

Proyecto Fin de Carrera

Desarrollo de un sistema de detección de
espurios a partir del hardware del anillo de
Usenns ®

Autor

Daniel Estepa Sanz

Director

Eduardo Mariano Horna Prat

Ponente

Luis Montesano del Campo

Escuela de Ingeniería y Arquitectura

Febrero 2015

*A mis padres, por su esfuerzo y sacrificio en el pasado
y el presente para proporcionarme un futuro*

Resumen

La detección de cambios en el estado fisiológico y psicológico de una persona, provocados por un estímulo concreto, es un campo que está siendo ampliamente explorado en los últimos tiempos, debido a sus posibles aplicaciones en ámbitos como el de la salud o el marketing. Para ello, se necesitan datos que sean recogidos por sensores biométricos, los cuales cada vez están haciéndose más ligeros y precisos. Dos de estos biosensores son los llamados **GSR** (Galvanic Skin Response), que registra cambios en la conductancia de la piel originados por la activación y desactivación de las glándulas sudoríparas y **BVP** (Blood Volume Pulse), un detector del ritmo cardíaco mediante luz infrarroja. En este proyecto se utilizarán estos dos sensores, con el objetivo de desarrollar un sistema capaz de detectar anomalías de tipo fisiológico y psicológico inducidas en un sujeto. Para ello, se creará una base de datos de estos sujetos formada por medidas normales, donde no se les haya inducido ningún tipo de estrés y, a partir de esa base de datos se creará un clasificador capaz de dictaminar si una nueva muestra es normal o anómala.

Se comenzará decidiendo si pueden utilizarse unos sensores desarrollados por **BitBrain Technologies** implantados en un dispositivo llamado **Anillo de Usenns**. Para ello serán comparados con una tecnología similar de la marca **TMSI®**, cuyos sensores, de tipo médico, son ampliamente utilizados en mediciones de laboratorio con un alto rendimiento en la medición de estas señales biológicas. Una vez realizado este paso se readaptarán los sensores elegidos de forma que se ajusten a los requisitos del sistema que se quiere crear, creándose para ello un nuevo hardware y software.

Siguiendo un protocolo de toma de datos periódicas se obtendrá una base de datos de varios sujetos, que serán consideradas como medidas normales. Por otro lado, se realizarán a los mismos sujetos pruebas de tipo físico y cognitivo en las que se tomará una medida también, considerada anómala. Una vez recopilados todos los datos, se extraerán varias características de las señales, que serán tratadas de diversas formas para dar lugar a varios posibles clasificadores. Una vez desarrollados los clasificadores, éstos serán comparados y evaluados para determinar cuáles son los más adecuados dependiendo de la aplicación para la que se vayan a utilizar, y se seleccionará al más adecuado para el detector de anomalías fisiológicas y psicológicas que se quiere implementar. También será estudiado el rendimiento del sistema en el caso de no disponer de una base de datos previa del sujeto cuya muestra se quiere clasificar.

Para finalizar, serán detallados los resultados y conclusiones extraídos de este trabajo, así como las posibles aplicaciones y mejores futuras.

Índice

1. Introducción	1
1.1. Contexto	1
1.2. Estado del arte	1
1.3. Motivación	2
1.4. Objetivos	2
1.5. Organización de la memoria	4
1.6. Herramientas	5
2. Consideraciones previas de interés	7
2.1. Introducción	7
2.2. Búsqueda de patentes	7
2.3. Señales utilizadas	7
2.3.1. GSR	8
2.3.2. BVP	9
2.4. Sensores disponibles	9
2.4.1. TMSI	10
2.4.2. BitBrain	10
3. Comparativa de sensores	13
3.1. Introducción	13
3.2. Calibración de los sensores	13
3.2.1. Calibración del GSR	13
3.2.1.1. Interfaz Piel-Electrodo	14
3.2.2. Calibración del BVP	15
3.3. Medición experimental	16
3.4. Comparativa de señales	17
3.4.1. Correlación	17
3.4.2. RMSE	18
3.5. Elección de sensores	19
4. Toma de datos	21
4.1. Introducción	21
4.2. Requisitos del sistema	21
4.3. Hardware	22
4.4. Software	23
4.5. Protocolo de toma de datos	24
4.5.1. Tamaño de la muestra	24

4.5.2.	Toma de muestras normales	24
4.5.3.	Toma de anomalías	24
5.	Análisis	27
5.1.	Introducción	27
5.2.	Preprocesado	27
5.2.1.	BVP	27
5.2.2.	GSR	28
5.3.	Extracción de características	28
5.3.1.	BVP	28
5.3.2.	GSR	29
5.4.	Reducción de características	29
5.5.	Clasificadores	30
5.5.1.	Distancias	30
5.5.2.	Objetivo del clasificador	31
5.5.3.	Selección de clasificadores	33
6.	Resultados	35
6.1.	Introducción	35
6.2.	Método de evaluación	35
6.2.1.	Curva ROC	35
6.2.2.	Cross-validation	36
6.3.	Evaluación y comparación de clasificadores	37
6.4.	Mejoras	40
6.4.1.	Tiempo de grabación	40
6.4.2.	Número de muestras normales	40
6.5.	Comparación por tipo de anomalía	41
6.6.	Resultados para un nuevo sujeto	42
6.7.	Resumen de resultados	43
7.	Conclusiones y líneas futuras	45
	Bibliografía	47
A.	Evolución temporal del proyecto	49
A.1.	Hitos conseguidos durante el proyecto	49
A.2.	Diagrama de Gantt	50
B.	Creación de hardware y software	51
B.1.	Hardware	51
B.1.1.	Requisitos	51
B.1.2.	Diseño	51
B.1.3.	Construcción del prototipo	52
B.2.	Software	54
B.2.1.	Requisitos	54
B.2.2.	Uso del programa	55

C. Métodos matemáticos	59
C.1. Métodos para la comparación de señales	59
C.1.1. Coeficiente de correlación de Pearson	59
C.1.2. p-valor	60
C.1.3. RMSE	60
C.2. Algoritmos de transformación de características	61
C.2.1. Análisis de componentes principales	62
C.3. Algoritmos de selección de características	63
C.3.1. Criterio de separación	63
C.3.2. Búsqueda exhaustiva	64
C.3.3. SFS	65
D. Búsqueda de patentes	67
D.1. Búsqueda en castellano	67
D.1.1. Bases de datos	67
D.1.2. Parámetros de búsqueda	67
D.1.3. Patentes encontradas	68
D.1.3.1. Ejemplos	68
D.2. Búsqueda en inglés	69
D.2.1. Bases de datos	69
D.2.2. Parámetros de búsqueda	69
D.2.3. Patentes encontradas	69
D.2.3.1. Ejemplos	70
D.3. Conclusiones	71

Índice de figuras

1.1. Ejemplo de anomalía	3
1.2. Fases del proyecto	4
2.1. Señal de GSR descompuesta	8
2.2. Señal de BVP	9
2.3. Componentes utilizados de TMSI	11
2.4. Anillo de Usenns	11
3.1. Calibración del GSR con resistores	14
3.2. Señales de GSR para ambos sensores	15
3.3. Artefactos de la señal de BVP	16
3.4. Tareas del experimento realizado	17
3.5. Correlación total de las señales	18
3.6. Correlación por tramos de las señales	18
3.7. NRMSE de las señales	19
4.1. Prototipo para el detector de anomalías	22
4.2. Instantánea de la interfaz del programa creado	23
5.1. Señal de BVP tras el preprocesado	28
5.2. Ejemplo de mal clasificador	32
5.3. Ejemplo de buen clasificador	32
6.1. Curva ROC	36
6.2. validación cruzada de K-iteraciones	37
6.3. Curvas ROC para todos los clasificadores	37
6.4. Curvas ROC para los mejores clasificadores	38
6.5. Puntos óptimos de los clasificadores	39
6.6. Evolución de la AUC según el tiempo de análisis	40
6.7. Evolución de la AUC según el número de muestras normales	41
6.8. ROC para cada tipo de anomalía	41
6.9. ROC para un nuevo sujeto	42
A.1. Diagrama de Gantt del proyecto	50
B.1. Esquema del hardware	52
B.2. Construcción del prototipo	53
B.3. Ventana principal del programa	56

B.4. Selección de las distintas opciones	56
B.5. Comienzo de la toma de medidas	57
B.6. Visor de las señales	57

Índice de tablas

3.1. Valores de los resistores con TMSI y el Anillo	14
4.1. Requisitos del sistema detector de anomalías	21
5.1. Clasificadores desarrollados	33
B.1. Requisitos del hardware	51
B.2. Requisitos del software	54

1. Introducción

1.1. Contexto

Este proyecto ha sido realizado en la *start-up* **BitBrain Technologies**, *spin-off* de la Universidad de Zaragoza, cuyos campos de trabajo se centran principalmente en el desarrollo de dispositivos con sensores biométricos para su aplicación en diversos ámbitos como la neurociencia y el neuromarketing. Eduardo Horna, responsable del departamento de Tratamiento de Señal de la empresa, ha sido el director del proyecto, con la supervisión de Javier Mínguez, CSO de la empresa y profesor en excedencia de la Universidad de Zaragoza. El ponente ha sido Luis Montesano, miembro del departamento de Informática e Ingeniería de Sistemas de la Universidad de Zaragoza. El periodo de estancia en la empresa ha sido de siete meses, desde comienzos de Julio 2014 hasta finales de Enero 2015.

