los sujetos se les enseñaban 3 letras M, N o W. Debían presionar el botón Go al ver las letras M y N, y el botón no Go al ver la letra W
- Nim: en este caso a los sujetos se les enseñaban una imagen con expresión neutra o de miedo durante unos milisegundos, es decir, no le daba tiempo a procesarla. Y tras esto se les enseñaba unos números y debían identificar si eran par o impar.

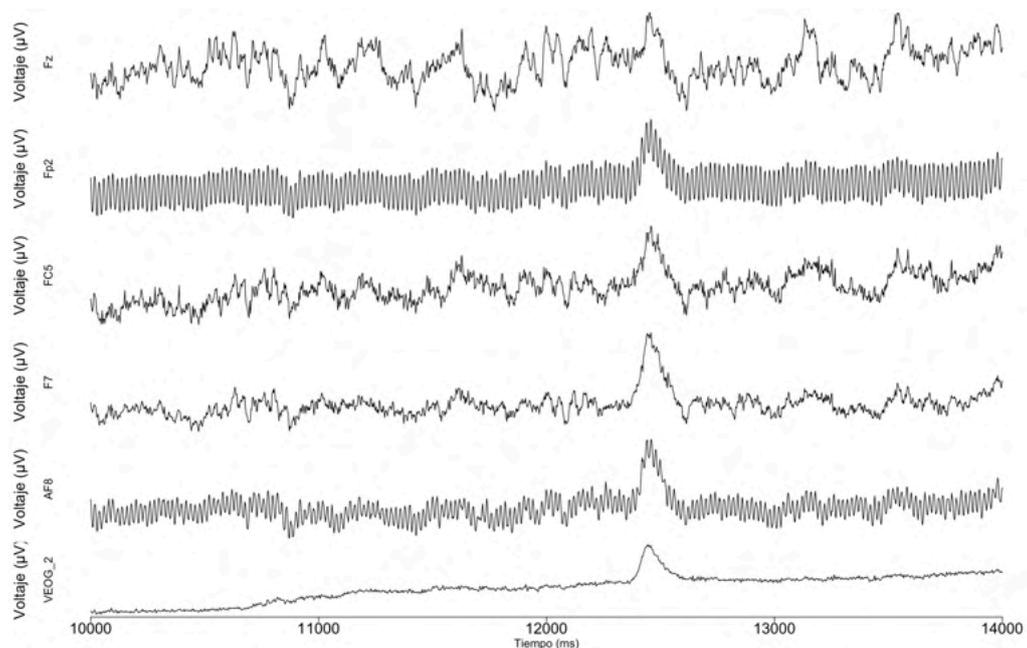
## **4.2 ESTRUCTURA DE LOS DATOS**

En este apartado se explica la estructura de los datos con los que se va a trabajar. Los datos se recogen con el programa Brain Vision Analyzer, que recoge en cada prueba 3 archivos diferentes, con distintas extensiones. Por un lado tenemos, “datos.vhdr” que contiene los metadatos, y enlaces a los otros dos archivos. Luego “datos.vmrk” que contiene los eventos que se producen en la prueba. Finalmente “datos.eeg” que contienen la señal en cada muestra

obtenida. Para leer y trabajar con estos datos en Rstudio, se utiliza la librería *eeeguna*. En especial la función `read_vhdr()`, que es capaz de leer los datos, y devuelve una lista formado por distintos *dataframes*.

Una vez comprobado que la librería funcionaba correctamente, se procedió a leer los datos de los 14 sujetos, para cada uno de las pruebas. Se creó una lista por individuo que contenía los datos de las tres pruebas (Ekman, Go/NoGo, y NIM). Por otro lado, cada una de estas pruebas estaba formada por 3 *dataframes* diferentes. Una de ellas que contenía la información sobre los estímulos mostrados durante el experimento y recogía las respuestas de los individuos. Otra, estaba formada por 66 variables, que incluían los 64 canales del EEG, una variable identificadora y una variable temporal. Cada una de estas variables tenía entre 70.000-80.000 observaciones. Por último, una lista que contenía información general de la prueba del EEG.

Mas allá para comprender cómo eran los datos que se iban a utilizar, se procede a representarlos. En la Figura 4 podemos ver un ejemplo de cómo serían las señales de 6 canales de un sujeto para una de las pruebas.



*Figura 4 Canales del electroencefalograma.*

### **4.3 ANÁLISIS EXPLORATORIO DE LOS DATOS**

En este apartado se explican los distintos análisis que se han hecho sobre el conjunto de datos originales, para su estudio y comprensión, como paso previo al procesado de señales, y la obtención de conclusiones.

Para el análisis se han realizado distintos análisis principales de componentes (PCA). PCA es un método estadístico que sirve para encontrar las transformaciones lineales que permitan reducir las dimensiones de los datos. Si se tiene un vector de variables  $[x_1, x_2, \dots, x_n]$ , el objetivo es encontrar una matriz ortonormal  $V$  de tamaño  $k \times n$ ,  $k < n$  tal que la proyección reducida  $x' = Vx$  pierda la menor cantidad de información posible utilizando para ello la matriz de covarianza [35].

#### **4.3.1 ANÁLISIS DE LAS DISTINTAS PRUEBAS**

En este apartado del análisis exploratorio, se realiza un análisis principal de componentes (PCA), sobre los datos de todas las pruebas de un sujeto. Esto se realiza para ver si las pruebas son lo suficientemente distinguibles entre si, y si se podrían clasificar los datos de un individuo en función de la prueba que realizó. En la Figura 5 se puede ver un gráfico de variables, para un sujeto. Se puede apreciar que los canales de cada prueba están correlacionados entre si, y son distinguibles de aquellos canales de otras pruebas. Las variables en azul muestran los canales de la prueba Nim, las variables rojas los canales de la prueba Go, y los variables verdes los canales de la prueba Ekman.

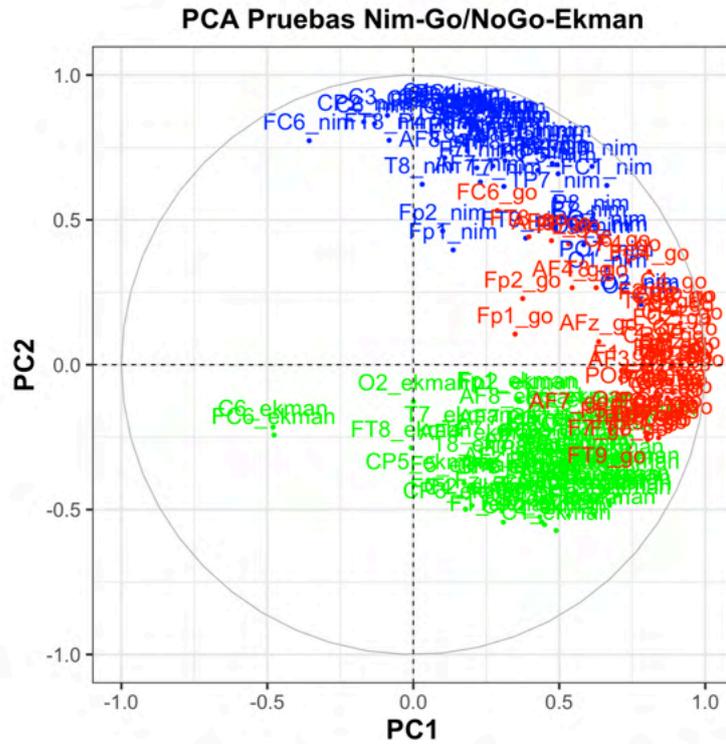


Figura 5 Gráfico de variables de PCA de un individuo.

#### 4.3.2 ANÁLISIS DE SUJETOS

En este apartado del análisis exploratorio, se realiza un análisis principal de componentes, utilizando como datos los 64 canales del electroencefalograma de 3 sujetos diferentes de la prueba Ekman.

En la Figura 6 se puede ver un gráfico de variables de las componentes principales 2 y 3 del análisis. En azul se tienen las señales de un individuo, en rojo el de otro y en verde los del tercero. Se observa que los datos para cada individuo son bastante distinguibles. Con esto se concluyó que los cambios en el potencial eléctrico de cada individuo son muy personalizados, y por tanto, el valor la actividad eléctrica del cerebro también es muy específica para cada sujeto.

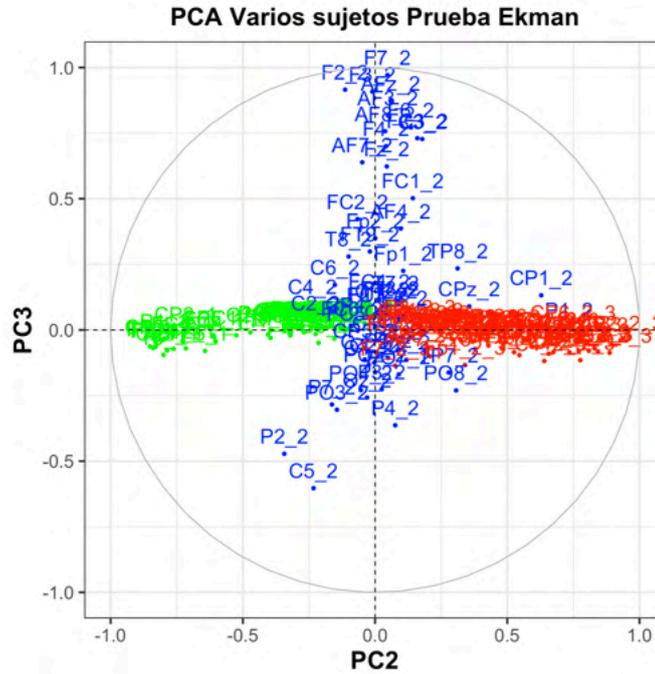


Figura 6 Componentes 2 y 3 del PCA para 3 sujetos de la prueba Ekman.

En la Figura 7 se ve el mismo grafico de variables que el anterior, pero en este caso, se han coloreado las variables en función del resultado de psicopatía que obtuvo el sujeto. En azul, los sujetos que dieron alto grado rasgos HCU, y en rojo un sujeto que obtuvo bajos resultados LCU. Se puede apreciar que con PCA es factible distinguir los distintos individuos, pero no es fácil clasificarlos en función de su resultado en la prueba de rasgos *Callous-unemotional* con los datos directamente sacados del EEG. Esto se debe a que el valor de las señales es muy diferente según los sujetos, y no presentan relación a primera vista en función de sus resultados a la prueba.

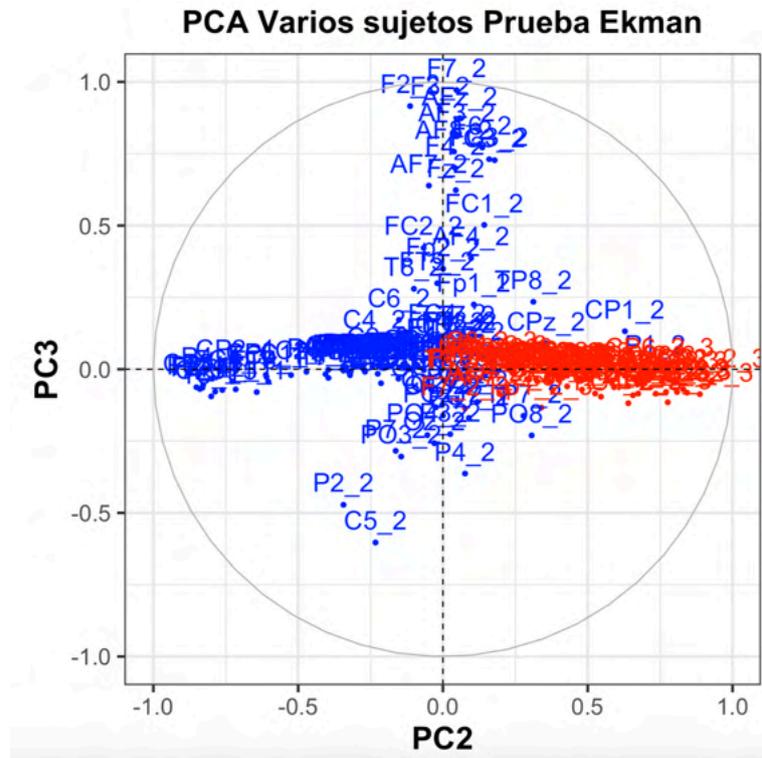


Figura 7 PCA de 3 sujetos de La prueba Ekman

### 4.3.3 REDUCCIÓN DE DIMENSIONES CON PCA

En este apartado del análisis exploratorio, se realizó un análisis principal de componentes utilizando 62 de los 64 canales posibles de un individuo. Se descartaron los canales que seguían el parpadeo de los ojos, puesto que distorsionaban las señales. En la Figura 8 se puede ver la correlación de cada canal con las componentes principales del análisis y se observa que las señales de la prueba están muy correlacionadas entre si.

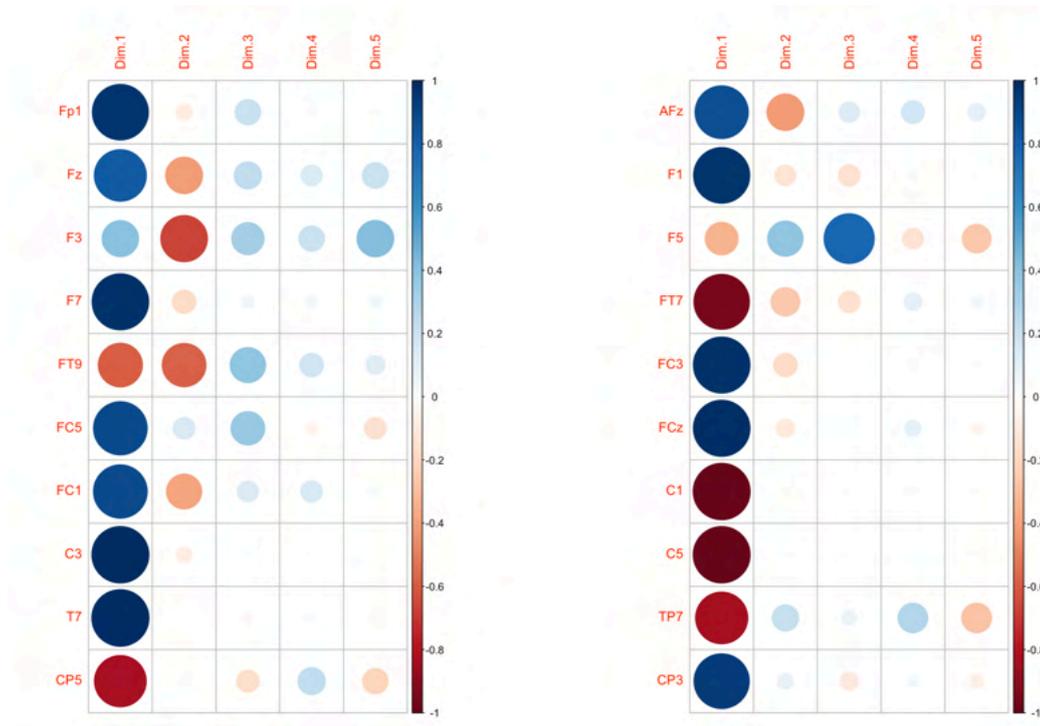


Figura 8 Gráfico de correlaciones de canales del EEG respecto a las componentes principales del análisis.

#### 4.4 ANÁLISIS DE LOS RESULTADOS DE LA PRUEBA EKMAN

En este apartado se ha estudiado tanto la naturaleza como los tiempos de las respuestas de los distintos individuos a los eventos neutrales y de miedo que se presentaron en la prueba Ekman.

En primer lugar, se observó el número de aciertos, errores y respuestas no contestadas de cada sujeto a los eventos de la prueba. Esto se hizo para ver si había alguna diferencia notable entre los sujetos que presentaban alto grado de rasgos HCU, y aquellos que no la presentaban, puesto que en teoría, las personas que presentan psicopatía tienen mayor dificultad para distinguir las emociones. La Tabla 6 recoge los resultados de las pruebas. Por un lado, en naranja se encuentra los sujetos que obtuvieron altos resultados de psicopatía (HCU), y en azul aquellos cuyos resultados fueron bajos (LCU). Se ve que no hay grandes

diferencias en los resultados de ambos grupos, y por ello no se pueden clasificar acorde a sus respuestas. Pero se pueden distinguir dos sujetos cuya tasa de respuesta esta muy por debajo de la media, y podrían distorsionar los datos.

| Individuo HCU  | Aciertos | Fallos | NC |
|----------------|----------|--------|----|
| <b>Sergio</b>  | 66       | 10     | 4  |
| <b>Samuel</b>  | 53       | 27     | 0  |
| <b>Luis</b>    | 51       | 21     | 8  |
| <b>Wender</b>  | 55       | 24     | 1  |
| <b>Hector</b>  | 41       | 39     | 0  |
| <b>Claudio</b> | 60       | 20     | 0  |
| <b>Mikel</b>   | 66       | 10     | 4  |

| Individuo LCU    | Aciertos | Fallos | NC |
|------------------|----------|--------|----|
| <b>Alejandro</b> | 63       | 16     | 1  |
| <b>Andrei</b>    | 47       | 29     | 4  |
| <b>David</b>     | 58       | 21     | 1  |
| <b>Diego</b>     | 16       | 61     | 1  |
| <b>David</b>     | 56       | 21     | 3  |
| <b>Jeison</b>    | 60       | 18     | 2  |
| <b>Gabriel</b>   | 26       | 47     | 7  |

*Tabla 6 Tasas de respuestas de la prueba Ekman. (Naranja HCU; Azul LCU).*

Además, para comprender mejor las respuestas, las características y el desempeño de cada individuo en la prueba Ekman, se ha procedido a calcular la exactitud, sensibilidad y especificidad de cada sujeto. Para ello se ha definido la siguiente matriz de confusión:

|                        |        | Respuesta del sujeto                           |   |
|------------------------|--------|--|---|
|                        |        | Miedo  | Neutro  |
| Evento/Imagen mostrada | Miedo  | <b>Acierto (TM)</b><br>Miedo bien identificado | <b>Fallo (FM)</b><br>Miedo mal identificado     |
|                        | Neutro | <b>Fallo (FN)</b><br>Neutro mal identificado   | <b>Acierto (TN)</b><br>Neutro bien identificado |

La exactitud se ha calculado para conocer la proporción de eventos que el individuo es capaz de identificar correctamente, para ello se ha usado la siguiente ecuación:

$$Exactitud = \frac{TM + TN}{TM + TN + FN + FM}$$

Además, se ha calculado la sensibilidad y la especificidad. La sensibilidad nos indica la capacidad de los individuos para detectar eventos de miedo y la especificidad la capacidad para detectar imágenes neutras.

$$\text{Sensibilidad} = \frac{TM}{TM + FM}$$

$$\text{Especificidad} = \frac{TN}{TN + FN}$$

En la Tabla 7 se pueden ver los parámetros calculados:

| Individuo HCU | Exactitud | Sensibilidad | Especificidad |
|---------------|-----------|--------------|---------------|
| Sergio        | 87%       | 88%          | 86%           |
| Samuel        | 34%       | 35%          | 32%           |
| Luis          | 71%       | 64%          | 80%           |
| Wender        | 70%       | 68%          | 72%           |
| Héctor        | 51%       | -            | 51%           |
| Claudio       | 75%       | 85%          | 70%           |
| Mikel         | 87%       | 88%          | 86%           |

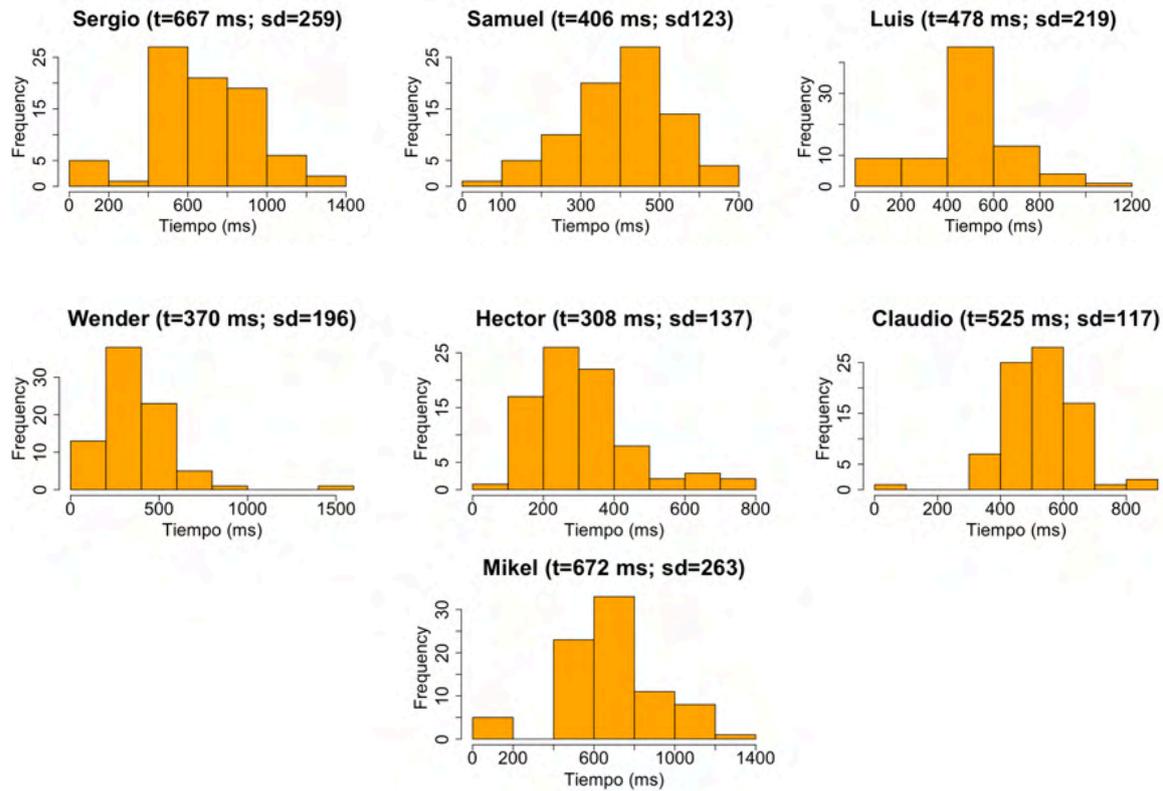
| Individuo LCU | Exactitud | Sensibilidad | Especificidad |
|---------------|-----------|--------------|---------------|
| Alejandro     | 80%       | 81%          | 79%           |
| Andrei        | 62%       | 60%          | 64%           |
| David         | 73%       | 90%          | 67%           |
| Diego         | 21%       | 17%          | 24%           |
| Iván          | 73%       | 74%          | 72%           |
| Jeison        | 77%       | 72%          | 83%           |
| Gabriel       | 36%       | 36%          | 34%           |

*Tabla 7 Exactitud, sensibilidad y especificidad.*

Analizando estos parámetros se puede ver que algunos sujetos como Samuel, Diego y Gabriel tienen muy baja exactitud. Destaca que pertenecen a distintos grupos, por lo que se plantea la hipótesis de que, al realizar la prueba, estos individuos hayan estado distraídos, o que hayan entendido incorrectamente la prueba, ya que sus resultados no son coherentes con el resto de las personas. Por otro lado, Héctor ha respondido a todos los eventos con el mismo botón (Neutro), y por ello ha obtenido 51% de exactitud, con esto se concluye que no ha realizado la prueba correctamente.

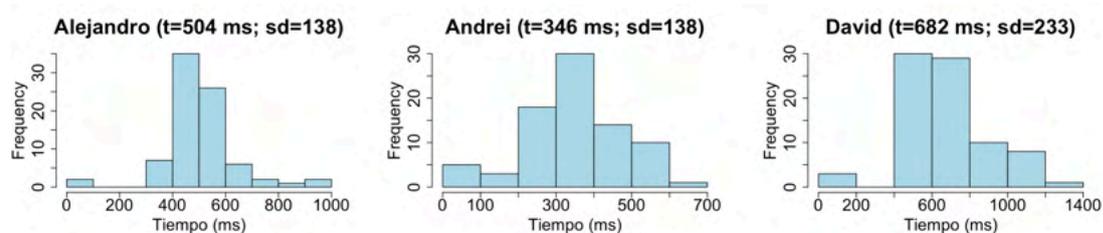
Tras esto, se procedió a estudiar los tiempos de respuesta de cada uno de los sujetos. Para ello se realizó un histograma de estos, y se calculó el tiempo medio y la desviación típica. Esto se hizo para comprobar si había mayor impulsividad en algunos miembros de los grupos y para corroborar las hipótesis anteriores. Por un lado, en la Figura 9 se tienen los

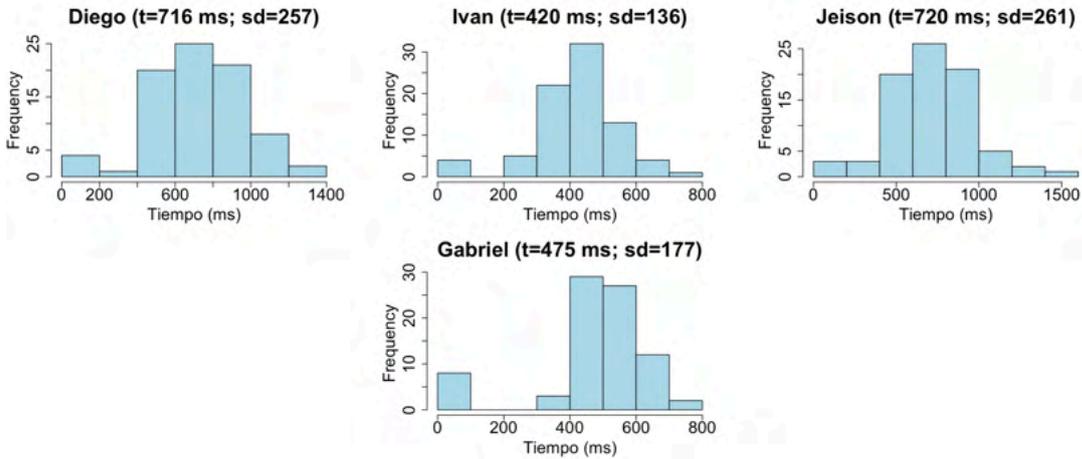
histogramas de tiempo de respuesta, el tiempo medio de respuesta y la desviación típica de cada sujeto del grupo con rasgos HCU.



*Figura 9 Histogramas, tiempos medios y desviación típica de los tiempos de respuesta de cada sujeto con altos rasgos CU.*

Por otro lado, en la Figura 10 se tienen los histogramas de tiempo de respuesta, el tiempo medio de respuesta y la desviación típica de cada sujeto con rasgos LCU.





*Figura 10 Histogramas, tiempos medios y desviación típica de los tiempos de respuesta de cada sujeto con bajos rasgos CU*

De las figuras anteriores, se extraen las siguientes conclusiones. En primer lugar, es coherente que Héctor sea el sujeto con los tiempos de respuesta más bajos, puesto que, como se había dicho antes, no realiza la prueba con seriedad, por lo que siempre tiene la misma respuesta y es casi instantánea. En segundo lugar, la hipótesis de que Diego no ha entendido bien el experimento es coherente, puesto que sus ratios son muy bajos, pero sus tiempos de respuestas muy altos. Es decir, se piensa las respuestas, pero a la hora de responder pulsa la tecla equivocada. Y, por último, se concluye que ninguno de los grupos muestra mayor impulsividad a la hora de identificar las imágenes, ya que en general los tiempos son muy similares para todos los individuos.

Para continuar con el análisis de las respuestas de los individuos se decidió analizar la tendencia en la cual cada uno de los sujetos fallaba, resultados que se pueden ver en la Tabla 8. Es decir, si el sujeto respondía miedo a un estímulo neutral, o neutral a un estímulo del miedo. Esto se realizó con la idea de que los sujetos con alto grado de psicopatía tuviesen mayor dificultad para identificar miedo, y su tendencia de fallo fuese más distinguible ante este tipo de estímulos. También está la posibilidad de que algunos individuos tengan mayor dificultad para identificar imágenes neutrales, y esto implicaría que existe una sobre activación a estos estímulos, y presentaría otros rasgos.

| Individuo HCU  | Neutro mal identificado | Miedo mal identificado | Individuo LCU    | Neutro mal identificado | Miedo mal identificado |
|----------------|-------------------------|------------------------|------------------|-------------------------|------------------------|
| <b>Sergio</b>  | 4                       | 6                      | <b>Alejandro</b> | 7                       | 9                      |
| <b>Samuel</b>  | 30                      | 23                     | <b>Andrea</b>    | 16                      | 13                     |
| <b>Luis</b>    | 15                      | 6                      | <b>David</b>     | 2                       | 19                     |
| <b>Wender</b>  | 13                      | 11                     | <b>Diego</b>     | 30                      | 31                     |
| <b>Héctor</b>  | 0                       | 39                     | <b>David</b>     | 10                      | 11                     |
| <b>Claudio</b> | 4                       | 16                     | <b>Jeison</b>    | 12                      | 6                      |
| <b>Mikel</b>   | 4                       | 6                      | <b>Gabriel</b>   | 28                      | 19                     |

*Tabla 8 Resumen de la tendencia de identificación errónea de cada individuo.*

A la vista de los resultados, se concluye que ningún grupo tiene mayor dificultad para identificar un tipo imágenes, es decir, que el grupo con rasgos HCU no presenta muchas diferencias respecto del otro. Omitiendo sujetos como Samuel, Héctor o David, cuyos resultados no son un buen ejemplo del resto del grupo, por los hechos que se han apuntado anteriormente. Se observa que existe un mayor número de sujetos de ambos grupos que tienen dificultad para identificar las imágenes de miedo. Estos puede deberse a los ambientes a los que han sido expuestos los sujetos. Sería interesante comparar estos grupos con individuos que no se encuentren en centros de menores.