1.2. Estado del arte

Rosalind W. Picard presentó en 1995 a la comunidad científica [1] el término *Affective Computing* o *Informática Emocional*, definida como *computación relacionada, derivada o influenciada a partir de las emociones*. En este artículo Picard explicó con diversos ejemplos la necesidad de que en la interfaz humano-computador los segundos sean capaces de identificar las emociones de los primeros. Dada la gran importancia del componente emotivo en todos los ámbitos de la vida humana, la *Informática Emocional* no solo permitiría una mejora en esta relación entre humanos y máquinas, sino que podría llegar a dotar a los ordenadores de *emociones* que mejorarían su toma de decisiones. En artículos posteriores como [2], Picard abordó este tema ya desde un punto de vista más técnico, desarrollando protocolos y técnicas para la identificación de emociones humanas a partir de diversos sensores biométricos.

Desde entonces han sido muchos los productos desarrollados relacionados con la identificación de emociones humanas. Además, la aparición de dispositivos *wearables* (incorporados en alguna parte del cuerpo o la ropa) que permiten tomar datos de las personas de forma instantánea y masiva, ha ocasionado un incremento de este tipo de productos en los últimos años. Aunque sus ámbitos de aplicación son numerosos, dos son principalmente los campos en los que más aparecen: *e-health* (cuidados sanitarios apoyados en tecnologías de la información y las comunicaciones) y *neuromarketing*, disciplina que busca recoger mediante tecnología las reacciones y emociones que genera un recurso publicitario en un sujeto.

Los sensores utilizados para el desarrollo de estos productos varían según la aplicación, pero generalmente suelen ser algunos de los que se nombran a continuación: **GSR** (*Respuesta Galvánica de la Piel*), **BVP** (*Volumen de Pulso Sanguíneo*), **AR** (*Actividad Respiratoria*), **EEG** (*Electroencefalograma*), **fMRI** (*Resonancia Magnética Funcional*), **EMG** (*Electromiograma*) y **Eye-Tracker** (*rastreador del movimiento ocular*).

La empresa **BitBrain Technologies** ha desarrollado varios productos de este tipo como son la tecnología *Brain-Up*, que utiliza un casco de EEG con el objetivo de conseguir una mejora cognitiva del sujeto o el *Anillo de Usenns*, que cuenta con sensores de GSR y BVP y puede ser utilizado para la medición de emociones en diversos ámbitos.

1.3. Motivación

BitBrain Technologies ha desarrollado unos sensores de GSR y BVP conjuntamente con una placa de adquisición de datos cuyas aplicaciones son muy variadas, como las ya nombradas en la sección 1.2 de e-health y neuromarketing. Otra aplicación que la empresa estaba interesada en estudiar, y que será la desarrollada en este proyecto, era la de identificar si una persona está o no por encima de su nivel normal de estrés, con lo que el objetivo sería la creación de un dispositivo capaz de identificar este estado anómalo en un sujeto. Para ello, se crearían situaciones generadoras de estrés fisiológico y psicológico, siendo éstas las que el dispositivo debería detectar.

En caso de éxito en el desarrollo de este sistema el objetivo a largo plazo, fuera del ámbito de este proyecto, sería su aplicación en la detección de anomalías más concretas que pudieran interesar a un cliente, dado que empresas externas ya habían manifestado su interés por la aplicación de este dispositivo en ámbitos más específicos.

Como punto de partida se planteó la creación de un hardware con sensores de GSR y BVP con el que pudieran grabarse las señales biológicas captadas por estos sensores de forma fácil y rápida, solamente apoyando la mano sobre el dispositivo. Los sensores a utilizar deberían ser evaluados para estudiar si aquellos desarrollados por BitBrain eran adecuados para su utilización o, por el contrario, era necesario utilizar unos ya contrastados.

1.4. Objetivos

El objetivo de este proyecto fin de carrera es el desarrollo de un sistema capaz de detectar datos espurios, tratándose éstos de anomalías inducidas de tipo fisiológico y psicológico. Para ello se recopilarán muestras de varios sujetos en un estado normal, en el que no se introduzca ningún tipo de estrés. Tras realizar un tratamiento y análisis de estas señales, se generarán varios clasificadores que, ante una nueva muestra, sean capaces de identificarla como **normal** o **anómala**, con el mayor índice de acierto posible. Por lo tanto se buscará maximizar las diferencias entre muestras normales y anómalas, de forma que estas últimas sean lo más fácilmente identificables posible para el clasificador.

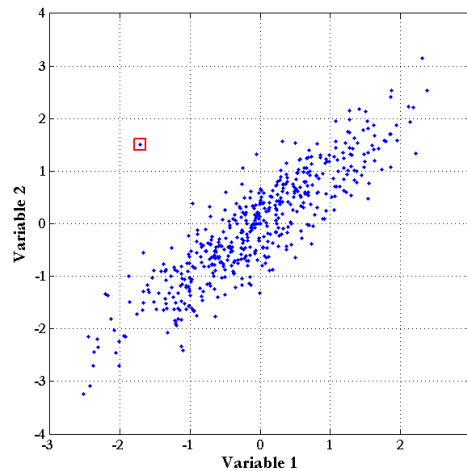


Figura 1.1: Ejemplo de dato anómalo o espurio: El objetivo es llevar nuestras anomalías a un espacio donde sean tan fácilmente identificables como el punto con el recuadro rojo. Tomando cualquiera de las dos variables por separado ese punto no es distinguible del resto de valores, pero utilizando ambas se convierte en un espurio fácilmente identificable.

Además, dado que se ha planteado como posible objetivo futuro la introducción de este dispositivo en un entorno comercial, es imprescindible realizar en primer lugar un estudio de patentes que refleje la presencia actual en el mercado de algún producto de características similares.

Los objetivos parciales que deben cumplirse par la creación del sistema detector de anomalías son:

- Búsqueda de patentes existentes similares al dispositivo que se quiere desarrollar.
- Selección de los sensores a utilizar.
- Desarrollo de la instrumentación (hardware y software) para la toma de medidas.
- Creación de una base de datos de medidas normales.
- Realización de dos tipos de anomalías inducidas en todos los sujetos: estrés fisiológico y estrés psicológico.
- Tratamiento de los datos recopilados y extracción de características.
- Creación de varios sistemas de clasificación mediante la combinación de esas características.
- Comparación de los diferentes clasificadores desarrollados.
- Evaluación de los resultados.

En la Figura 1.2 se representa un esquema de todas las tareas realizadas para la creación del sistema detector de anomalías.

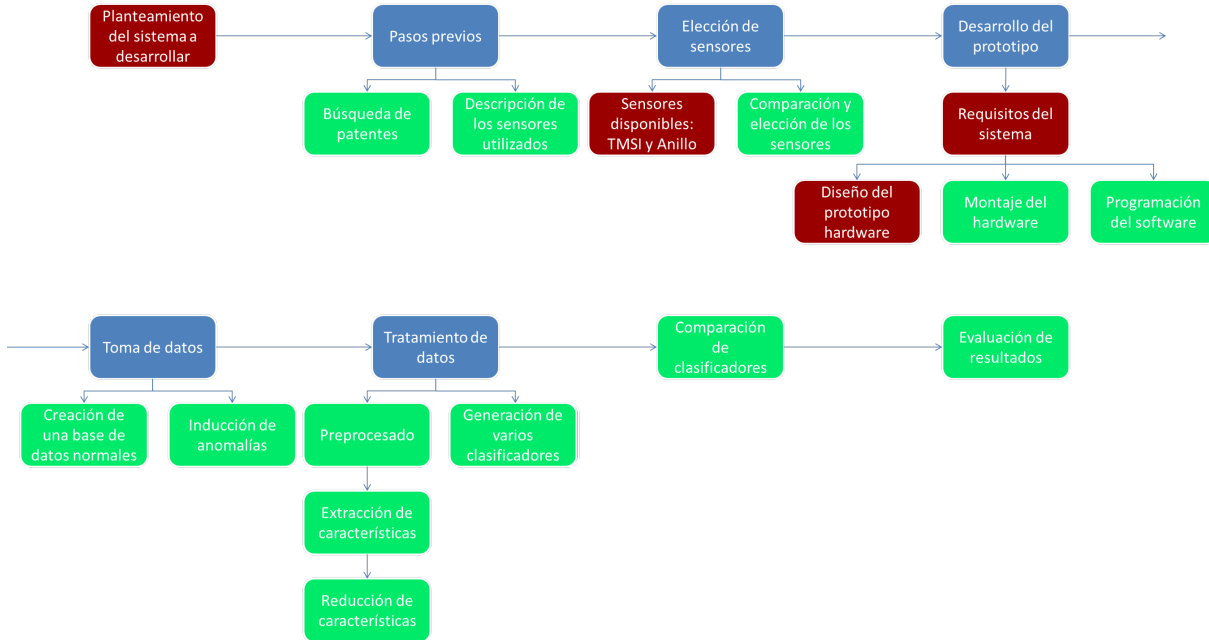


Figura 1.2: Fases del proyecto. En verde se representan las tareas realizadas en esta memoria, y en rojo aquellas que fueron realizadas por otras partes de la empresa.

1.5. Organización de la memoria

La memoria ha sido dividida en siete capítulos, siendo éste el primero donde se introducen el proyecto y sus contenidos. A continuación se explica el contenido del resto de capítulos:

- **Capítulo 2: Consideraciones previas de interés**, se explicará cómo ha sido realizado el estudio de patentes y serán presentados los sensores y señales de interés.
- **Capítulo 3: Comparativa de sensores**, los sensores desarrollados por BitBrain serán comparados con unos ya contrastados para elegir cuáles serán utilizados en el desarrollo del sistema.
- **Capítulo 4: Toma de datos**, se explicará la implementación del hardware y software necesario así como el protocolo seguido para la toma de datos normales y anómalos.
- **Capítulo 5: Análisis**, los datos recopilados serán tratados y se extraerán características de estos datos, que serán combinadas de varias formas para el desarrollo de los clasificadores.
- **Capítulo 6: Resultados**, los distintos clasificadores serán comparados y evaluados, exponiéndose los resultados más importantes.