Finalmente, tras realizar el análisis exploratorio se concluyó que no existe gran diferencia entre los individuos de ambos grupos. Asimismo, se ve que la clasificación basada en el cuestionario “Inventory of Callous–Unemotional traits”, y por el cual hemos agrupado a los distintos individuos, no correlaciona claramente con las tendencias a fallar en el test Ekman a la que se sometió a los individuos. Una de las principales razones por las que se cree que no existe una correlación clara es por la voluntad y seriedad con la que los sujetos han realizado las pruebas. Por todo esto, en el capítulo 7 y 8 se realizará un análisis detallado de las señales cerebrales, con el objetivo de identificar factores inconscientes ante los estímulos neutrales y de miedo, que no serían capaz de controlar los individuos.

## Capítulo 5. PROCESADO DE LAS SEÑALES

En este capítulo, se explicará el procedimiento seguido para el procesado de las señales que servirán para el estudio de los capítulos posteriores. En primer lugar, se explicará qué se define como evento. Tras esto, se explicarán las distintas partes del procesado de la señal, tales como el filtrado de datos y la segmentación.

### 5.1 EVENTO

Un evento es un período de tiempo que es de interés en una prueba experimental. En el caso del experimento Ekman, el sujeto debe presionar distintos botones en función de si observa una imagen que representa miedo o una imagen neutra. Un evento, como se ve en Figura 11 sería el período de tiempo desde que se presenta el estímulo (la imagen), hasta un tiempo posterior, en el cual, el sujeto ya ha introducido su respuesta.

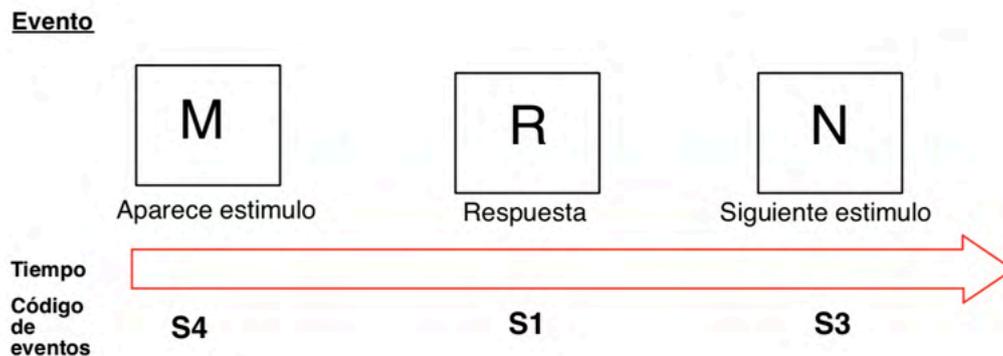


Figura 11 Evento de una parte del experimento Aman.

Los códigos que se han seguido en la prueba ekman para indicar el tipo de estímulo y la respuesta del sujeto son las siguientes:

- Estímulo S4: código que indica que la imagen presentada, contiene rasgos neutrales.
- Estimulo S3: código que indica que la imagen presentada, contiene rasgos de miedo.
- Respuesta S2: código que indica que el individuo identifico la imagen como neutral.

- Respuesta S1: código que indica que el individuo identifico la imagen como de miedo.

En la Figura 12 se puede ver una parte del canal AF8, en el que se aprecian los distintos estímulos que se producen durante un experimento. Un evento, en este caso, correspondería con la señal durante el periodo de tiempo entre estímulos, es decir, el primer evento sería la señal de S4 a S3, y el segundo evento de S3 a S4. Además, se puede apreciar las respuestas S1 del sujeto a los estímulos presentado.

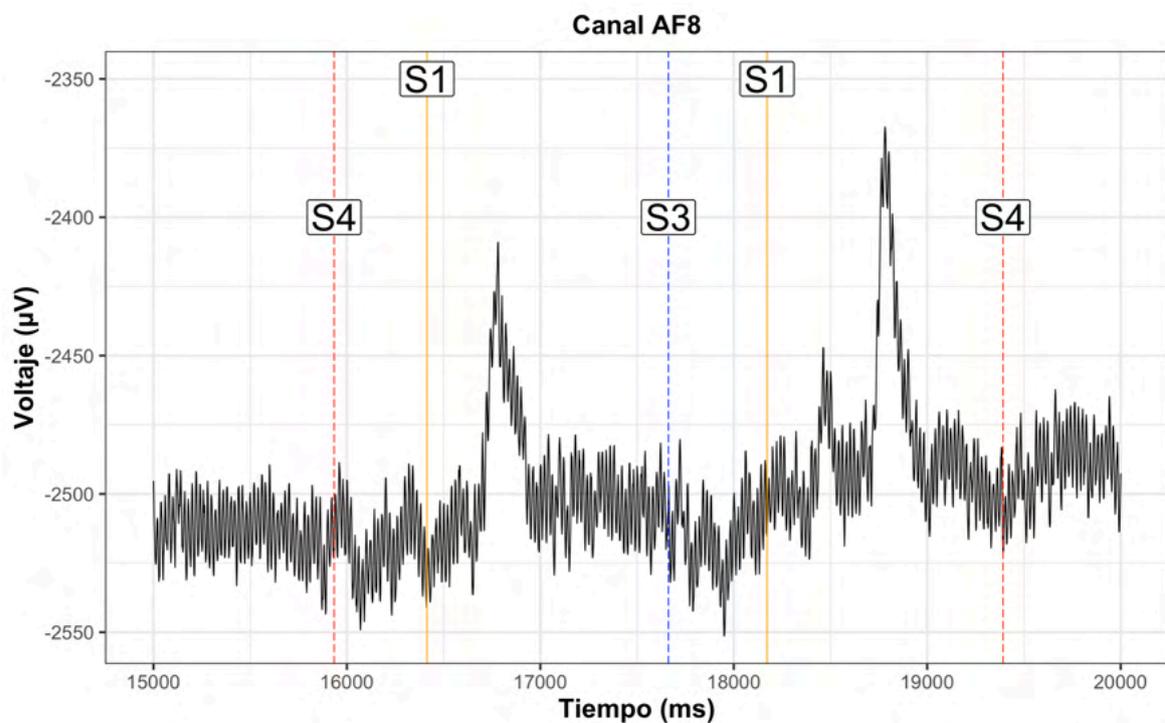
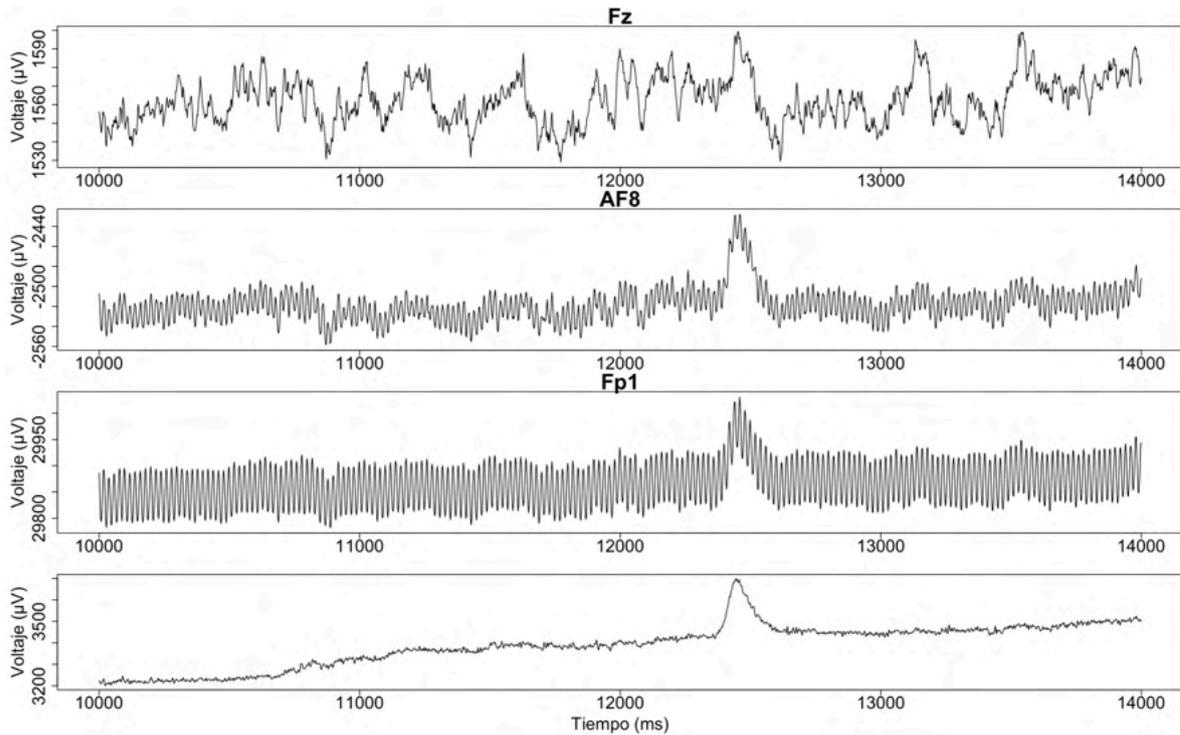


Figura 12 Canal AF8 del experimento Ekman.

## 5.2 SEÑALES SIN PROCESAR

En este apartado se expondrá cómo son las señales que se obtienen a partir de los canales del encefalograma. En la Figura 13 se ven señales de 4 canales diferentes. Los canales Fp1, F7 y AF4 son canales que miden la actividad cerebral, y el canal VEOG\_1 mide la actividad ocular. En la Figura 13 se observa que las señales tienen una componente importante a alta

frecuencia. Por otro lado, se aprecian unos picos en todos los canales, estos picos son debidos a la actividad ocular, y no interesan para nuestro análisis, puesto que distorsionan la señal y no proporcionan información relevante sobre la actividad cerebral.



*Figura 13 Canales Fp1, F7, AF4 y VEOG\_1 del experimento Ekman.*

### **5.3 FILTRADO DE ARTIFICIOS DE LAS SEÑALES**

En este apartado, se explicará el proceso por el cual se han eliminado los artificios que se pueden ver en las señales. Un artificio es cualquier potencial eléctrico que no se origina en el cerebro. Existen dos tipos básicos de artificios: a) los fisiológicos, que se deben al funcionamiento normal del cuerpo (parpadeo, respiración, movimiento de la lengua, etc.), y b) los no fisiológicos que se deben al ruido ambiente e instrumentaría utilizada/o [36]. Para este estudio, se ha decidido eliminar principalmente los artificios fisiológicos, en especial, los parpadeos.

En primer lugar, se ha hecho un análisis independiente de componentes (ICA). Esto es un método estadístico cuyo objetivo es encontrar una representación lineal de unos datos multivariantes, y en los que las componentes son lo más independientes posible, es decir, maximizar la independencia mutua de las componentes. Como se puede ver en la literatura [37], el análisis independiente de componentes ha demostrado tener la capacidad de separar las señales EEG en componentes independientes, siendo algunas componentes la actividad neuronal de distintas partes del cerebro y otra debida a parpadeos, ruidos, o movimientos musculares.

En el caso de las señales de un electroencefalograma tenemos una señal de  $n$  canales  $x = [x_1, x_2, x_3, \dots, x_n]$  procedente de  $m$  fuentes independientes  $s = [s_1, s_2, s_3, \dots, s_m]$ .

$$x = sA$$

Donde  $A$  es una matriz de transformación lineal de  $m \times n$  dimensiones. El estudio ICA busca encontrar una matriz de transformación lineal que permita obtener una versión original de la fuente, que será igual excepto por la escala.

$$xW = u \rightarrow W \approx A^{-1}$$

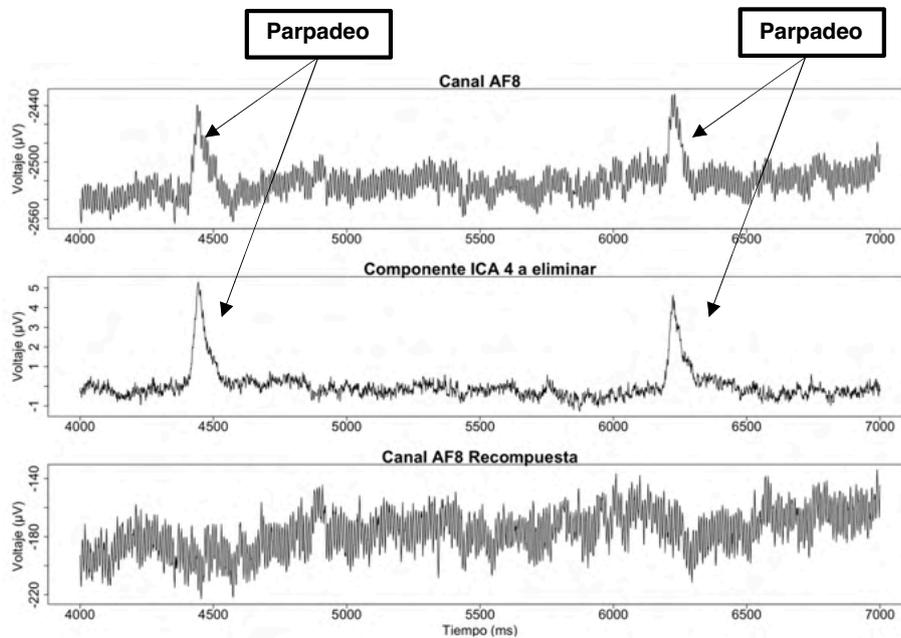
Donde la matriz resultante  $u$  será una versión de la fuente original:

$$u = xW = (sA)W = s' \approx s$$

Y donde las señales originales,  $s$ , deben ser estadísticamente independiente en todo instante de tiempo y la mezcla de las fuentes lineal.

Este concepto que se acaba de explicar es el que se ha utilizado para obtener las distintas fuentes independientes, procedentes tanto de la actividad cerebral, como de los distintos artificios. Una vez se obtuvieron las componentes, se analizaron manualmente, para distinguir aquéllas que se deben a los parpadeos, y las que se deben a otros artificios. Estas componentes se anularon, y tras esto se recompusieron las señales originales, eliminando así los picos que se debían a los artificios fisiológicos.

En la Figura 14 se puede ver el proceso que se ha seguido. La primera gráfica muestra una parte del canal AF8, en la que se pueden apreciar unos picos que se deben principalmente al parpadeo del sujeto y serán éstos los que se deben eliminar. Para ello se ha realizado un análisis independiente de componentes, y se puede evaluar que la componente cuatro identifica los parpadeos, y será la que debe eliminarse. Finalmente, en la última gráfica se puede ver el canal AF8 tras recomponer las señales, y se aprecia que ya no hay artificios fisiológicos destacables.



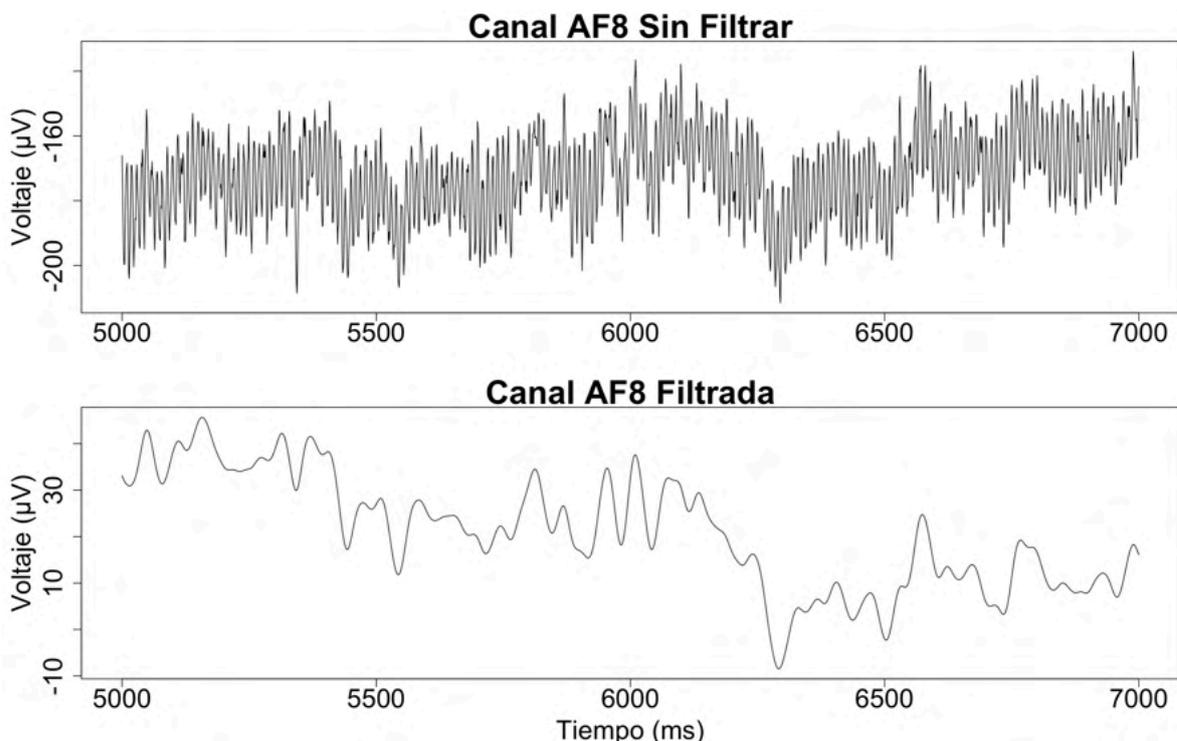
*Figura 14 Proceso ICA para la eliminación de artificios.*

## **5.4 FILTRADO DE LOS DATOS**

Una vez eliminados los distintos artificios, se procede al filtrado de la señal, para eliminar el ruido de alta frecuencia que distorsiona nuestra señal. Para ello se realiza un filtro paso banda, con frecuencias entre 3-18 Hz. Se han elegido estas frecuencias porque son las que incluyen las frecuencias de las distintas ondas cerebrales de interés [38]:

- Theta (3,5-8Hz): memoria plástica, imaginación, inspiración creativa y capacidad de aprendizaje.
- Alfa (8-13Hz): estados de relajación
- Betas inferiores (12-16Hz): analiza, cálculos, pensamientos del ambiente externo.

En la Figura 15 se pueden distinguir por un lado, las señales antes de ser filtradas, y por otro, las señales una vez que han pasado por el paso banda. Se observa que se reduce mucho el ruido de alta frecuencia.



*Figura 15 Representación del proceso de filtrado de señales.*

## **5.5 EPOCHING/SEGMENTACIÓN DE LOS DATOS**

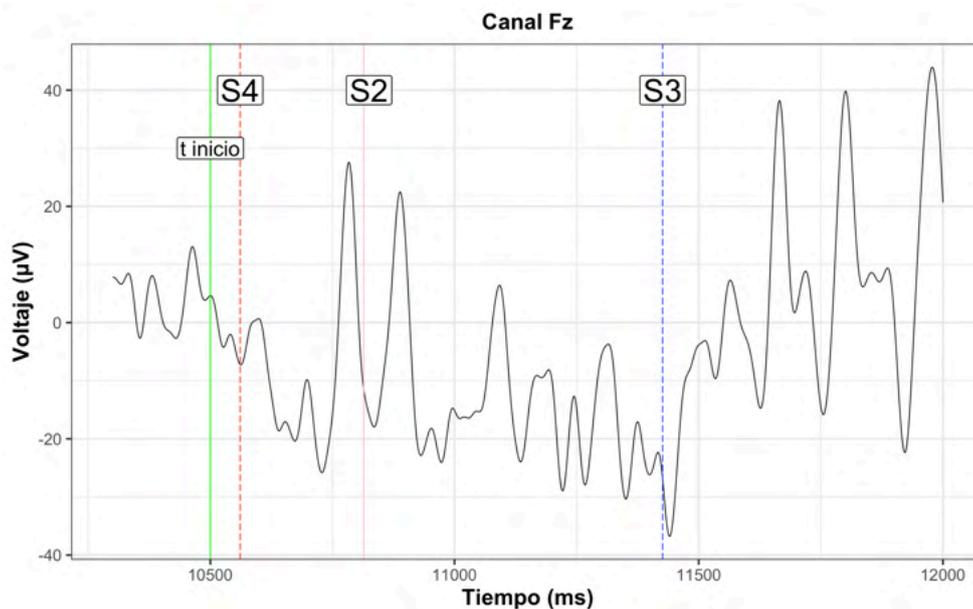
En este apartado, se refiere el procedimiento por el que se han segmentado los datos. Se ha utilizado un proceso de segmentación de datos, llamado epoching. Como ya se ha explicado, durante el experimento, a los sujetos se les presentan una serie de estímulos. Epoching consiste en extraer de los datos, unos intervalos de tiempo específicos alrededor de estos

estímulos. Estos intervalos, se llaman “epochs”, y son segmentos de datos que suelen tener un tiempo fijo. En este caso, se ha segmentado cogiendo 100ms antes de que aparezca el estímulo, y el tiempo hasta que se presenta el siguiente estímulo.

En el experimento Ekman, se presentan 80 estímulos, por lo que una vez segmentado los datos tendremos 62 canales diferentes, divididos a su vez cada uno, en 80 epochs, algunos correspondientes al estímulo S4 (neutro) y otros al S3 (miedo).

El proceso que se ha seguido es el siguiente:

1. Se parte de la señal original (que se puede ver en Figura 16);
2. Sobre esta identificamos los tiempos en los cuales se producen estímulos
3. Se escoge el intervalo de tiempo que interesa. En este caso sería desde  $t_{inicio}$  (100 ms antes del estímulo S4) hasta el estímulo S3.



*Figura 16 Canal Fz, con tiempos de un epoch.*

Una vez seleccionados los intervalos de tiempo, se segmentan los datos, y se obtienen los epochs como los que pueden verse en la Figura 17.

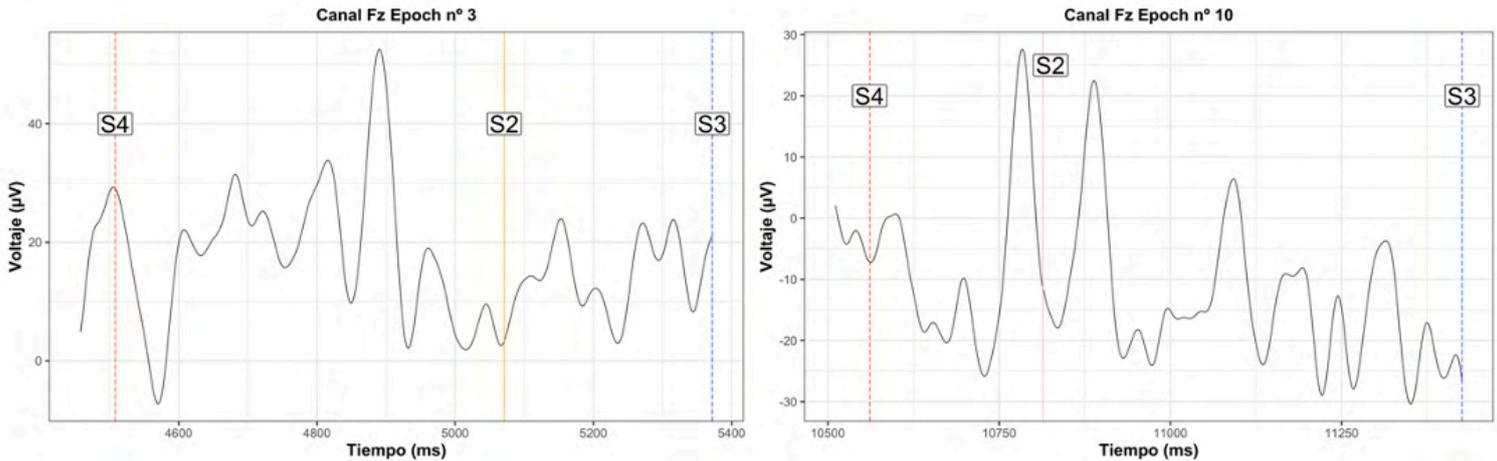


Figura 17 Epoch n° 3 y n° 10 del canal Fz.

Una vez que se tienen los epochs se realiza una corrección para que todos estén en la misma escala. Para ello se resta al epoch la media de la señal, un poco antes de que esta comience, y así conseguimos escalar todos los epochs para el siguiente paso.

Finalmente, tras realizar la corrección sobre los epochs se pasa a la parte de análisis y clasificación de las señales, que será objeto de estudio en los capítulos posteriores. En el Capítulo 6. se obtendrán los potenciales relacionados con eventos, con los cuales se intentará identificar diferencias entre los individuos. En el Capítulo 7. se intentará extraer características de las epochs, para luego utilizarlas en los modelos de clasificación.

## **Capítulo 6. POTENCIAL RELACIONADO CON EVENTOS**

En este capítulo, se realizará un análisis de las señales, para lo que se usarán los potenciales relacionado con eventos. Estos se obtendrán para cada sujeto, y para la prueba en su totalidad, y servirán para realizar comparaciones entre distintos sujetos.

### ***6.1 POTENCIAL RELACIONADO CON EVENTOS***

En las pruebas realizadas, se utiliza un electroencefalograma con el que se pueden obtener señales que representan las variaciones de voltaje en el cerebro. Durante estas pruebas, se presentan una serie de estímulos a los sujetos, y lo que interesa para el estudio es su respuesta a estos. Por ello, se realiza una segmentación de los datos en distintos epochs. Es decir, se observa una parte de la señal del EEG definida entre dos tiempos  $t_1$ (tiempo antes del estímulo) y  $t_2$ (tiempo tras el estímulo). Esto constituye un potencial relacionado con evento, es la respuesta cerebral de un sujeto ante un evento. Sin embargo, surge el problema de que estos cambios en el potencial son muy pequeños, y que además las señales aparecen con mucho ruido, y distintos artificios que han de ser eliminados, mediante el procesamiento de las señales como se explica en el Capítulo 5.

### ***6.2 MEDIA DE LOS SUJETOS***

En este apartado, se realiza la media de todos los epochs para cada sujeto y así se obtiene una señal más limpia y suave que facilitará el estudio de la prueba. En la Figura 18 se puede ver el promedio de los epochs para algunos canales. Se aprecia que las señales quedan más limpias, y con menos artificios y ruido.

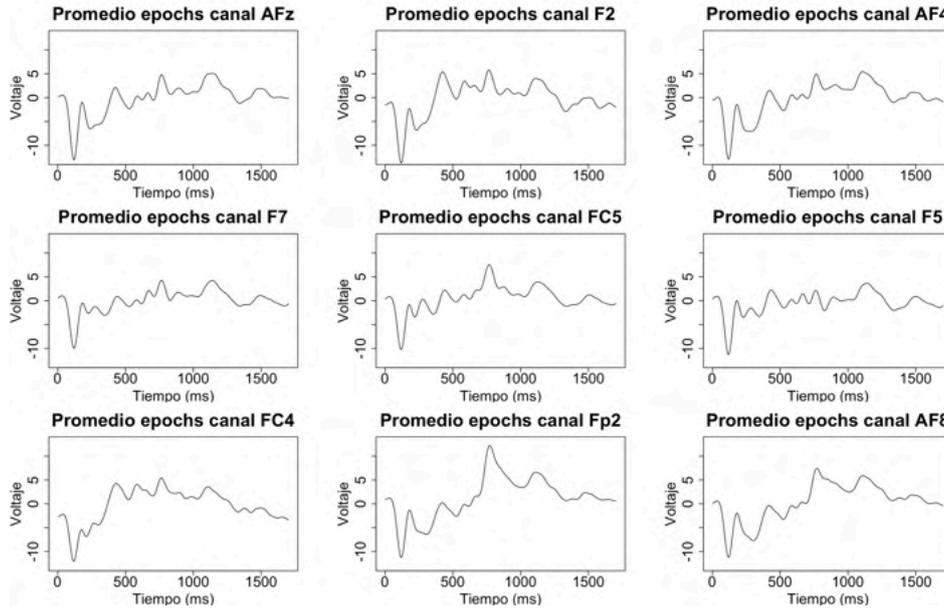


Figura 18 Promedio de todos los epochs para algunos canales.

Además de realizar la media del conjunto de epochs, se ha realizado la media del conjunto de epochs según el estímulo, es decir, por un lado, la media de todos los epochs que corresponden con un estímulo S4 y por otro lado con un estímulo S3. Finalmente, por cada sujeto obtendríamos 3 señales para cada canal, que se pueden ver en la Figura 19; en este caso, para el canal Fz.