Para finalizar la memoria en el **Capítulo 7: Conclusiones y líneas futuras**, se recopilarán los objetivos conseguidos y se nombrarán posibles mejoras y aplicaciones futuras.

1.6. Herramientas

La toma de medidas se realizó mediante la utilización del programa **Neurolab®**, desarrollado por **BitBrain Technologies**. Para el análisis de todos los datos se utilizó la herramienta matemática **MATLAB®**, y el software de interacción con el usuario fue realizado en **Python**.

2. Consideraciones previas de interés

2.1. Introducción

Como paso previo al desarrollo de un producto que podría salir al mercado en un futuro, es necesaria la realización de un estudio de patentes (sección 2.2) para documentarse sobre la existencia de productos similares al de interés. Por otro lado, es conveniente conocer un poco más en profundidad tanto las señales (sección 2.3) como los sensores (sección 2.4) que serán utilizados con posterioridad en el diseño.

2.2. Búsqueda de patentes

Antes de comenzar con el desarrollo de un producto es necesaria una búsqueda de productos similares ya existentes en el mercado por diversos motivos, como los que se exponen a continuación:

- Evitar la vulneración de patentes ya existentes, que de no ser tenidas en cuenta podrían impedir la comercialización del dispositivo una vez desarrollado.
- Como estudio de mercado, la constatación de un vacío de productos similares aumentaría las posibilidades de lograr un buen impacto comercial.
- Utilización de patentes existentes (en el caso de que las hubiera) para la introducción de mejoras o nuevos elementos en nuestro producto, siempre respetando esas patentes.

Por todas las razones arriba descritas, el primer paso para la creación del dispositivo detector de anomalías ha sido la realización de un estudio de patentes a nivel tanto nacional como internacional, que puede ser visto en detalle en el Anexo D. Esta investigación otorgó la conclusión de que ningún producto con el que se pudiera entrar en conflicto había sido desarrollado, lo que permite avanzar a la fase de desarrollo sin riesgo de vulnerar patentes existentes.

2.3. Señales utilizadas

Los sensores elegidos para la realización del sistema han sido los de **GSR** y **BVP**, que permiten medir la respuesta galvánica de la piel y el pulso sanguíneo respectivamente. A continuación

se realizará una explicación más detallada de cómo se toman y qué son exactamente las señales que registran estos sensores.

2.3.1. GSR

Varios nombres se utilizan para nombrar a este sensor, tales como **EDA** (*Electrodermal Activity*), **EDR** (*Electrodermal Response*), o el que será utilizado de aquí en adelante, **GSR** (*Galvanic Skin Response, Respuesta Galvánica de la Piel*). Su labor es la medición del nivel de activación de las glándulas sudoríparas, el cual ha sido directamente relacionado con variaciones en el estado físico o anímico de las personas, lo que ha llevado a su utilización de manera extendida en el campo de la neurociencia [3].

El sensor consta de dos electrodos que se colocan sobre la piel del sujeto. Su posicionamiento puede variar, aunque la colocación estándar suele ser en los dedos de la mano, donde los cambios de actividad en la sudoración son más acusados. Concretamente, suelen ser colocados en las segundas falanges de los dedos índice y corazón o corazón y anular.

Se trata de una medición **activa**, ya que se introduce una pequeña corriente entre los dos electrodos. De esta forma, se puede calcular el valor de la resistencia eléctrica de la piel, que varía según el nivel de activación de las glándulas sudoríparas. A partir de ahí puede ser extraído el valor de conductancia de la piel (medido en Siemens, $S = \Omega^{-1}$), o lo que es lo mismo, la inversa de la resistencia de la piel. El rango de esta señal suele estar en torno a los microSiemens (μS).

La señal de GSR puede ser dividida en dos partes:

- **Parte tónica:** Refleja las variaciones lentas producidas en la señal a lo largo del tiempo.
- **Parte fásica:** Muestra cambios rápidos de la señal que suelen ser producidos como respuesta a algún estímulo. Estas respuestas pueden ser de dos tipos: **Relacionadas con algún evento** (*ER, Event-Related*) cuando se deben a algún estímulo concreto que puede ser identificado, o **No Específicas** (*NSR, Non-Specific Responses*), cuando ocurren en ausencia de estímulos identificables.

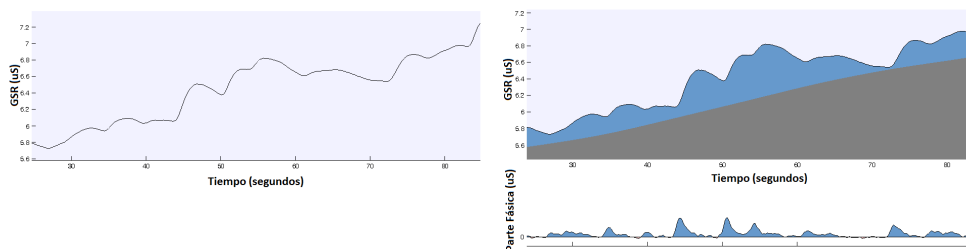


Figura 2.1: A la izquierda se ve una señal típica de GSR durante un intervalo de 60 segundos. A la derecha esa señal ha sido descompuesta en su parte tónica (zona sombreada de la gráfica superior) y fásica (gráfica inferior).

Dado que existe una relación entre las respuestas del *Sistema Nervioso Simpático (SNS)* y el nivel de *activación emocional*, el sensor de GSR puede ser utilizado para medir el nivel de excitación emocional de una persona, aunque no es posible la detección de la emoción concreta.

Como libro de referencia sobre la activación electrodérmica, su explicación fisiológica, conexión con el sistema nervioso y aplicaciones en diversos campos véase [4].

2.3.2. BVP

El sensor de BVP (*Blood Volume Pulse, Volumen de Pulso Sanguíneo*), también conocido como fotoplestismograma (*PPG*, por sus siglas en inglés), es capaz de extraer la onda de pulso sanguíneo, a partir de la cual se puede medir el ritmo cardíaco (Heart Rate (**HR**)) y la variabilidad del ritmo cardíaco (Heart Rate Variability (**HRV**)), parámetros relacionados directamente con el sistema nervioso autónomo [5], entre otras características.

Para ello el sensor es colocado típicamente en la primera falange del dedo índice (el lóbulo de la oreja puede ser otro buen lugar de posicionamiento), y una señal de origen óptico es registrada por el sensor. Dos componentes son utilizados: un LED que ilumina la piel y un fotodiodo que capta la luz. Este fotodiodo puede ser colocado al lado del LED, con lo que medirá la cantidad de luz **reflejada**, o al otro lado del dedo, captando la luz **transmitida** a través de todos los tejidos del dedo. Ambas colocaciones permiten la extracción de la onda cardíaca mediante los cambios en la absorción de la luz, que se deben a las variaciones en la cantidad de sangre que circula por el dedo a lo largo del ciclo cardíaco. El rango de la señal obtenida se encuentra en los μV .

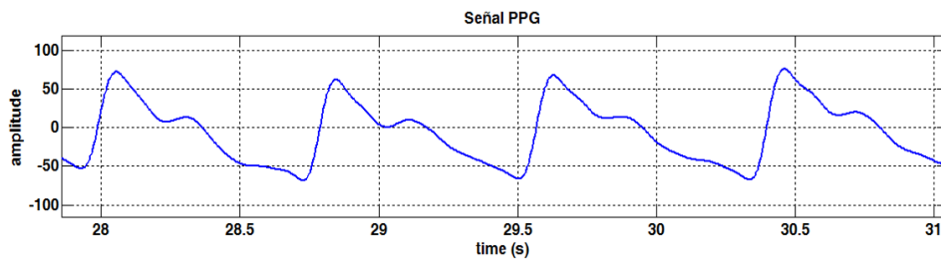


Figura 2.2: Señal de BVP.

2.4. Sensores disponibles

La fabricación de dispositivos para la medición de señales biológicas ha venido incrementándose en los últimos tiempos, especialmente unida a la aparición de las tecnologías *wearables*. Estos dispositivos pueden ser clasificados en dos categorías:

- **Instrumentación de laboratorio o médica:** pensada para realizar medidas en un entorno estático con alta fiabilidad y precisión. Este es el caso de la tecnología de **TMSI** o **GTec**.
- **Instrumentación portátil:** Su gran ventaja radica en que el sujeto puede estar en movimiento, permitiendo la toma de medidas en cualquier lugar y situación. Por ello, el tamaño y peso de este tipo de dispositivos debe ser mucho menor que los de tipo médico, con las consecuentes limitaciones de diseño que ello implica. Se trata del caso de los ya comentados *wearables*, siendo el **Anillo de Usenns** un ejemplo de este tipo de productos.

El detector de anomalías que queremos desarrollar deberá llevar los sensores ya descritos de GSR y BVP. La empresa **BitBrain** dispone de unos sensores de este tipo de reciente creación. Aunque lo preferible sería que utilizáramos éstos para el desarrollo de nuestro producto, lo cierto es que aún no han podido ser comparados contra otros de calidad contrastada para comprobar su correcto funcionamiento. Por ello, a continuación se describirán más en detalle tanto estos sensores como los de **TMSI**, y en el Capítulo 3 serán comparados y evaluados, decidiéndose cuáles serán usados para el desarrollo del sistema de interés.