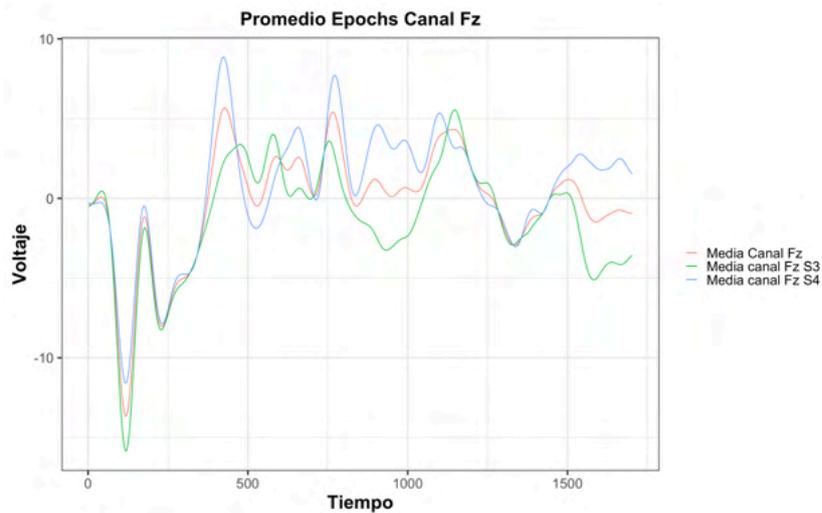
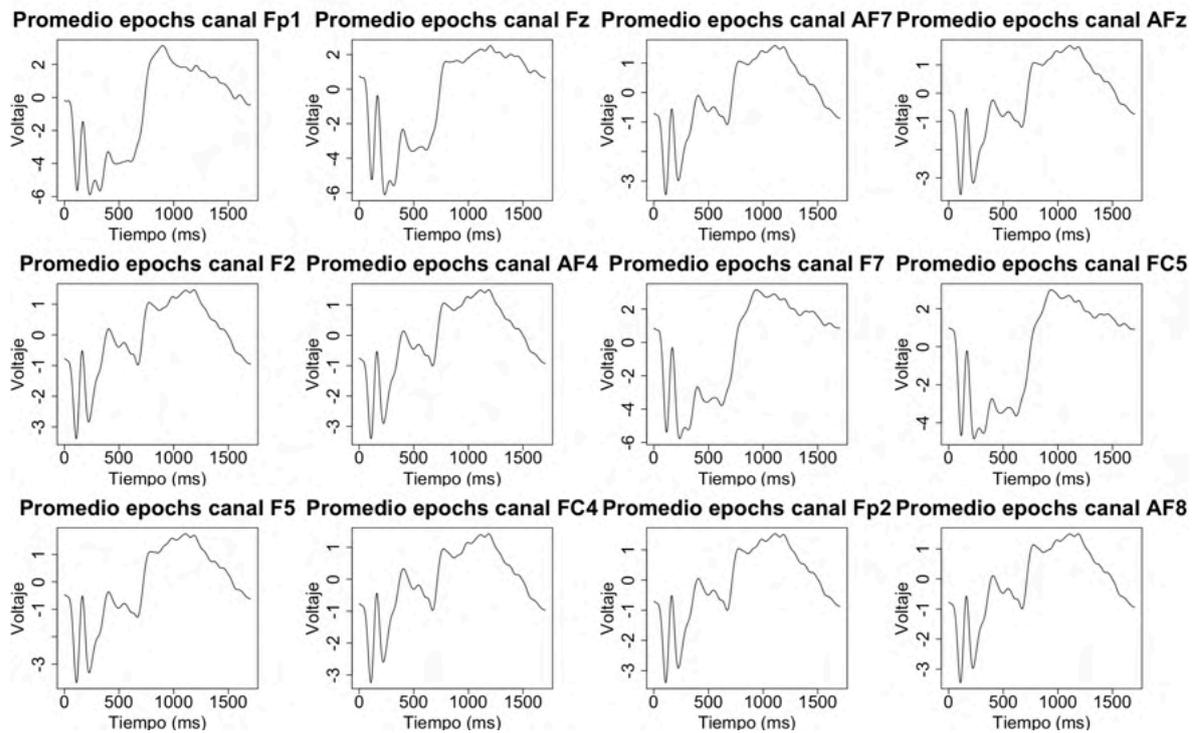


Figura 19 Media canal Fz total, frente al estímulo S3 y S4.

### 6.3 *MEDIA DE LAS PRUEBAS*

Una vez realizado el promedio para cada sujeto, se ha realizado una media de todos los sujetos del experimento para cada uno de los canales. En Figura 20 se pueden ver algunos canales. Se aprecia que la señal cambia respecto la anterior.



*Figura 20 Promedio de todos los sujetos para distintos canales.*

Asimismo, se ha podido comparar la media del grupo con la media de cada individuo, y ver si hay relaciones según el grupo al que pertenecen. A continuación, se observan los gráficos de líneas de dos canales, y se aprecia que las señales de algunos sujetos están bastante agrupadas. En verde, se encuentra el promedio de la señal para el grupo, en rojo aquellas señales de sujetos que obtuvieron altos resultados de psicopatía, y en azul los que obtuvieron bajos resultados.

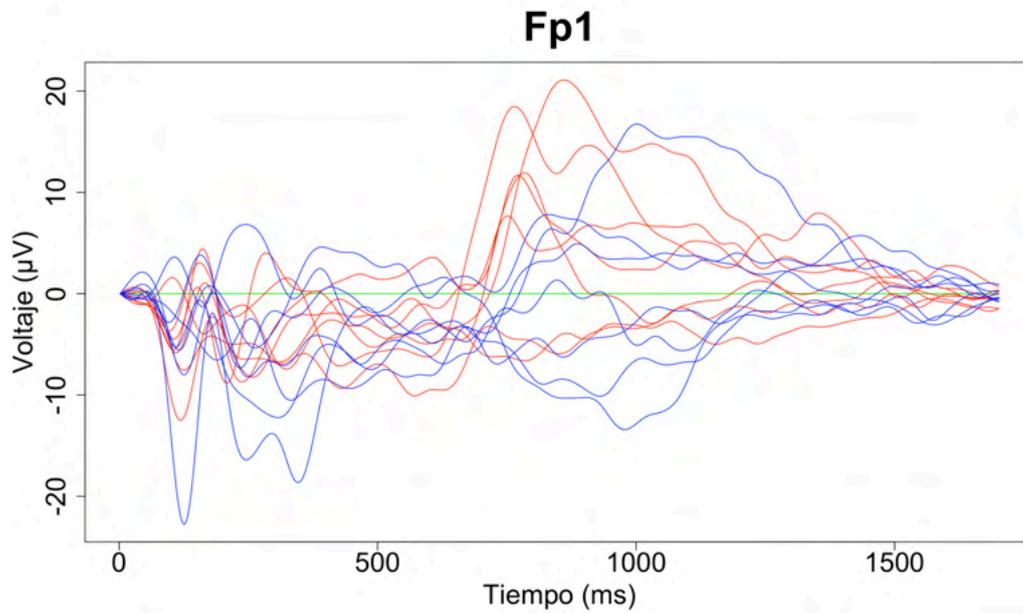


Figura 21 Media de epoch de cada sujeto vs Media de epoch de todo el grupo para el canal Fp1.

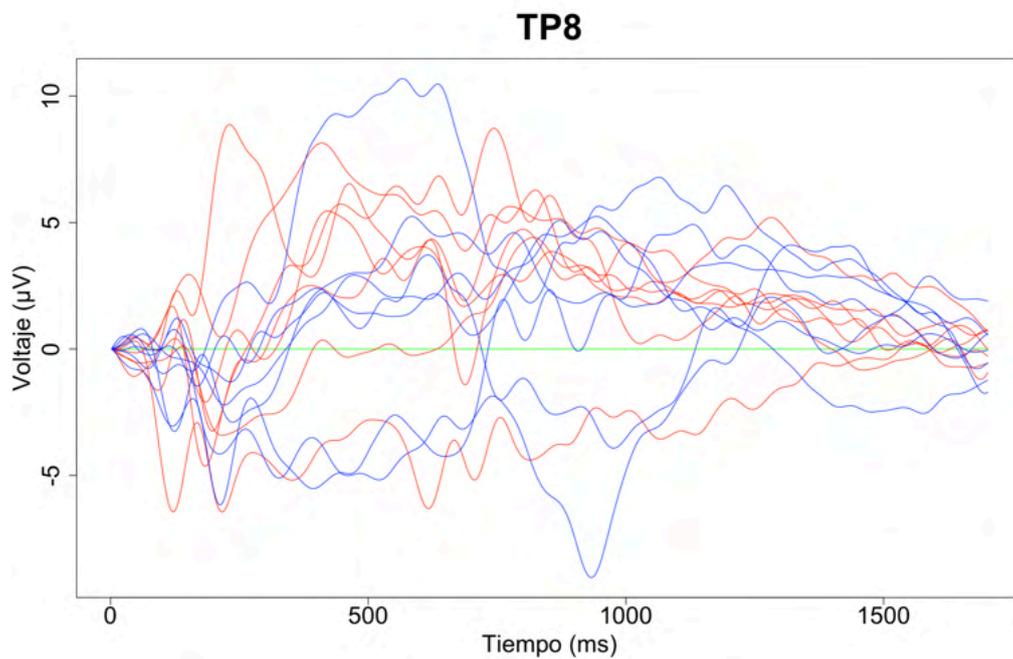


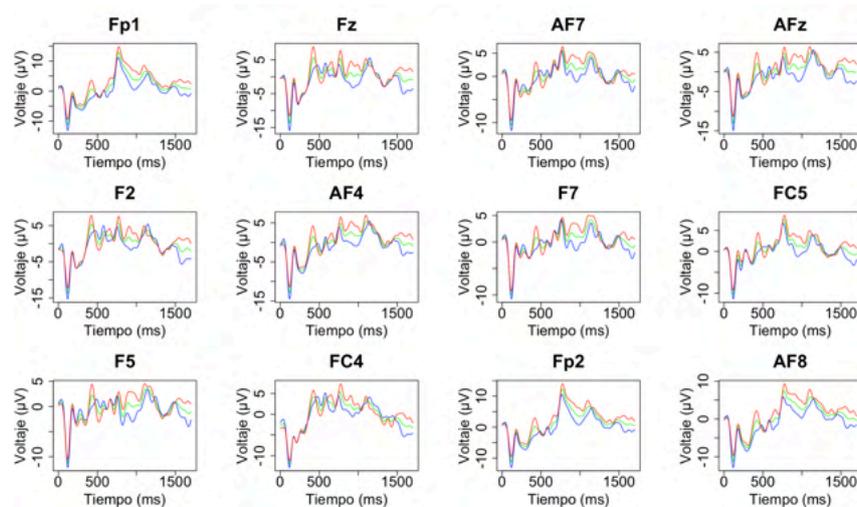
Figura 22 Media de epoch de cada sujeto vs Media de epoch de todo el grupo para el canal F5.

## **6.4 ANÁLISIS DE POTENCIALES RELACIONADOS CON EVENTOS**

En este apartado se realizará un análisis de potenciales relacionados con eventos. En el análisis de EEG en la neurociencia, se estudia la señal tras los estímulos. Se estudian los cambios en el potencial cerebral negativos y positivos, de estos se mira el instante en el que producen, la amplitud y el área que ocupan. Estos marcadores permiten comparar ERP entre si. La nomenclatura de estos componentes son N100,N200...donde la N indica que el cambio en el potencial es negativo, y el número el instante de tiempo en el que ocurre. Por otro lado, para cambios de potencial positivo se utiliza P50,P100...

### **6.4.1 POTENCIALES RELACIONADOS CON EVENTOS POR SUJETO**

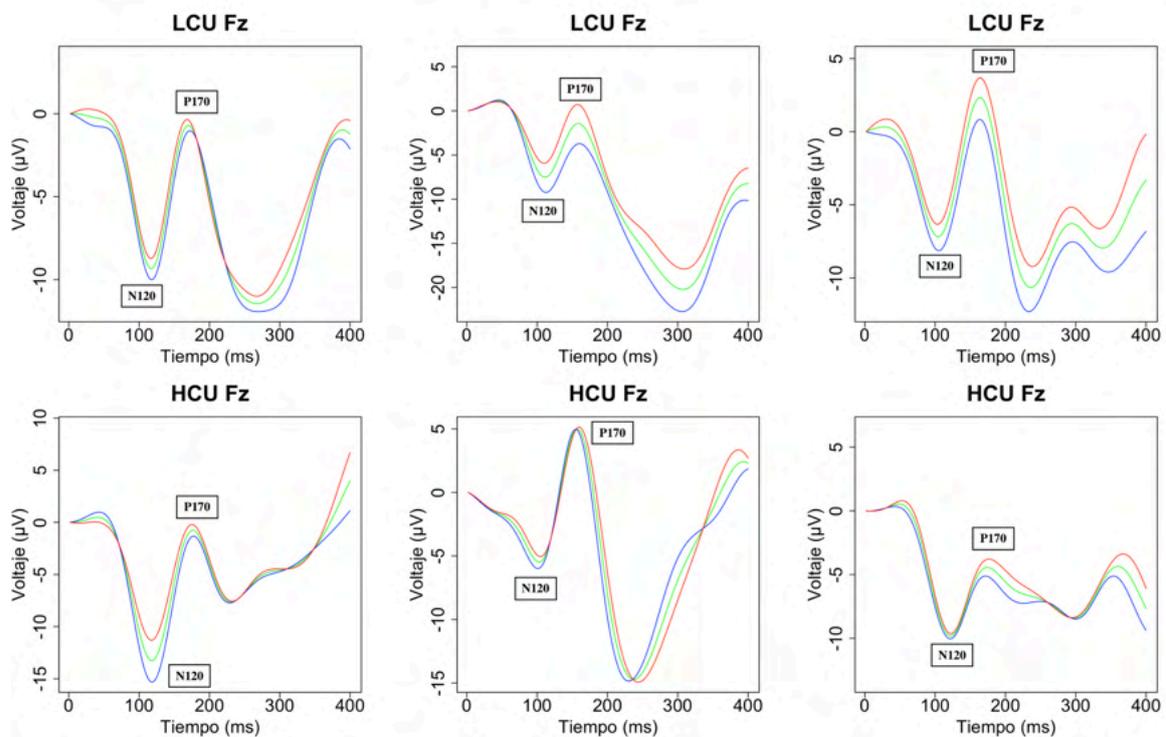
En este apartado se analizaran los ERP obtenidos de cada sujeto, donde compararemos la media de todos los epochs, frente a la media de las epochs que contiene solo estímulos S4 y S3. En la Figura 23 se pueden ver varios canales que son los más representativos. Donde las señales verdes son el promedio de la prueba total, las rojas el promedio de eventos S3, y las azules promedio de eventos S3. Se puede constatar que las diferencias entre la media de la prueba y de cada estímulo son bastantes parecidas para todos los canales, y bastante destacables.



*Figura 23 ERP para un sujeto. (Verde: media de toda la prueba; Azul: promedio de eventos S3; Rojo: promedio de eventos S4).*

Las señales anteriores son para todo el intervalo de tiempo tras el estímulo. A continuación, se pueden ver los 300ms para los canales Fz y AF4 de varios sujetos diferentes.

En la Figura 24, se observa que los estímulos S3 (miedo) producen una componente N120 de mayor amplitud en el canal Fz frente a la media de la prueba y el estímulo S4 (neutro). Esta diferencia se puede apreciar para todos los sujetos, independientemente de sus resultados en las pruebas de psicopatía. Además, para la componente P170, se percibe una mayor amplitud para el estímulo S4.



*Figura 24 ERP del canal Fz para sujetos LCU y HCU.*

Por otro lado, en la Figura 25, se ven las señales para el canal Cz. De este destacar los resultados más homogéneos para el grupo de bajo grado en psicopatía, donde la componente N120 tiene mayor amplitud para el estímulo S4. Sin embargo, para el grupo de alto grado en psicopatía se puede constatar que esta componente varía entre sujetos.