2.4.1. TMSI

La empresa **TMSI** (*Twente Medical Systems International*) es una compañía especializada en el desarrollo de amplificadores y sensores de alta calidad para la adquisición de datos en aplicaciones fisiológicas. Su tecnología es ampliamente utilizada en entornos biomédicos y reconocida como precisa y fiable incluso en ambientes hostiles, con alta presencia de artefactos. La frecuencia de muestreo es de 256 Hz, y el amplificador va conectado mediante una interfaz USB a un ordenador para la transmisión de los datos.

En la figura 2.3 vemos el amplificador que se ha utilizado, el cual permite la conexión de múltiples tipos de sensores. Nosotros utilizaremos los de BVP y GSR aunque podrían ser conectados también otros tipos de sensores como EMG y hasta 32 canales de EEG.

2.4.2. BitBrain

La empresa BitBrain ha desarrollado dos sensores de BVP y GSR junto con una placa de adquisición de datos con frecuencia de muestreo de 32 Hz, pensados para su instalación en instrumentación de medida portátil. Uno de esos dispositivos, el cual se utilizará para la medición de calidad de los sensores es el **Anillo de Usenns**. Este dispositivo va sujeto a los dedos índice y corazón mediante una correa, permitiendo la adquisición cómoda de esas señales. Además, dispone de un acelerómetro que registra los movimientos del anillo y de la placa de adquisición de datos que transmite los datos a un ordenador vía Wi-Fi, lo que hace al aparato totalmente inalámbrico.

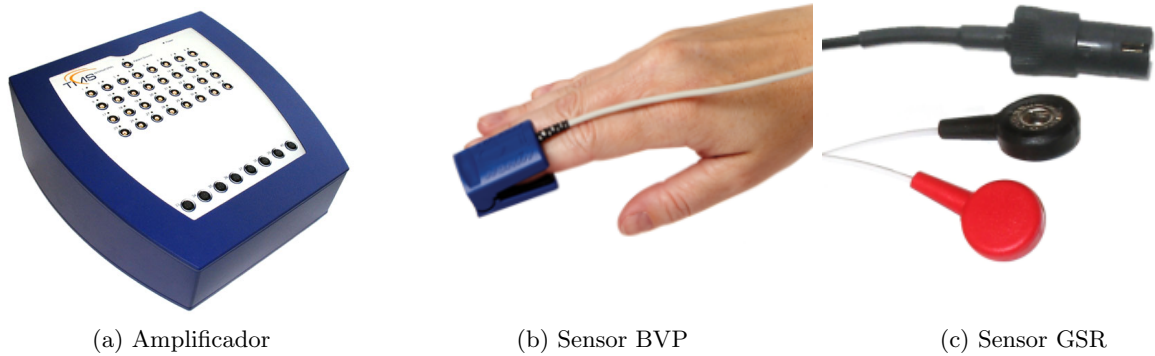


Figura 2.3: Componentes utilizados de TMSI: sensor BVP (2.3b), que es colocado en la primera falange del dedo índice y sensor de GSR (2.3c), formado por dos electrodos posicionados en las segundas falanges de los dedos corazón y anular. Ambos van conectados al amplificador de la figura 2.3a, que a su vez transmite los datos via USB a un ordenador.



Figura 2.4: Anillo de Usenns. En la figura podemos observar el sensor de actividad cardiovascular (BVP), que se sitúa en la primera falange del dedo índice, el de actividad electrodermica de la piel (GSR), cuyos electrodos se posicionan sobre las segundas falanges de los dedos índice y corazón, y el acelerómetro o sensor de movimiento. Además, incorpora un módulo Wi-Fi para el envío de datos.

3. Comparativa de sensores

3.1. Introducción

En el capítulo anterior se han presentado los dos tipos de sensores disponibles, unos de tipo médico de la marca **TMSI** y otros desarrollados por **BitBrain**. A lo largo de este apartado serán comparados con el objetivo de determinar si podemos utilizar los segundos para la implementación del detector de anomalías o, por el contrario, será necesaria la utilización de los sensores de TMSI.

La comparativa se ha basado en el artículo de referencia [6], donde un sensor portátil de GSR es comparado con uno de tipo médico. Se seguirá un esquema similar al del artículo citado, realizándose también una comparación de los sensores de BVP, la cual no tiene lugar en ese artículo.

Como primer paso se llevará a cabo una calibración de los sensores (sección 3.2) para garantizar su correcto funcionamiento, grabándose posteriormente una medida experimental descrita en la sección 3.3. Una vez que se hayan adquiridos las señales, éstas serán comparadas (sección 3.4), y se determinará qué sensores serán utilizados (sección 3.5), en base a los resultados obtenidos.

3.2. Calibración de los sensores

Antes de que se realice una grabación, es imprescindible comprobar el buen funcionamiento de los sensores y calibrarlos para que su rendimiento sea el mejor posible.

3.2.1. Calibración del GSR

Tal y como fue descrito en la sección 2.3.1, el GSR es un sensor que mide la resistencia de la piel. Por lo tanto, es posible realizar una calibración del sensor con resistores fijos cuyo valor es conocido. Se han realizado mediciones en el rango típico de la piel humana, que va desde los 100 k Ω hasta los 4 M Ω .

En la Tabla 3.1 puede observarse el error obtenido al medir los distintos valores resistivos con los sensores de TMSI y el Anillo. Este error se ha calculado con respecto al valor real de los resistores, medido con un polímetro de laboratorio de precisión 0.9 %.

Resist. teórica	Resist. real	Valor TMSI	Error TMSI	Valor Anillo	Error Anillo
100 k Ω	99.7 k Ω	100.076 k Ω	0.3780 %	90.995 k Ω	8.7309 %
200 k Ω	199.1 k Ω	199.059 k Ω	0.0204 %	191.033 k Ω	4.0515 %
300 k Ω	298.8 k Ω	298.223 k Ω	0.1931 %	290.924 k Ω	2.6358 %
400 k Ω	398.7 k Ω	397.578 k Ω	0.2814 %	392.011 k Ω	1.6776 %
512 k Ω	511 k Ω	509.151 k Ω	0.3619 %	505.976 k Ω	1.1782 %
1000 k Ω	995 k Ω	990.440 k Ω	0.4583 %	992.049 k Ω	0.2966 %
2000 k Ω	1990 k Ω	1975.736 k Ω	0.7168 %	1995.539 k Ω	0.2784 %
3000 k Ω	2986 k Ω	2957.455 k Ω	0.9560 %	2998.649 k Ω	0.4236 %
4000 k Ω	3986 k Ω	3937.312 k Ω	1.2215 %	4005.343 k Ω	0.4853 %

Tabla 3.1: Valores de los resistores con TMSI y el Anillo.

En la Figura 3.1 se puede observar el error que se obtiene en el artículo de referencia [6], en el que se realizó la misma prueba, y el obtenido con los sensores de TMSI y el Anillo. Para valores inferiores a 1 M Ω TMSI obtiene un error menor, mientras que partir de 1 M Ω el Anillo consigue un error muy por debajo del de TMSI, y bastante similar al del artículo de referencia. Los valores de resistencia de la piel humana suelen situarse siempre por encima de 1 M Ω , intervalo en el que el Anillo tiene un error muy pequeño.

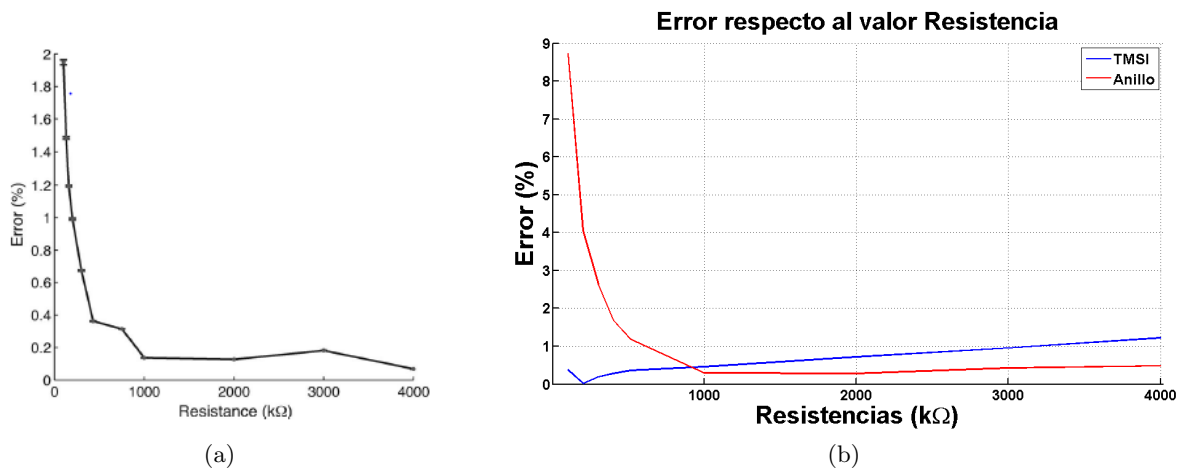


Figura 3.1: Curva de error para: 3.1a el artículo de referencia [6] y para 3.1b el anillo y TMSI.

Como resultado de esta prueba el departamento de electrónica pudo mejorar el funcionamiento del sensor ajustando alguno de los componentes del Anillo para minimizar el error, así como plantear futuras mejoras que pudieran minimizarlo aún más.