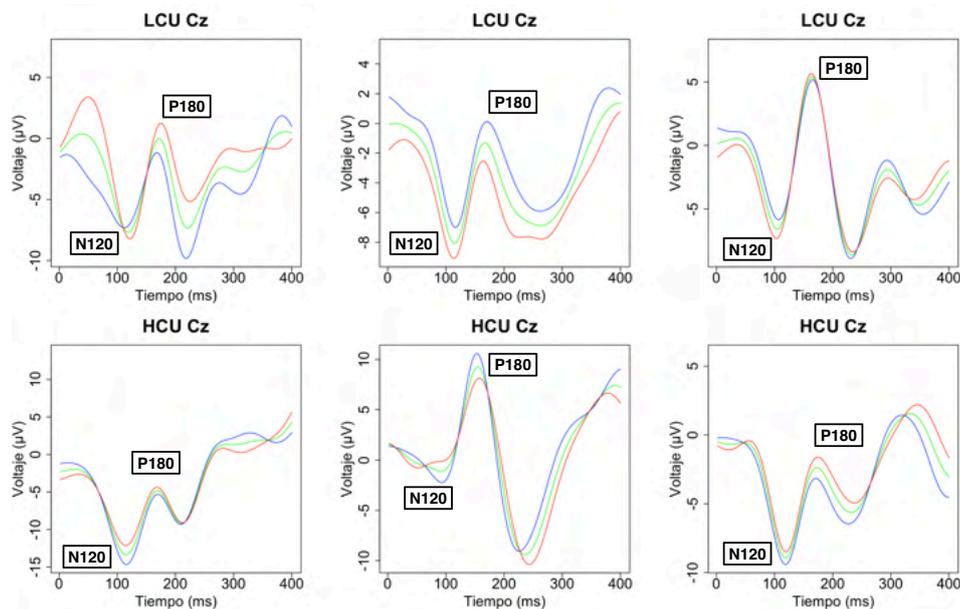


Figura 25 ERP del canal Cz para sujetos LCU y HCU.

Se han observado las componentes N120 y P170 para todos los canales, y se puede apreciar mucha diferencia entre canal, siendo las diferencias en amplitudes más homogéneas para algunos canales frente a otros.

Por otro lado, en la Figura 24 se muestra el canal Fz que está ubicado en la zona frontal que es la encargada de procesar emociones, por ello resulta interesante que la respuesta a los estímulos S3 (miedo) tenga mayor amplitud para la mayoría de los sujetos. Sin embargo, en la Figura 25 se muestra un canal de la zona parietal que se encarga de los sentidos, y posición del cuerpo. Por tanto, es coherente que las señales para ambos grupos sean bastante variadas, puesto que dependerá de otros factores al margen de la prueba Ekman.

Con esto se concluye que cada sujeto procesa los estímulos de forma diferente, independientemente del grupo al que pertenezcan, y por tanto conviene comparar los sujetos en base al total de la prueba y no separarla en estímulos. Esto se verá en el siguiente apartado.

## 6.4.2 POTENCIALES RELACIONADOS CON EVENTOS DE LA PRUEBA

En este apartado, se analizará los potenciales relacionados con eventos de la prueba Ekman en su totalidad. Se comparará la respuesta de los distintos sujetos, en función de al grupo que pertenecen. En la Figura 26, se pueden ver las señales de varios canales, donde la señal verde es la media de todos los sujetos, la roja la del grupo de HCU y la azul la del grupo LCU.

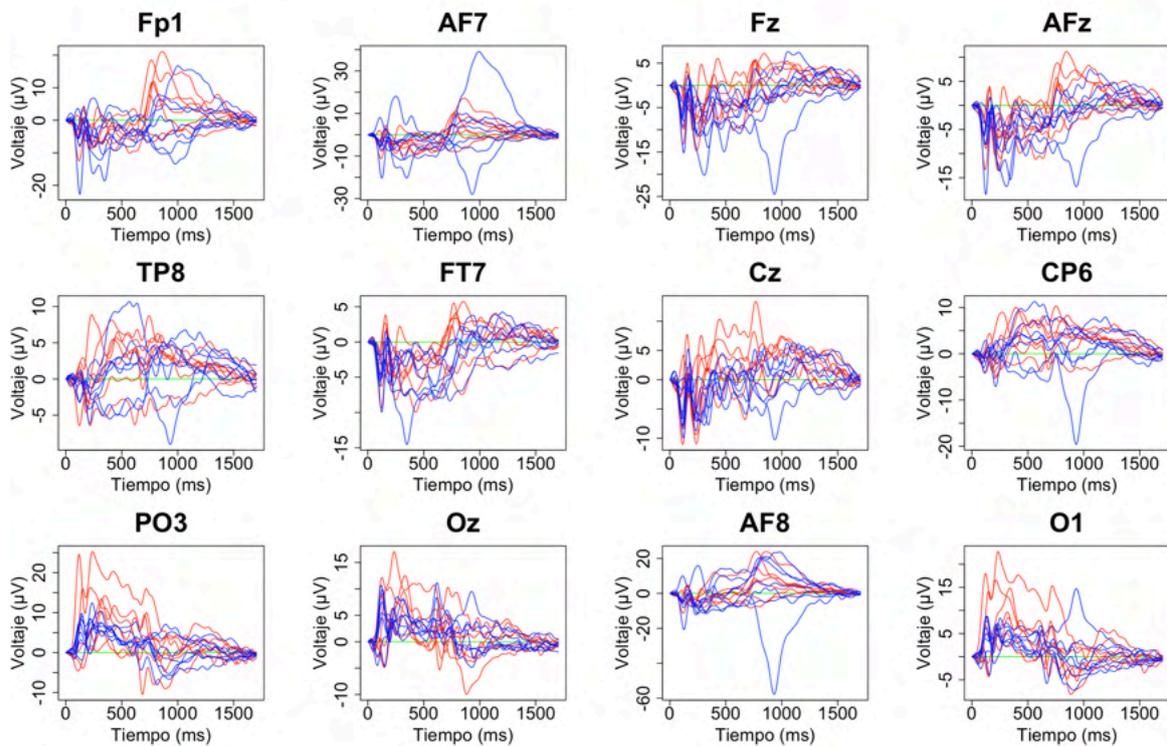


Figura 26 ERP para varios canales de individuos HCU (rojo) y LCU (azul).

Se puede ver que para algunos canales, los miembros de cada grupo están más concentrados, por ello se ha realizado la media de cada grupo y se ha analizado la componente N120 para estudiar las diferencias de los ERP de cada grupo. Además, se puede apreciar en el canal AF7 y Oz una señal que destaca sobre el resto, esta pertenece a David y no se ha tomado para realizar la media, puesto se ha considerado como un *outlier*.

Para estudiar las componentes, se ha decidido realizar un análisis para cada zona de la corteza del cerebro. En la Figura 27 se observan los canales asociados con el lóbulo frontal, en la

Figura 28 con el lóbulo parietal, en la Figura 29 con el lóbulo temporal, en la Figura 30 con el lóbulo occipital.

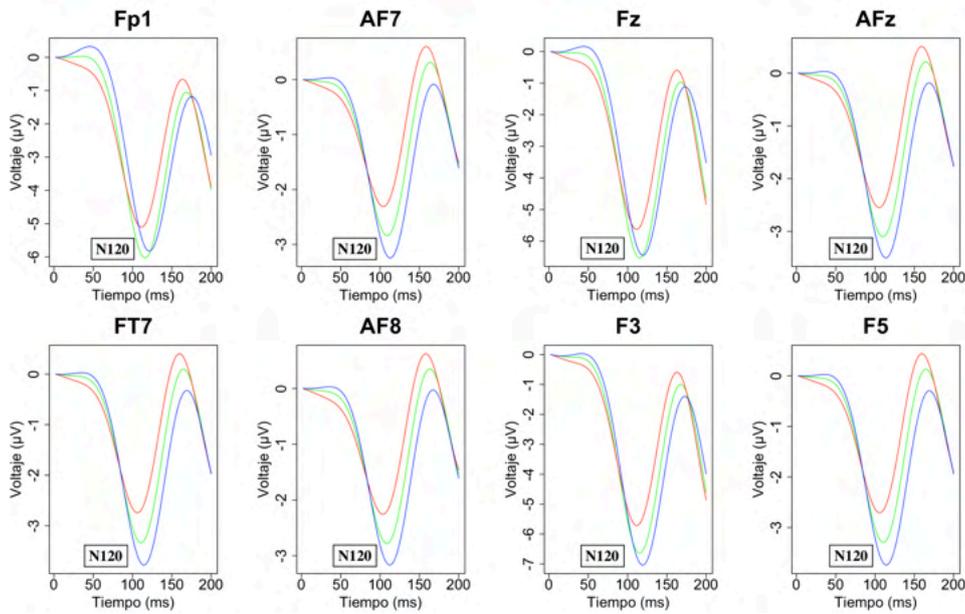


Figura 27 ERP asociado al lóbulo frontal.

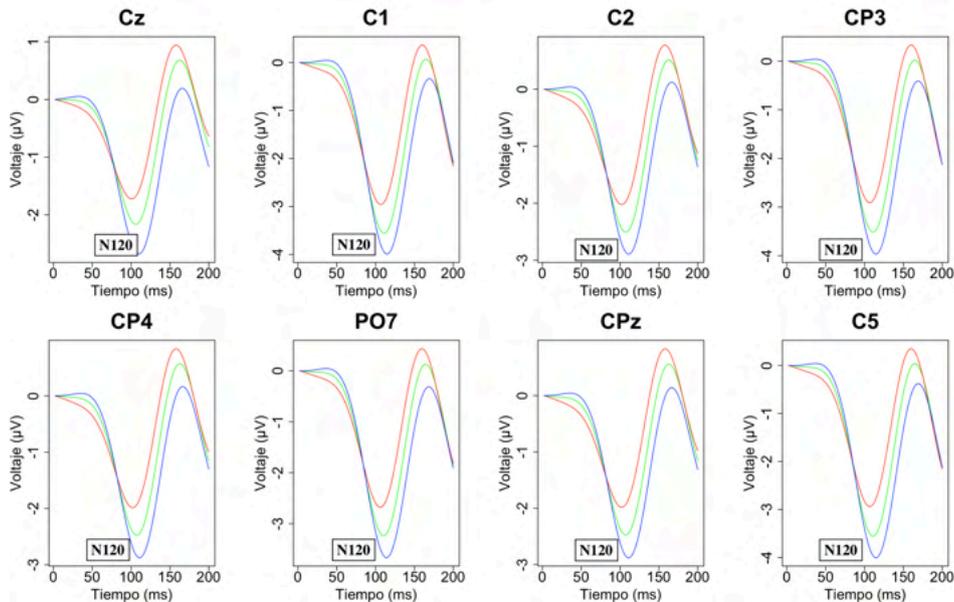


Figura 28 ERP asociado al lóbulo parietal.

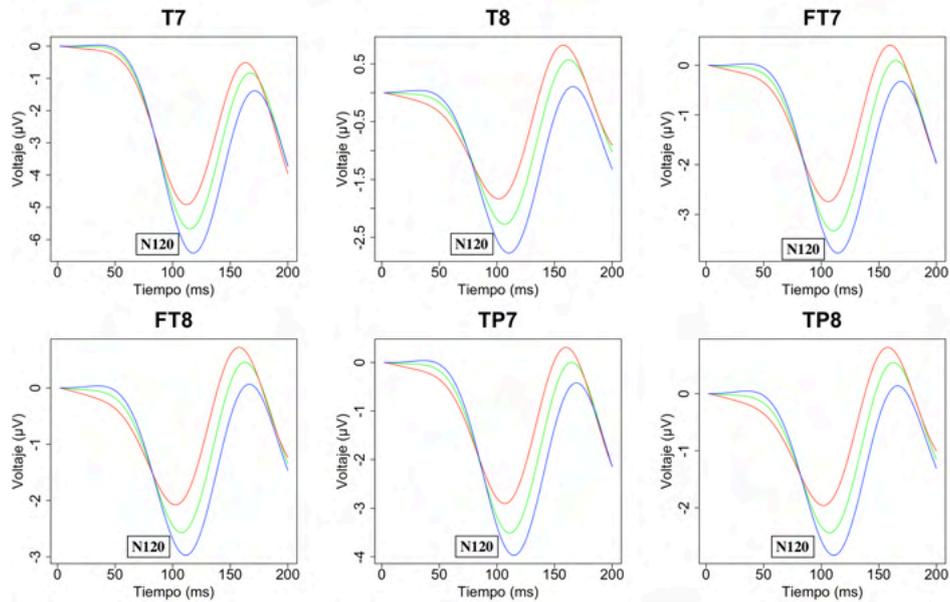


Figura 29 ERP asociado al lóbulo temporal.

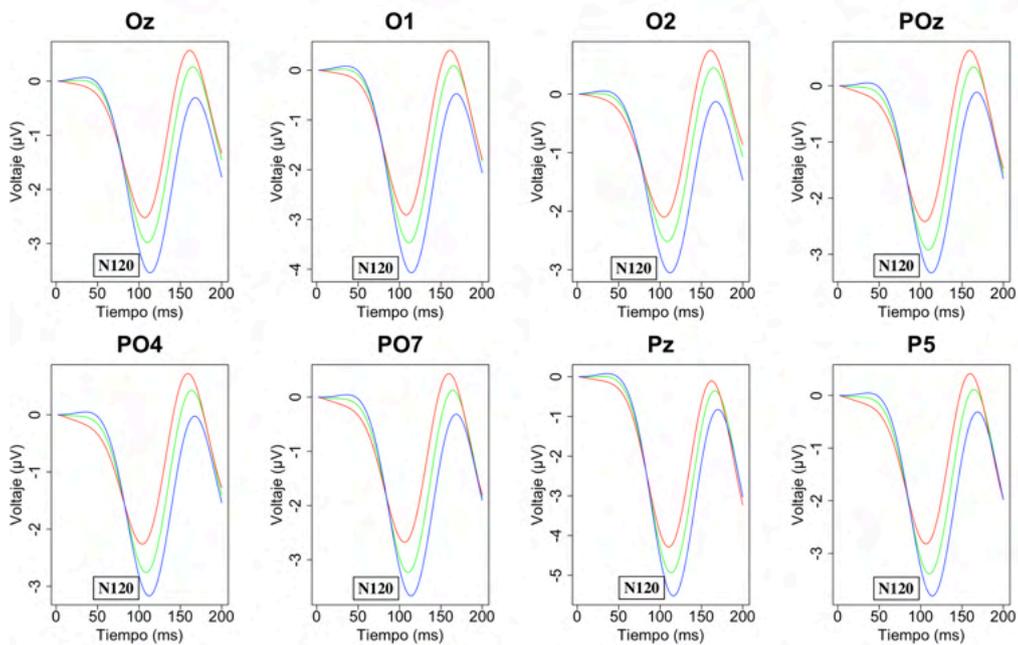


Figura 30 ERP asociado al lóbulo occipital.