3.2.1.1. Interfaz Piel-Electrodo

La calibración del sensor de GSR ha sido realizada mediante resistores fijos. Sin embargo, antes de que estos sensores sean utilizados en una persona, algunas cuestiones de interés deben ser descritas:

- La interfaz Piel-Metal que se da al colocar los electrodos en la mano de una persona es mucho más compleja que la interfaz Metal-Metal realizada en la calibración con resistores [7]. Esta interacción es dependiente en gran medida del metal utilizado en el sensor, así como del tipo de piel (un dedo donde la piel es más dura tiene diferente resistencia que otro con la piel más blanda). De los dos sensores disponibles, el de TMSI utiliza acero como material de sus electrodos mientras que el del Anillo utiliza plata. Esto provoca diferencias en la señal obtenida (Figura 3.2), y una normalización de las señales se hace imprescindible para que éstas puedan ser comparadas.
- Ambos sensores no pueden ser colocados en los mismo dedos ya que se originarían interferencias que afectarían a las señales. La repetición del experimento con diferentes sensores tampoco es posible, debido a que la intensidad de la respuesta a los estímulos disminuye con el número de repeticiones, por lo que las señales no serían comparables. La solución radica en la colocación de los sensores en manos diferentes. Sin embargo, existen diferencias entre la mano izquierda y derecha [8], [9] a la hora de registrar una señal de GSR. Por lo tanto, no deben esperarse señales idénticas al grabar en manos distintas, incluso aunque los sensores fueran los mismos. Esto también puede visualizarse de forma gráfica en la Figura 3.2.

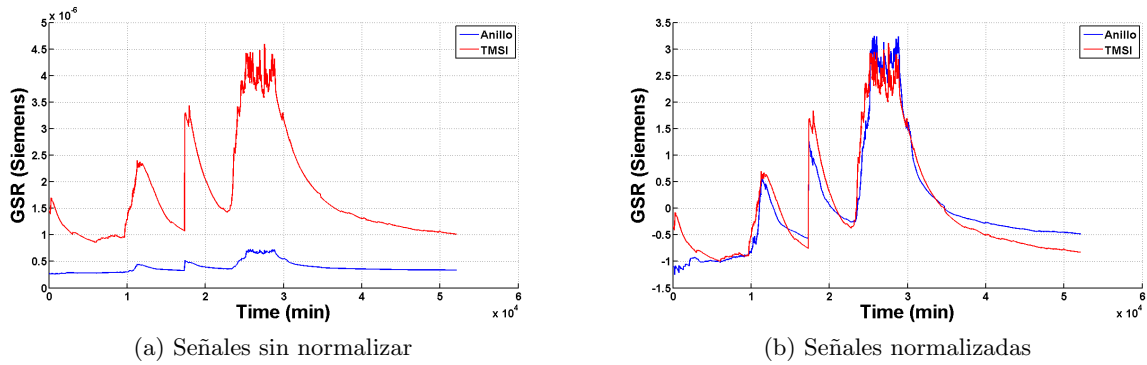


Figura 3.2: Una normalización de las señales de ambos sensores es necesaria debido al diferente metal utilizado por sus electrodos.

En definitiva, tanto el material del que están fabricados los sensores como la zona de colocación de éstos aportan una variabilidad a las señales que deberá ser tomada en cuenta a la hora de interpretar la comparación entre ellas.

3.2.2. Calibración del BVP

Debido a que el sensor de BVP realiza medidas ópticas, no es posible la realización de una calibración similar a la del GSR, donde la comparación se realizaba con un parámetro fijo conocido como eran los resistores. En este caso, la calibración se ha realizado en dos pasos: en primer lugar, se comprueba que el sensor está correctamente conectado y funcionando. Posteriormente, una vez colocado en el dedo del sujeto y con la ayuda de un visor que permita observar la señal en tiempo real, el sensor es ajustado hasta la obtención de una señal correcta. En la Figura 3.3

se muestra la forma que debe de tener la señal cuando el sensor está bien colocado así como alguno de los artefactos más comunes provocados por su mal posicionamiento.

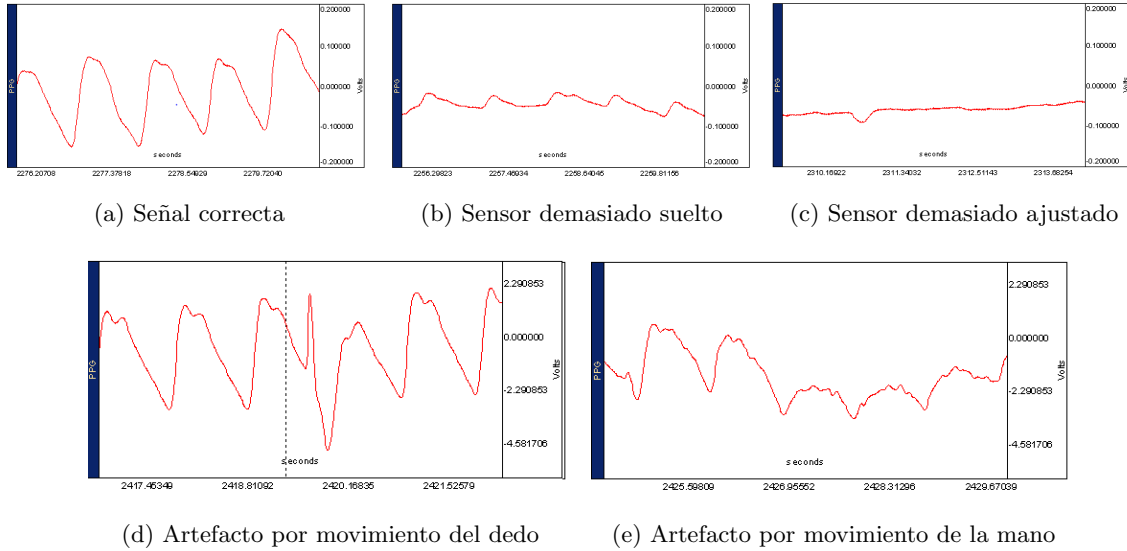


Figura 3.3: Señal de BVP correcta y afectada por algunos de los artefactos más comunes.

Desgraciadamente, aunque el sensor se haya ajustado correctamente al principio de la grabación, la aparición de artefactos por movimientos del dedo y mano del sujeto (Figura 3.3d y 3.3e) es muy común, especialmente en grabaciones largas como la que será realizada. Como consecuencia, una inspección visual de la señal grabada será necesaria con el objetivo de identificar y eliminar artefactos de las señales, lo cual implica que la longitud de éstas se verá reducida y no será exactamente la misma para todos los sujetos, ya que dependerá del número de artefactos encontrados.

3.3. Medición experimental

A continuación se describe el protocolo seguido para la grabación de una medida experimental:

- **Número de sujetos:** 5.
- **Colocación de los dispositivos:** Sensores de TMSI en la mano derecha tal y como se describe en la Figura 2.3 y anillo de Usenns en la mano izquierda (Figura 2.4). También se ha utilizado un fotodiodo de TMSI para la posterior sincronización de las señales, y dos ordenadores para la adquisición de datos de cada dispositivo.
- **Experimento:** se realizaron 3 tareas de 9 minutos cada una, formadas cada una por 3 minutos de relajación, 3 de una tarea física (pedaleo en una pedalina), cognitiva (test de Stroop, una prueba generadora de estrés cognitivo que explicaremos más en detalle en el capítulo 4) o emocional (ver una serie de vídeos), seguidos de otros 3 minutos de relajación. Un esquema general puede ser observado en la Figura 3.4.

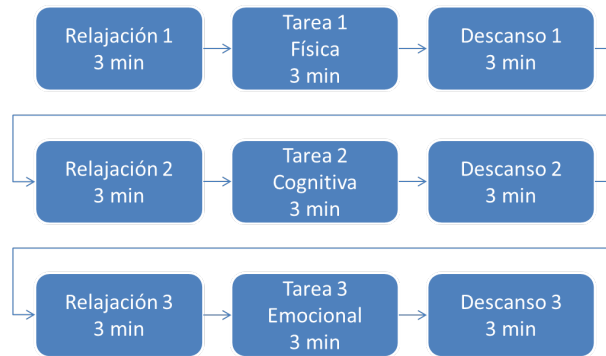


Figura 3.4: Tareas del experimento realizado.

3.4. Comparativa de señales

Para la comparación de las señales ha sido utilizada la misma métrica que en [6], el **Coeeficiente de correlación de Pearson**, parámetro que refleja el grado de relación entre dos variables cuantitativas. Como test estadístico ha sido calculado el **p-valor**, teniendo en cuenta que una correlación será significativa si su p-valor es menor a 0.05. En el caso de todas las correlaciones que serán mostradas en los siguientes apartados, su p-valor ha sido < 0.001 .

Una explicación más detallada del coeficiente de correlación de Pearson y del p-valor puede ser encontrada en el Anexo C.

Cabe señalar que, dado que la frecuencia de muestreo de ambos dispositivos es distinta tal y como se detalló en la sección 2.4, ha sido necesario realizar un diezmado de la señal de TMSI para reducir la frecuencia de muestro a 32 Hz antes de poder practicar cualquier comparación entre las señales.

3.4.1. Correlación

En la Figura 3.5 se presenta la correlación total de las señales recogidas para cada sujeto por ambos tipos de sensores. Sin embargo, la correlación de señales de larga duración y baja frecuencia puede dar como resultado un valor muy alto, pero pasar por alto zonas de peor correlación. Por ello, en la Figura 3.6 se han dividido las señales en tramos más pequeños, con el fin de comprobar si la correlación desciende excesivamente en algún punto concreto de la señal. Dado que después de la eliminación de artefactos la duración de las señales para cada sujeto no es idéntica, esta división se ha realizado en fragmentos de duración 20 % de la longitud total de cada señal.

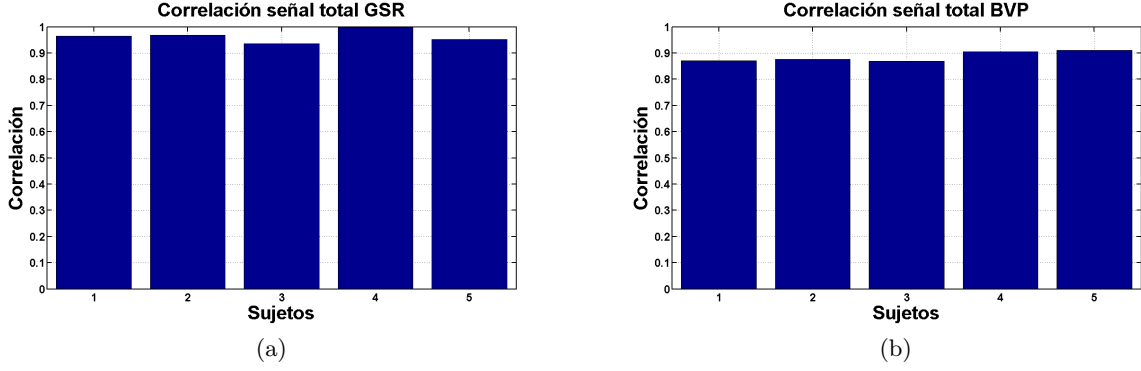


Figura 3.5: Correlación total de las señales de BVP y GSR por sujeto.

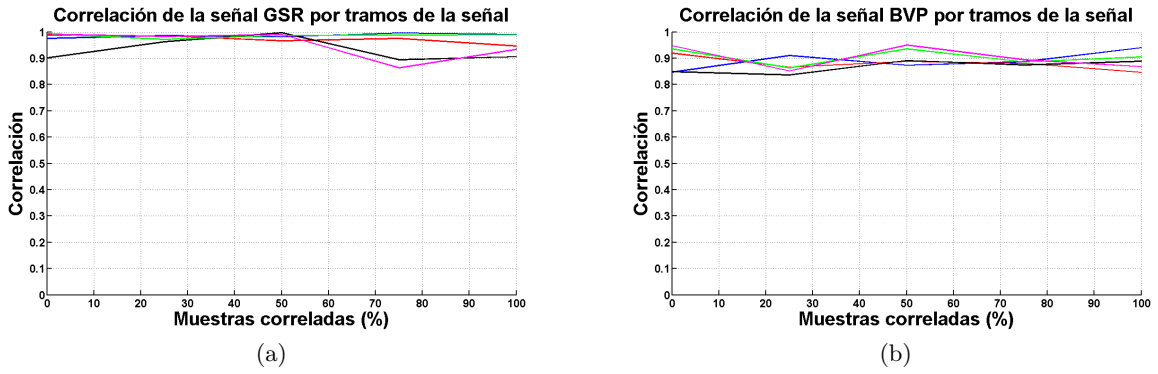


Figura 3.6: Correlación por tramos de las señales de BVP y GSR por sujeto, p-valor < 0.001.

3.4.2. RMSE

Otro parámetro de comparación de señales que suele ser utilizado es el **RMSE** (*root-mean square error*) o raíz del error cuadrático medio, que representa el error residual medio entre una señal y otra:

Se quiere calcular el error de la señal de n puntos recogida por el anillo respecto a la señal de referencia, que será la de TMSI. En este caso la ecuación para calcular el RMSE queda:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_{t,Anillo} - y_{t,TMSI})^2}{n}} \quad (3.1)$$

Si realizamos una normalización de este valor respecto al rango de valores observados en la variable de referencia obtenemos el **NRMSE**:

$$NRMSE = \frac{RMSE}{y_{max,TMSI} - y_{min,TMSI}} \quad (3.2)$$

Para ambos sensores el valor del NRMSE es muy bajo (Figura 3.7), siempre menor al 4 %, siendo aún más pequeño en la señal PPG donde como máximo llega al 0.1 %. El hecho de que el error sea más grande en la señal de GSR se debe a que ambas señales pueden tener diferencias debido al lugar de posicionamiento y material de los sensores tal y como se describió en la sección 3.2.1.1, y este parámetro castiga más a pequeñas diferencias puntuales entre las señales que otros como la correlación. Aun así pueden considerarse valores bastante pequeños que reflejan la gran similitud de las señales.

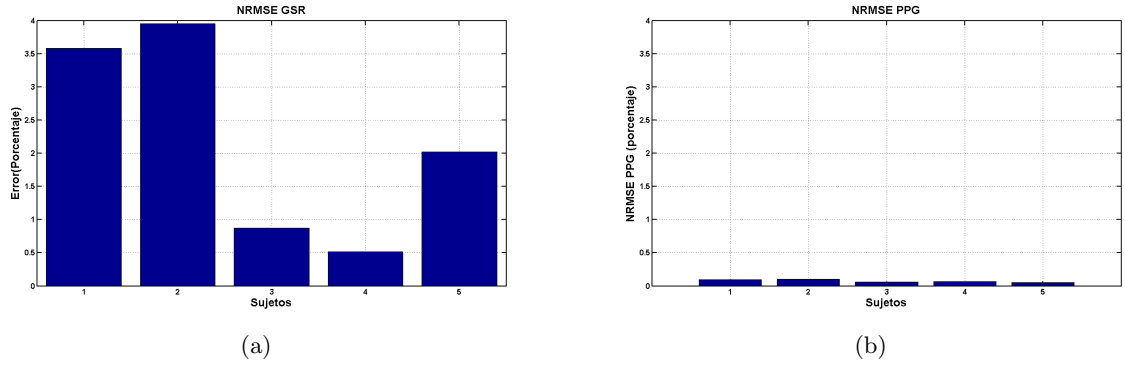


Figura 3.7: NRMSE de las señales de GSR y PPG.

3.5. Elección de sensores

Después de realizar las comparaciones anteriores entre las señales del anillo y de TMSI, se han extraído las siguientes conclusiones:

- Los altos índices de correlación apuntan que ambas señales son equivalentes a nivel global con un gran nivel de similitud, lo cual nos indica que estamos registrando los mismos procesos fisiológicos.
- Se observan porcentajes muy bajos en el NRMSE, con lo que podemos constatar que tampoco existen diferencias puntuales muy grandes entre las señales. En el PPG esto es muy importante ya que provocaría grandes desigualdades en el cálculo del ritmo cardíaco y otras características. Respecto al GSR estas diferencias puntuales son más asumibles, incluso normales, al tratarse de sensores fabricados con materiales distintos y colocados en manos diferentes, con lo que las señales diferirán ligeramente.

Considerando lo arriba comentado, las dos pruebas que hemos realizado nos confirman que los sensores del anillo y TMSI son equivalentes, habiéndose comprobado a nivel global y puntual. Por lo tanto, podemos afirmar que los sensores de BitBrain pueden ser utilizados en cualquier ámbito de aplicación con total garantía de que el registro de señal es comparable al de un dispositivo médico ya contrastado como TMSI.

4. Toma de datos

4.1. Introducción

Una vez que los sensores han sido elegidos, es posible comenzar con el desarrollo del detector de anomalías. A lo largo de este capítulo describe en primer lugar los requisitos de diseño que debe tener el dispositivo (sección 4.2), para posteriormente detallar como se han cumplido cada uno de estos requisitos, a nivel de hardware (sección 4.3) y de software (4.4). Para finalizar el capítulo se detalla el protocolo seguido para la medición de datos normales y anómalos en la sección 4.5.

4.2. Requisitos del sistema

El sistema detector de anomalías contaba con unos requisitos a nivel de hardware y software impuestos por la empresa que se debían cumplir. Además, también fueron fijados el número de sujetos y muestras que debían tomarse. Todos estos requisitos se resumen en la tabla 4.1, y serán descritos más detalladamente en las secciones posteriores.

	Requisitos
Hardware	Apoyo cómodo de la mano derecha
	Misma placa y sensores que los del Anillo de Usenns
	Sensor de BVP en la primera falange del dedo índice
	Electrodos de GSR en segundas falanges de dedos índice y corazón
	Tapa sobre los dedos que proporcione oscuridad al sensor de BVP
	Material y diseño que imposibilite movimientos indeseados
Software	Creación de un programa para la adquisición de datos
	Fácil e intuitivo de usar para un usuario estándar
	Etiquetado de las medidas por sujeto, momento del día y tipo
Toma de medidas	Dos tipos de medidas: normales y anómalas
	Toma de medidas normales dos veces al día: mañana y tarde
	70 medidas normales por sujeto
	6 anomalías por sujeto: 3 físicas y 3 anómalas

Tabla 4.1: Requisitos del sistema detector de anomalías.

4.3. Hardware

Para la creación de un prototipo que cumpliera con las especificaciones descritas de la Tabla 4.1, el departamento de diseño de la empresa generó un esquema que recogiera estos requerimientos. A partir de ese esquema se procedió al montaje del hardware, contándose con la colaboración del departamento de electrónica para el soldado y ensamblado de los cables y sensores. Se pueden ver el plano y el resultado final en la Figura 4.1, y una explicación más detallada del montaje del prototipo puede encontrarse en el Anexo B.