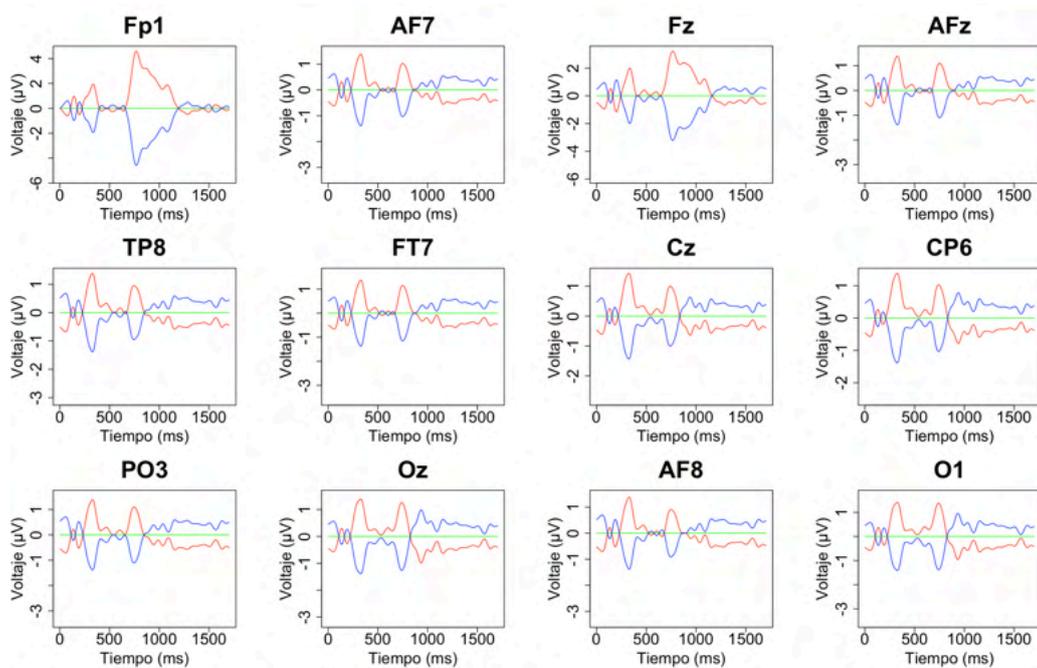
Tras realizar la media de las señales se pueden apreciar diferencias entre los dos grupos. Se constata que la componente N120 para el grupo con HCU (rojo) tiene una mayor amplitud

que el grupo LCU (azul), esto concuerda con la hipótesis de que los sujetos con HCU tienen dificultad para procesar las emociones, ya sean de miedo o neutras.

Asimismo, se observa mayor diferencia en los grupos para las zonas temporales que están relacionadas con la afectividad, y el reconocimiento de rostros, y para la zona frontal que procesa las emociones.

Todo esto concuerda con lo visto en la literatura [32], que indica que los individuos con HCU tienen problema para identificar emociones a través de la expresión facial. Un problema muy parecido al de las personas con la amígdala dañada, y consistente con estudios sobre el papel de la amígdala en el desarrollo de la psicopatía. Esto ha llevado a científicos a crear la hipótesis que la dificultad para identificar emociones en la infancia, provoca en el desarrollo de los individuos dificultades para la empatía.

Además de comparar los ERP se decidió restar a las señales de cada grupo la media total del experimento para ver si existían diferencias significativas. Esto se puede ver en la Figura 31 para algunos canales.



*Figura 31 ERP de varios canales restándole el promedio del experimento.*

Aquí puede apreciarse, que las señales rojas contribuyen positivamente por encima de la media para tiempos cortos y las azules, en general para tiempos largos. Esa diferencia puede tener que ver con la activación de regiones de "consciencia" (tiempos largos) frente a inconscientes (tiempos cortos).

## Capítulo 7. EXTRACCIÓN DE CARACTERÍSTICAS Y CLASIFICACIÓN

En este capítulo, se realizará un proceso de extracción de características de las señales cerebrales. Tras esto se usarán métodos para la reducción de las dimensiones de los datos, y finalmente se hará un modelo de clasificación, que permita separar a los individuos según sus características.

### **7.1 EXTRACCIÓN DE CARACTERÍSTICAS**

*Feature extraction* es un proceso por el cual a partir una serie de datos (señales de EEG) se obtienen una serie de “*features*” o características, que permitan comprender mejor los datos.

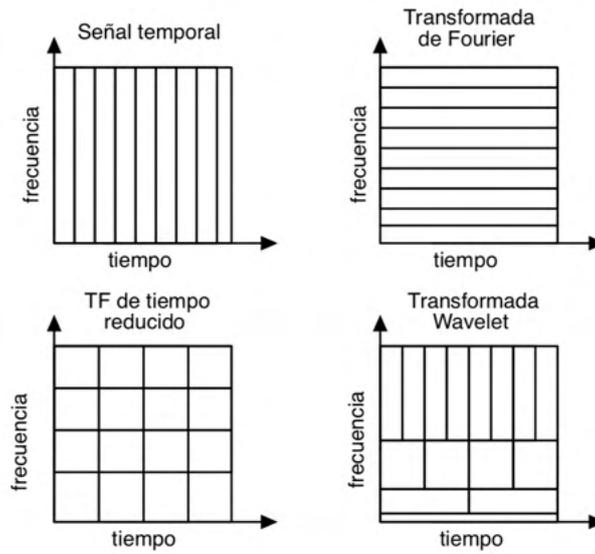
#### **7.1.1 TRANSFORMADA DISCRETA WAVELET (DWT)**

La transformada discreta wavelet, es un método muy utilizado para el análisis de señales no estacionarias, que permite una representación en tiempo y frecuencia. Debido a la naturaleza no estacionaria de los EEG, DWT es una transformada muy utilizada para su estudio [39].

Por un lado, si se tienen señales en el dominio del tiempo, se puede decir que se tiene información detallada de la señal en función del tiempo, pero no se tiene información de la señal en frecuencia. Por otro lado, si se utiliza la transformada de Fourier se conoce bien la señal en el dominio de la frecuencia, pero no en el dominio del tiempo.

La transformada wavelet es una función que permite obtener información de la señal, tanto en el dominio del tiempo y como de la frecuencia. Para ello, emplea regiones de tamaño variable en el análisis de las señales. Usa regiones grandes para bajas frecuencias, con la idea que, a bajas frecuencias la señal cambia a menor velocidad y por ello es necesaria menor información de la señal en el tiempo. Y regiones pequeñas para altas frecuencias, puesto

que a altas frecuencias la señal cambia más rápidamente en el tiempo, y es necesaria mas información en el dominio del tiempo, sin embargo, se pierde precisión en el dominio de la frecuencia [40].



*Figura 32 Resoluciones en tiempo y frecuencia de distintas transformaciones en comparación con una señal en el dominio del tiempo [41].*

Para la descomposición wavelet, se usará una señal que se proyectará sobre serie de funciones ortonormales. Estas funciones ortonormales se obtienen a partir de la *mother wavelet*  $\Psi_{a,b}(t)$  :

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \cdot \Psi\left(\frac{t-b}{a}\right)$$

Donde  $a, b \in R$ ,  $a \neq 0$  son los parámetros de escala y traslación, respectivamente y  $t$  el dominio del tiempo. Por tanto, a medida que aumenta  $a$  la función es más estrecha [42].

La transformada continua wavelet (CWT) de una función será el producto escalar de la señal a transformar, y la wavelet correspondiente.

$$W_{\Psi}(f)(a, b) = \langle f(t), \Psi_{a,b}(t) \rangle$$

La transformada discreta wavelet (DWT) será la transformada discreteada. Para la realizar la DWT se usa una serie de filtros paso altos y paso bajos. Donde el filtro paso alto es la *mother wavelete* discreta  $g(n)$  y el paso bajo es un filtro espejo  $h(n)$ . Las salida de los filtros paso alto y bajo, serán los coeficientes aproximados y detallados, respectivamente. Se representan como A1 y D1, y el proceso se repetirá sobre la salida A1 hasta obtener el nivel deseado [43].

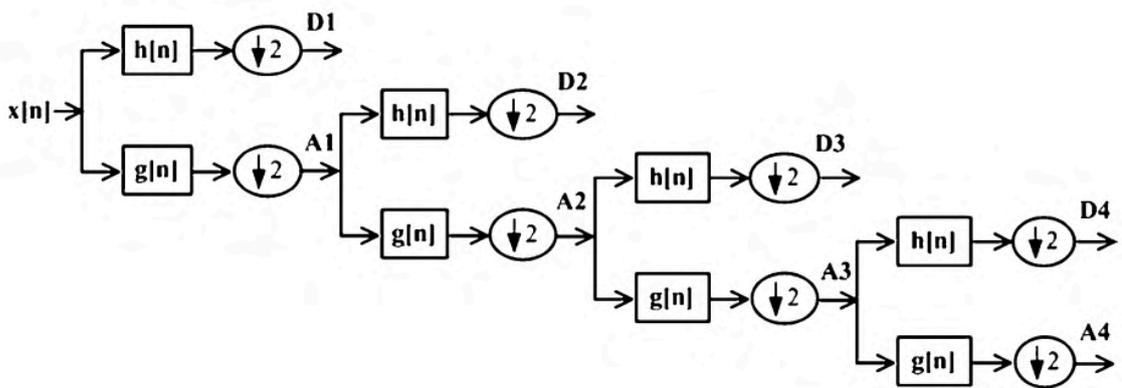


Figura 33 Descomposición de DWT, donde  $\downarrow 2$  significa muestreo reducido a la mitad

## 7.2 EXTRACCIÓN DE CARACTERÍSTICAS

En este proyecto, se realiza la DWT para obtener características de las señales del EEG, a distintas frecuencias. Para ello, se usará una *mother wavelete* daubechies (db4), puesto que se ha visto que proporciona buenos resultados para señales de EEG [44]. En Figura 34 se puede ver la *mother wavele* (*d4*), y la función de escalado.

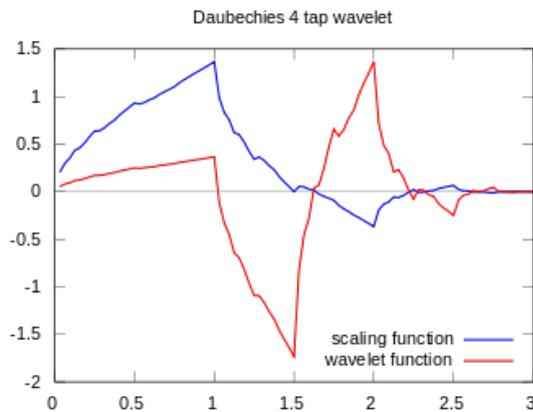


Figura 34 Mother wavelete para daubechies 4, y la función de escalado [45].

Se realizó una DWT de nivel cinco sobre las señales de cada canal previamente filtradas entre 0.05-50 Hz y divididas en distintos epoch. Con esto se obtienen los coeficientes detallados en distintos rangos de frecuencia que se ven en la Tabla 9, que corresponden aproximadamente con rangos de frecuencia de interés; el nivel 5 coincide aproximadamente con el rango de frecuencias delta (1-3)Hz; el nivel 4 coincide aproximadamente con el rango de frecuencias theta (3.5-8)Hz; el nivel 3 coincide aproximadamente con el rango de frecuencias alpha (8-13)Hz; el nivel 2 coincide aproximadamente con el rango de frecuencias beta (13-25)Hz; el nivel 1 coincide aproximadamente con el rango de frecuencias gamma (25-100)Hz;

| SEÑAL   | RANGO DE FRECUENCIA |
|---------|---------------------|
| NIVEL 1 | 25-50 Hz            |
| NIVEL 2 | 12,5-25 Hz          |
| NIVEL 3 | 6,25-12,5 Hz        |
| NIVEL 4 | 3,125-6,25 Hz       |
| NIVEL 5 | 0-3,125 Hz          |

*Tabla 9 Niveles de la descomposición DWT*

Una vez se obtienen los coeficientes para cada rango de frecuencias se calcula la *energía* de cada nivel para cada epoch y canal:

$$E_j = \sum_n |x[n]|^2$$

A partir de esto somos capaces de calcular la *energía total* de la señal:

$$E_{tot} = \sum_j E_j$$

Con esto normalizamos los valores y obtenemos la *energía wavelete relativa* para cada nivel [42]:

$$P_j = \frac{E_j}{E_{tot}}$$

Siendo esta la distribución de la probabilidad de la energía, que servirá para caracterizar algunos fenómenos en tiempo y frecuencia.

A continuación, se calcula la *entropía* de Shannon [42], [46]:

$$W_e = - \sum_j P_j \cdot \ln(P_j)$$

Donde  $W_e$  sirve como medida del orden/desorden de la señal, que proporcionara información sobre la dinámica de la señal [42].

Finalmente, para la construcción del clasificador se usarán la energía de cada nivel para así tener variables que representan distintos rangos de frecuencia del cerebro, y por tanto distintas características. Además se usará la característica de entropía.

### 7.3 *CLASIFICADOR SUPPORT VECTOR MACHINE*

En este apartado se explicará el modelo de clasificador que se ha utilizado, y los resultados que se han obtenido. Una vez extraída las características de las señales que se explica en el apartado anterior, se procedió a realizar el modelo de clasificador, que será un clasificador SVM.

El objetivo de un clasificador SVM es encontrar el hiperplano de n-dimensiones (n: número de características) que mejor sea capaz de clasificar los objetos [47].

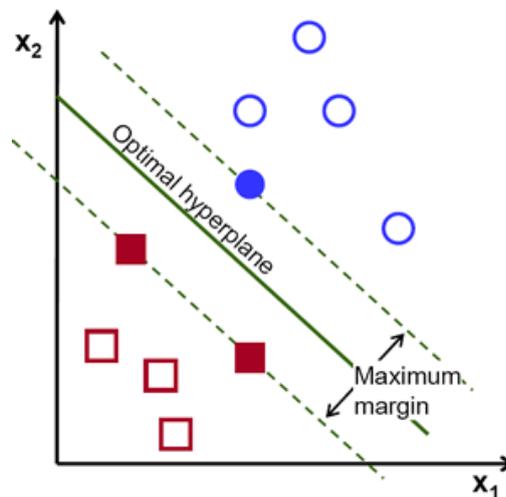


Figura 35 Ejemplo de clasificador SVM en 2-D.

Para ello SVM busca el hiperplano que maximiza la distancia de los objetos a clasificar a este. Es decir, los planos son líneas de decisión, clasificando variables en función del lado del plano en el que se encuentre los objetos.

Para este proyecto, y para la construcción del clasificador se ha utilizado los valores de la entropía de cada canal y la energía de cada nivel para cada canal. Se han realizado varias combinaciones de estas variables hasta obtener el clasificador con mejor exactitud. Además se ha ido optimizando el clasificador, variando parámetros de estos, con el objetivo de aumentar la exactitud. Por esto, se han realizado 3 clasificadores SVM, uno lineal, otro polinómico y finalmente uno radial, eligiendo finalmente aquel con mejores resultados.

Debido al reducido número de individuos a analizar, se ha utilizado como método de *cross-validation*, la forma *leave-one-out cross validation*. Donde para cada observación se realiza un modelo de clasificación y se compara la predicción del modelo de esa observación con el valor real de esta.

Una vez se han realizado los distintos modelos de clasificación se ha utilizado la curva AUC-ROC para compararlos y elegir el que ofrezca mejores resultados. En la se puede ver la curva para los distintos modelos.