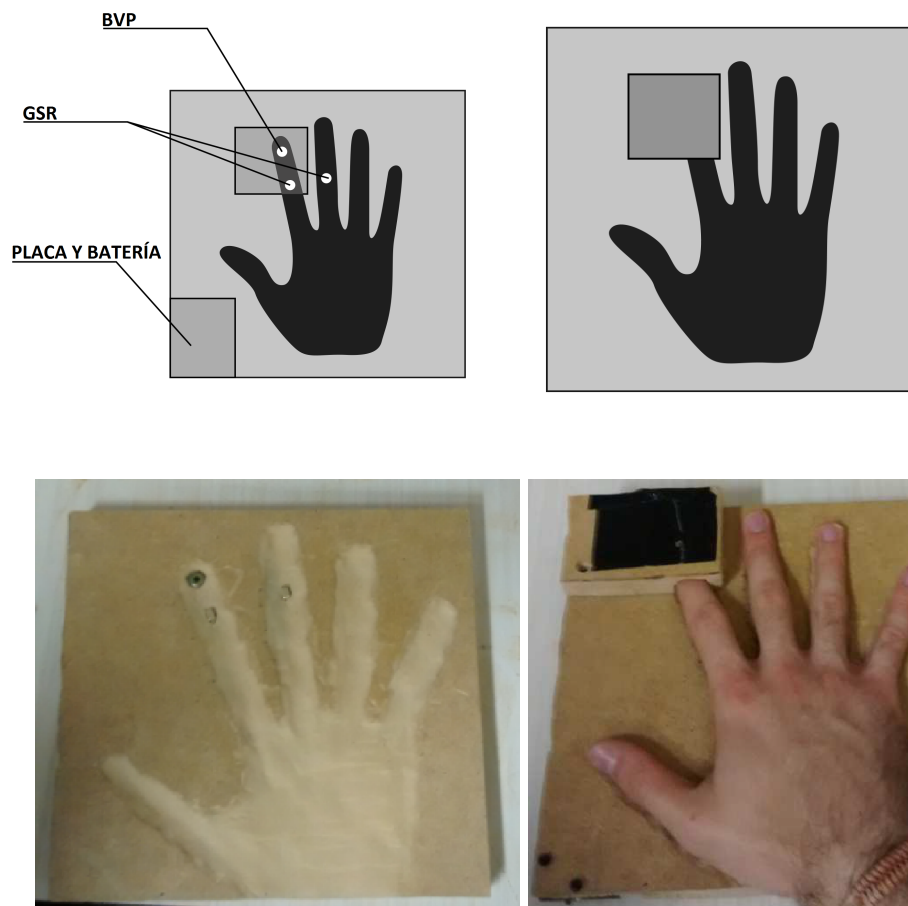


Figura 4.1: Prototipo para el detector de anomalías.

4.4. Software

Complementando al hardware descrito en la sección anterior, fue creado un programa en **Python** que cumpliera con lo descrito en la Tabla 4.1. El programa se encarga de la grabación y etiquetado de las señales para su posterior almacenamiento, lo que permite tomar las medidas a cualquier persona, incluso al propio sujeto cuyas señales se van a grabar si así se desea. En el Anexo B se detalla como ha sido desarrollado este programa, cuya interfaz puede ser observada en la Figura 4.2.

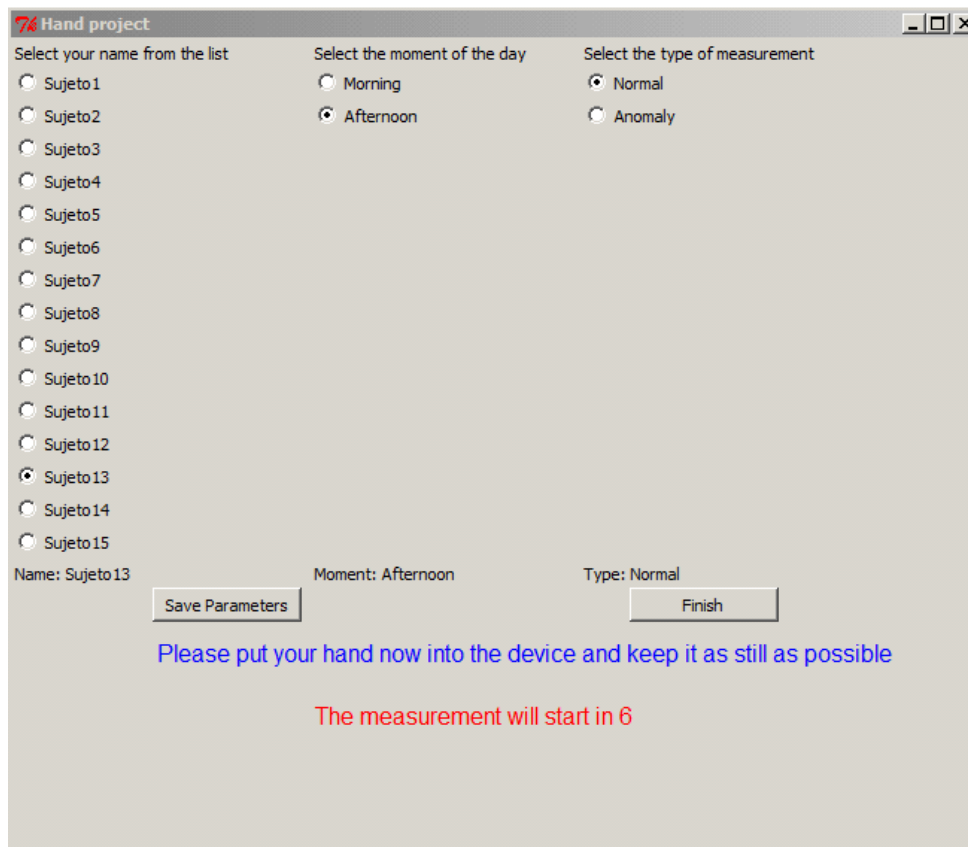


Figura 4.2: Instantánea de la interfaz del programa creado.

4.5. Protocolo de toma de datos

Durante el intervalo de tiempo en el que se llevó a cabo el proceso de toma de medidas se realizaron dos tipos de grabaciones: *normales*, en las que no se realizaba ninguna actividad especial y *anómalas*, dónde el sujeto llevaba a cabo alguna tarea que le generara estrés físico o cognitivo.

4.5.1. Tamaño de la muestra

- **Número de sujetos:** 13.
- **Número de muestras normales por sujeto:** 70.
- **Número de muestras anómalas por sujeto:** 6, 3 de estrés físico y 3 de estrés cognitivo.

4.5.2. Toma de muestras normales

Uno de los objetivos detallados en la sección 1.4 era la creación de una base datos con muestras normales de cada sujeto, durante las cuales no se estuviera realizando ningún tipo de actividad estresante. Para ello, se tomaban medidas de los sujetos dos veces al día, en el momento de comenzar y terminar la jornada laboral respectivamente. Este proceso se repitió durante los tres meses que duró aproximadamente completar la base de datos. Cada vez que se tomaba una medida se realizaba una grabación de duración 60 segundos, aunque ésto no quiere decir necesariamente que el tiempo utilizado para el análisis posterior sea ese, como se verá en el capítulo 5.

4.5.3. Toma de anomalías

Durante los meses en los que se estaban recopilando las muestras normales, ocasionalmente se realizaba la grabación de alguna de las anomalías que queríamos detectar. En concreto, han sido 6 el número de anomalías tomadas por sujeto, tres de estrés físico o fisiológico y tres de estrés cognitivo o psicológico. Las anomalías de tipo físico tienen un impacto claro en el ritmo cardíaco del sujeto, acelerándolo, con lo que *a priori* deberían ser relativamente sencillas de detectar. Por otro lado, las originadas por estrés cognitivo no parecen tan fáciles de detectar, dado que no se puede decir que afecten a un parámetro corporal de forma tan evidente como las físicas. Por lo tanto, resultará de interés realizar ambos tipos de anomalías y comparar la dificultad de detectar cada una de ellas. A continuación se explica más detalladamente como se han inducido estos datos anómalos:

- **Anomalías físicas:** el sujeto subía y bajaba 5 veces un tramo de escaleras. La medida, de 60 segundos de duración, se realizaba justo **después** de que se hubiera realizado esta tarea. Estudios como [10] hablan de la alteración de parámetros biométricos como la Variabilidad del Ritmo Cardíaco debido a pequeños esfuerzos físicos.
- **Anomalías cognitivas:** todos los sujetos han realizado tres pruebas diferentes que se describirán a continuación. En todos los casos las medidas de 1 minuto se han tomado **durante** la realización de la prueba pertinente.
 - *Test de Stroop* [11], utiliza el conocido como efecto **Stroop**, o interferencia cognitiva que se produce en un sujeto cuando lee el color del que está pintada una palabra, pero el significado de esa palabra indica un color distinto. Por ejemplo, leyendo el color de la siguiente palabra: **VERDE**. La prueba consiste en leer el color del mayor número posible de palabras en un minuto.
 - *Prueba de sustracción* [12], el sujeto debe restar en voz alta un número primo a un número grande de forma continua. Por ejemplo, 17 a 1431 (1414, 1397, 1380...) durante el minuto que dura la medida. Para incrementar la dificultad se cambian los números cuando el sujeto se ha acostumbrado.
 - *Prueba matemática* [13], consistente en la resolución del mayor número de pequeños problemas matemáticos en un minuto, relacionados con cálculo de porcentajes, divisiones, etc. Los problemas usados fueron los mismos para todos los sujetos.

5. Análisis

5.1. Introducción

A lo largo de este capítulo se explicará el análisis realizado a los datos obtenidos, y su tratamiento para la creación de los distintos clasificadores. Como primer paso se aplicará un preprocesado a los datos grabados (sección 5.2) antes de proceder en el apartado siguiente a la descripción de todas las características extraídas. Sin embargo, estas características pueden ser reducidas o combinadas para la creación de distintos clasificadores, por lo que en la sección 5.4 se detallan los métodos de reducción de características empleados.

Para acabar el capítulo, en la sección 5.5 se define el criterio de clasificación utilizado y se enumeran todos los clasificadores generados, que serán comparados en el capítulo 6.