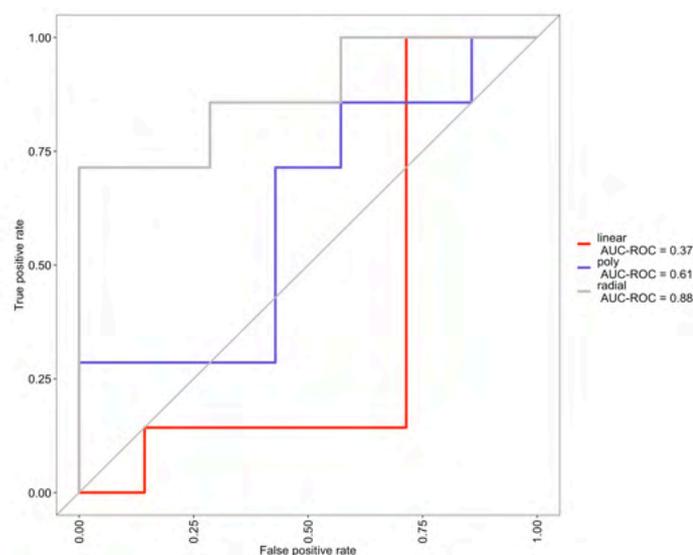


Figura 36 Curva AUC-ROC de los distintos modelos de clasificación.

Viendo la curva AUC-ROC, se observa que el mejor modelo de clasificación SVM es el radial. Obteniendo un área bajo la curva del 88%, unos resultados bastante buenos teniendo en cuenta la naturaleza de las señales. Asimismo, se observa que puesto que el SVM radial proporciona mejores resultados que el lineal. Esto muestra claramente que no hay una relación lineal entre el valor de la entropía de cada canal y el nivel de psicopatía de los sujetos.

## **Capítulo 8. CONCLUSIONES Y TRABAJOS FUTUROS**

En este apartado se expondrán las conclusiones que se han obtenido del proyecto, y se propondrá posibles trabajos futuros.

### **8.1 CONCLUSIONES**

En este proyecto se realizó un estudio sobre las señales cerebrales obtenidas de sujetos en un reformatorio, y se intentó realizar un modelo de clasificación que permitiese distinguir aquellos sujetos con rasgos HCU, y aquellos con rasgos LCU. Para ello se realizaron 3 estudios diferentes, primero se estudiaron las respuestas a la prueba ekman y se intentó ver si esto permitía distinguir a los individuos. En segundo lugar, se realizó un estudio de los potenciales relacionados con eventos y se compararon estas señales entre los distintos sujetos de estudio. Finalmente, se extrajeron características de las señales se trato de crear un modelo de clasificación automático.

Del estudio de las respuestas de la prueba Ekman se concluyó, que estas no permitían distinguir a los individuos. Esto se debió a que los resultados de las pruebas mostraron muchos fallos por parte de todos los sujetos, lo que supone que la prueba podría ser complicada. Además, se pudo ver que algunos individuos no entendieron bien la prueba, y sus respuestas parecían estar al revés, es decir, que presionaban el botón que no debían para cada imagen. A mayores, algunos sujetos realizaron la prueba sin seriedad, respondiendo a todas las cuestiones lo mismo o aparentemente aleatoriamente. También la dificultad para identificar las imágenes puede venir de la poca socialización a la que se enfrentan ambos grupos, y por ello la tasa de error en las respuestas es tan alta [32].

Con el estudio de los ERP de cada sujeto, se llegó a la conclusión de aquellos con rasgos HCU tenían mayor dificultad para procesar las imágenes. Y las señales del EEG para la componente N120 presentaban menor activación para el grupo HCU frente al grupo LCU.

Esta menor activación se puede ver a lo largo de toda la corteza cerebral, pero especialmente en la zona frontal, que es la encargada de las emociones.

Con la extracción de características, se llegó a la conclusión de que estas si permitían la diferenciación de sujetos, es decir, que las señales de cada sujeto son bastante diferentes. Pero aun así existe dificultad a la hora de construir un clasificador, debido a la naturaleza de las señales donde los cambios son muy sutiles, y a la pequeña muestra con la que se ha trabajado. Con todo ello, se consiguió crear un clasificador SVM, con forma radial que permite cierta clasificación de los individuos, con un área bajo la curva del 88%. Pero este modelo no permitiría sustituir las pruebas que existen actualmente, aunque un mayor desarrollo podría permitir su uso como herramienta complementaria.

## **8.2 TRABAJOS FUTUROS**

Como se ha visto una de las principales limitaciones de este proyecto es la falta de sujetos, por ello como proyecto futuros se propondría usar una muestra mucho mayor que permitiría obtener unos mejores resultados. Además, sería interesante realizar este estudio con individuos con mayor socialización y que no se encontrasen en un reformatorio. Esto aumentaría la muestra y se podría utilizar como grupo de control.

Por otra parte, sería imprescindible evaluar el nivel de implicación de los sujetos en el test (algunos de ellos responden completamente al azar). No obstante, el análisis de la entropía de los canales combinado con la técnica de clasificación SVM radial, arroja resultados prometedores que avalan el uso de técnicas de *Machine Learning* como herramienta de análisis forense de encefalografía. [12].

## Capítulo 9. BIBLIOGRAFÍA

- [1] J. L. F. Febber. [En línea]. Available: <https://historiadelamedicina.wordpress.com/2015/05/21/hans-berger-1873-1941-y-la-electroencefalografia>.
- [2] Edgar Adrian – Biographical. NobelPrize.org. , [En línea]. Available: <https://www.nobelprize.org/prizes/medicine/1932/adrian/biographical/>.
- [3] F. Schirmann, «“The wondrous eyes of a new technology”—a history of the early electroencephalography (EEG) of psychopathy, delinquency, and immorality,» *Frontier in Human Neuroscience*, vol. 8:232, 2014.
- [4] R. C. H. S. Diego, «<https://www.rchsd.org/>,» [En línea]. Available: <https://www.rchsd.org/health-articles/electroencefalograma-eeeg/>.
- [5] Mayo Clinic, [En línea]. Available: <https://www.mayoclinic.org/es-es/tests-procedures/eeg/about/pac-20393875>.
- [6] M. Yesuron, «La psicopatía y su diagnóstico,» vol. 1, nº 1, 2017.
- [7] S. López, «Revisión de la psicopatía: Pasado, presente y futuro.,» *Revista Puertorriqueña de Psicología*, vol. 24, nº 2, pp. 1-16, 2013.
- [8] J. I. R. S. Carolina López Magro, «Revista Puertorriqueña de Psicología,» *Psicopatología Clínica Legal y Forense*, vol. 5, pp. 137-168, 2005.

- [9] L. M. González, «Un estudio bibliométrico sobre la Psicopatía en Colombia,» Universidad Cooperativa de Colombia , 2019.
- [10] P. J. H. G. Víctor Dujo López, «LA PSICOPATÍA EN LA ACTUALIDAD: ABORDAJE CLÍNICO-LEGAL Y REPERCUSIONES FORENSES EN EL ÁMBITO PENAL,» de *Psicopatología Clínica, Legal y Forense*, vol. 17, Madrid, 2017, pp. 69-88.
- [11] M. M. A. Cárceles, «La inadecuada identificación de la psicopatía con el trastorno antisocial de la personalidad,» *Revista Electrónica de Ciencia Penal y Criminología*, pp. 2-50.
- [12] G. M. Jorge O. Folino, «La Psicopatía, el MMPI y la Hare PCL R,» *Revista Argentina de Clínica Neuropsiquiátrica*, vol. 13, nº 2, p. 43 a 54, 2006.
- [13] A. X. Stewart, A. Nuthmann y G. Sanguinetti, «Single-trial classification of EEG in a visual object task using ICA and machine learning,» *Journal of Neuroscience Methods*, vol. 228, pp. 1-14.
- [14] A. S. K. A. Mohammad H. Alomari, «Automated Classification of L/R Hand Movement EEG Signals using Advanced Feature Extraction and Machine Learning,» *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 4, 2013.
- [15] L. A. Gemein, R. T. S. P. C. D. W. J. B. A. S.-B. F. H. y T. B. , «Machine-learning-based diagnostics of EEG pathology,» *NeuroImage*, vol. 220, 2020.
- [16] OMS, [En línea]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/epilepsy>.

- [17] A. H. Shoeb, «Application of Machine Learning To Epileptic Seizure Detection,» Massachusetts Institute of Technology, Tesis Doctoral, 2009.
- [18] D. N. B.-L. L. Xiao-Wei Wang, «Emotional state classification from EEG data using machine learning approach,» *Neurocomputing*, 2014.
- [19] O. A. Zoubi, C. W. Ki, R. T. Kuplicki y A. Hung-wen , «Predicting Age From Brain EEG Signals—A Machine Learning Approach,» vol. 10, p. 184, 2018.
- [20] Consejo general de psicología de España, [En línea]. Available: [http://www.infocop.es/view\\_article.asp?id=8176](http://www.infocop.es/view_article.asp?id=8176) .
- [21] H. H. Jasper, P. Solomon y C. Bradley, «Electroencephalographic analyses of behavior problem children.,» *American Journal of Psychiatry*, vol. 95, nº 3, pp. 641-658, 1938.
- [22] D. Hill y D. P. , «Reflections on One Hundred Capital Cases Submitted to Electroencephalography,» *Journal of Mental Science*, vol. 98, nº 410, pp. 23-43, 1952.
- [23] K. Margaret, «The electroencephalogram and disorders of behavior,» *The Journal of Nervous and Mental Disease*, vol. 124, nº 2, pp. 102-124, 1956.
- [24] H. Baumgartl, F. Dikici, D. Sauter y R. Buettner, «Detecting Antisocial Personality Disorder Using a Novel Machine Learning Algorithm Based on Electroencephalographic Data,» 2020.
- [25] F. Hasanzadeh, M. M. y R. R. , «Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal,» *Journal of affective disorders*, vol. 256, pp. 132-142, 2019.
- [26] D. Raab y H. B. , «Machine Learning Based Diagnosis of Binge Eating Disorder Using EEG Recordings,» 2020.

- [27] J. Liu, H. M. A. N. y M. L. , «Emotion Detection from EEG Recordings,» *Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 1722-1727, 2016.
- [28] RStudio, «<https://www.rstudio.com/>,» [En línea].
- [29] B. Nicenboim, *eeguana*.
- [30] N. E. Helwig, *eegKit*.
- [31] Á. Gonzalo, «[machinelearningparatodos.com](http://machinelearningparatodos.com/),» [En línea]. Available: <https://machinelearningparatodos.com/salario-data-scientist-ingeniero-analista-datos/>.
- [32] L. Halty, «Impairment in the processing of fear gaze in adolescents with callous–unemotional traits,» *Psychology, Crime & Law*, vol. 25:8, n° 792-802, p. DOI: 10.1080/1068316X.2019.1588970 , 2019.
- [33] N. T. Alexandra Alda, «Bitbrain,» [En línea]. Available: <https://www.bitbrain.com/es/blog/colocacion-electrodos-eeg>.
- [34] cerebr25bg, «cerebrosaludable,» [En línea]. Available: <https://cerebrosaludable.com/blog/nuestro-sistema-nervioso/>.
- [35] J. Himberg, J. Mantyjarvi y P. Korpipaa, «Using PCA and ICA for exploratory data analysis in situation awareness,» *Conference Documentation International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp. 127-131, 2001.
- [36] B. López, A. Úbeda, D. Planelles, P.-V. E. Iáñez y J. M. Azorín, «Aplicación de ICA para la eliminación de artefactos oculares y mandibulares en señales EEG durante la realización de tareas mentales.,» *Actas de las XXXV Jornadas de Automática*, pp. 3-5, 2014.

- [37] S. Mozaffar y D. W. Petr, «Artifact Extraction from EEG Data Using Independent Component Analysis,» Information Telecommunication and Technology Center, University of Kansas, Lawrence, 2002.
- [38] E. A. Sciotto y E. B. Niripil, «Ondas cerebrales, conciencia y cognición,» Organización para la Prevención y Promoción de la Salud en la Educación, 2018.
- [39] U. Orhan, M. Hekim y M. Ozer, «EEG signals classification using theK-means clustering and a multilayerperceptron neural network model,» *Expert Systems with Applications*, vol. 38, nº 10, pp. 13475-13481, 2011.
- [40] B. J. G. Menéndez, E. M. Ambrona y R. M. Fraile, «Optimización de la transformada Wavelet discreta,» 2005.
- [41] «<https://ataspinar.com/2018/12/21/a-guide-for-using-the-wavelet-transform-in-machine-learning/>,» 2020. [En línea]. [Último acceso: 11 07 2021].
- [42] O. A. Rosso, S. Blanco, J. Yordanova, V. Kolev, A. Figliola y M. Schürmann, «Wavelet entropy: a new tool for analysis of short duration brainelectrical signals,» *Journal of neuroscience methods*, vol. 105, nº 1, pp. 65-75, 2001.
- [43] H. U. Amin, A. S. Malik, R. F. Ahmad, N. Badruddin, N. Kamel, M. Hussain y W.-T. Chooi, «Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques,» *Australasian physical & engineering sciences in medicine*, vol. 38, nº 1, pp. 139-149, 2015.
- [44] D. N. ,.-L. L. Xiao-Wei Wang, «Emotional state classification from EEG data usingmachine learning approach,» vol. 129, pp. 95-106, 2014.
- [45] Handwiki, «[https://handwiki.org/wiki/Daubechies\\_wavelet](https://handwiki.org/wiki/Daubechies_wavelet),» [En línea].

- [46] U. K. Diego Galar, «Shannon Entropy,» eMaintenance, 2017.
- [47] R. Gandhi, «towards data science,» [En línea]. Available: <https://towardsdatascience.com/support-vector-machine-introduction-to-machine-learning-algorithms-934a444fca47>.
- [48] B. H. J. V. J. R.-D. F. J. Pérez, «Análisis de contraste de la estructura factorial de la PCL-R: ¿qué modelo se ajustamejor a los datos?,» vol. 7, nº 1, pp. 13-30, 2015.
- [49] L. Konicar, S. R. S. S. E. B. R. V. U. S. C. V. P. L. P. y L. P. , «Balancing the brain of offenders with psychopathy? Resting state EEG and electrodermal activity after a pilot study of brain self-regulation training,» *Plos one*, vol. 15, nº 1, 2016.

## ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS

Los objetivos de desarrollo sostenible son una serie de objetivos adoptados por los estados miembros de la ONU en 2015 con la idea de proteger el planeta y garantizar que todas las personas gocen de paz y prosperidad. Se tratan de 17 objetivos, que se pueden ver en la Figura 37.



Figura 37 Objetivos de desarrollo sostenible

Este proyecto se alinea principalmente con 2 de ellos. Primero, con el objetivo de salud y bienestar, cuyo fin es promover el bienestar y una vida sana para todas las personas. Este proyecto busca crear una herramienta capaz de ayudar a identificar aquellos sujetos que presentan rasgos psicopáticos, principalmente en adolescentes. Todo ello para poder tratar esta enfermedad y mejorar la calidad de vida de estos individuos.

En segundo lugar, el proyecto se alinea con el objetivo de reducir las desigualdades, puesto que entender mejor el comportamiento de los adolescentes en reformatorios, podría facilitar

*ANEXO I: ALINEACIÓN DEL PROYECTO CON LOS ODS*

---

el desarrollo de métodos que permitiesen su reincorporación en la sociedad, es decir, darle las mismas oportunidades que el resto de la sociedad.

Finalmente, se puede ver que este proyecto busca mejorar la identificación y tratamiento de individuos con rasgos de psicopatía. Asimismo, identificar y entender mejor el comportamiento de los adolescentes en reformatorios, con la idea de facilitar su reinserción a la sociedad.