5.2. Preprocesado

La realización de una serie de tratamientos previos sobre las señales recogidas por los dos sensores resulta indispensable para la obtención de unos resultados fiables. Estas operaciones serán descritas a continuación:

5.2.1. BVP

- **Reducción de la señal:** De los 60 segundos grabados solo será utilizado un intervalo de 30 para el análisis posterior. El criterio de selección de este intervalo es la ausencia en la medida de lo posible de artefactos como los descritos en la sección 3.2.2.
- **Filtrado:** Se aplica un filtro paso banda entre 0.5 y 2 Hz.
- **Detector de picos:** La señal es introducida a un algoritmo de detección de picos desarrollado por BitBrain, quedando tras el paso por ese algoritmo tal y como se observa en la Figura 5.1. El intervalo entre dos picos se denomina **IBI** (inter-beat-interval), y expresa el ritmo cardíaco instantáneo. Sirva como ejemplo, si entre dos picos el IBI es de 0.5 segundos el ritmo cardíaco en pulsaciones vendría dado por $60/\text{IBI}(\text{segundos}) = 120$ pulsaciones por minuto.

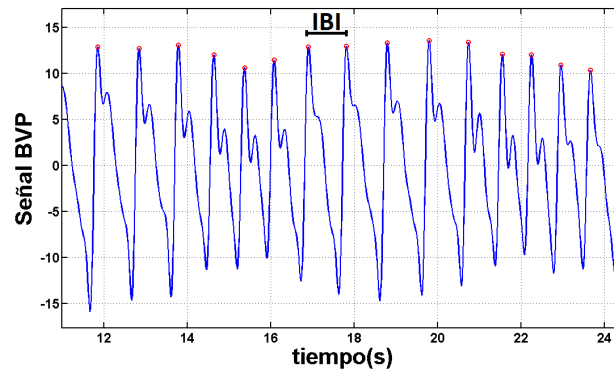


Figura 5.1: Señal de BVP tras el preprocesado. Los círculos rojos señalan los puntos identificados como picos por el algoritmo.

5.2.2. GSR

- **Reducción de la señal:** Con igual criterio de ausencia de artefactos, se selecciona un intervalo de 30 segundos de señal.
- **Filtrado:** Se aplica a la señal un filtro paso bajo con frecuencia de corte en 2 Hz.
- **Conversión:** Para la transformación de la señal de Voltios a Siemens, debe aplicarse una fórmula de conversión propia del hardware utilizado.

5.3. Extracción de características

En total han sido extraídas 16 características diferentes, 6 de la señal BVP y 10 de GSR, que se describen a continuación. Su selección ha sido el resultado de la búsqueda de las características usadas más comunes en varios artículos de referencia como [2], [14] o [15].

5.3.1. BVP

- **Ritmo cardíaco medio**, o media de todos los intervalos (IBI). Indica el ritmo cardíaco medio a lo largo de la grabación.
- **Mediana del ritmo cardíaco**, misma medida que la anterior, pero utilizando la mediana en vez de la media.
- **Desviación estándar del ritmo cardíaco**.
- **NN50**, parámetro que indica el número de intervalos IBI que han diferido en 50 o más ms respecto al intervalo anterior (indicador de cambios rápidos de pulso).
- **pNN50**, proporción de NN50 respecto al número total de intervalos IBI.
- **RMSSD** (*root mean square of successive differences*), raíz cuadrada de la media de los cuadrados de las sucesivas diferencias entre intervalos IBI adyacentes.

5.3.2. GSR

- **Media del GSR**, media de la señal total, parte tónica más fásica.
- **Desviación estándar del GSR**, de la señal total.
- **Media de la primera derivada del GSR**.
- **Media de la primera derivada del GSR normalizada por la desviación estándar del GSR**.
- **Media de la segunda derivada del GSR**.
- **Media de la segunda derivada del GSR normalizada por la desviación estándar del GSR**.
- **Media de la parte tónica**, solo de esa parte de la señal.
- **Desviación estándar de la parte tónica**.
- **Número de respuestas**, cantidad de estímulos (tanto específicos como no-específicos) en la parte fásica.
- **Área fásica**, o área bajo la señal fásica.

5.4. Reducción de características

En muchas aplicaciones, como la que estamos desarrollando, un alto número de posibles características pueden ser utilizadas. Sin embargo, ésto no implica que el mejor resultado vaya a ser obtenido necesariamente mediante la utilización de todas ellas. Algunas pueden ser redundantes entre sí, no aportando nada de información al clasificador o incluso empeorando su funcionamiento. También puede darse el caso de que una característica no sea de ninguna ayuda por sí sola y sin embargo, al combinarse con otras, resulte de alta utilidad (algunos ejemplos pueden verse en [16]). Por otro lado, el cálculo de más características siempre conllevará un mayor **coste computacional**. Solo con el objetivo de reducir ese tiempo de computación, merece la pena estudiar distintas posibilidades de reducción del número de dimensiones. Dentro de este campo podemos distinguir dos tipos de algoritmos:

- **Algoritmos de transformación:** las características existentes se transforman a un espacio dimensional más bajo, combinándose de forma que las nuevas variables contengan el máximo posible de la información original. El método utilizado en esta categoría será **PCA** [17], *Principal Component Analysis*, o *Análisis de Componentes Principales*.
- **Algoritmos de selección:** Parte de los datos disponibles se suministrarán a un algoritmo que, sabiendo las observaciones que pertenecen a cada clase, selecciona las características que más separan estas dos clases. Un 30 % de los datos recogidos (tanto normales como anómalos) se destinarán a estos algoritmos, y ya no podrán ser empleados para la clasificación. Se utilizarán dos métodos: **Búsqueda exhaustiva**, que evalúa todas las

combinaciones posibles (alto coste computacional) y **SFS** (*Sequential Forward Search*), con un coste de computación inferior.

Una explicación más detallada de estos algoritmos puede encontrarse en el Anexo C.

5.5. Clasificadores

La detección precisa de espurios es una de las tareas más importantes dentro del análisis de datos, debido a que su no consideración puede llevar a unos resultados, como mínimo, engañosos. Además, en numerosos ámbitos los propios espurios pueden resultar ser los puntos buscados ya que describen un comportamiento anómalo que muchas veces es de alto interés, como en el caso de la detección de intrusos o anomalías en el tráfico de una red (más ejemplos y una explicación más amplia pueden verse en [18]). Éste es también el caso que se da en el sistema que se está desarrollando en esta memoria.

La principal pregunta que surge es: ¿qué debe considerarse como espurio?. Aunque no existe una única definición generalmente aceptada, Hawkins [19] lo define como: *Un espurio es una observación que se desvía tanto de otras observaciones que levanta sospechas de que puede estar generada por otros mecanismos diferentes*. A partir de esta definición son varias las metodologías desarrolladas para la detección de espurios. En este trabajo se usará la llamada detección **basada en la distancia**, que computa un tipo de distancia para todos los puntos de la distribución y, dependiendo de lo grande que sea esta distancia, el punto es clasificado como normal o anómalo. Este método ha sido el elegido debido a que es adecuado para conjuntos de datos multidimensionales y es independiente de la distribución de los datos [20].

5.5.1. Distancias

El parámetro más importante a la hora de desarrollar el clasificador es la selección del tipo de distancia. La elección estándar en este tipo de métodos suele ser la de *Mahalanobis* [21]. La distancia de Mahalanobis de un punto con coordenadas en P dimensiones $x = (x_1, x_2, x_3, \dots, x_P)^T$ a un grupo de observaciones con media $\mu = (\mu_1, \mu_2, \mu_3, \dots, \mu_P)^T$ y matriz de covarianzas Σ se define como:

$$MD(x) = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)} \quad (5.1)$$

Sin embargo, la distancia de Mahalanobis cuenta con una desventaja muy grande si es aplicada tal y como se ha definido previamente. La media y la matriz de covarianzas de la distribución, necesarias para el cálculo de la distancia, son altamente sensibles a valores espurios, de forma que la distancia obtenida no resulta fiable. Por ello, han sido desarrollados varios algoritmos **robustos**, basados en esta distancia de Mahalanobis, que se definen como:

$$RD(x) = \sqrt{(x - \hat{\mu})^T \hat{\Sigma}^{-1} (x - \hat{\mu})} \quad (5.2)$$

donde $\hat{\mu}$ es la media y $\hat{\Sigma}$ la matriz de covarianzas, ambas calculadas de manera robusta por un algoritmo, siendo algunos de los más comunes discutidos en [22]. De todos ellos, el que será utilizaremos es el denominado **MCD**, *Minimum Covariance Determinant* (concretamente, el método y librería utilizados son aquellos descritos en [23]) que busca, de entre el conjunto de datos de tamaño total n , el subconjunto de tamaño h , con $h > n/2$, cuya matriz de covarianzas clásica tenga el menor determinante posible. De esta forma, $\hat{\mu} = \mu_{MCD}$ será la media de esos h puntos, y $\hat{\Sigma} = \Sigma_{MCD}$ será su matriz de covarianzas. El algoritmo MCD puede resistir $(n - h)$ datos espurios y, por lo tanto, ese valor h (o equivalentemente la proporción $\alpha = h/n$) determinará la robustez del algoritmo. Cuando se espera un alto nivel de contaminación por espurios α debería ser elegido con un valor cercano a 0.5. Sin embargo, para situaciones en las que no se espera una presencia excesiva de datos anómalos, un valor intermedio de $\alpha = 0.75$ es el recomendado, y el que ha sido elegido en esta memoria, ya que proporciona resultados precisos si el set de datos contiene hasta un máximo del 25 % de datos espurios, lo que es una condición bastante razonable para el sistema que se está desarrollando. Una explicación más profunda del algoritmo puede verse en [24].

5.5.2. Objetivo del clasificador

Los datos introducidos a nuestro sistema pueden ser *normales* o *anómalos*. El objetivo de los clasificadores es el de detectar si una muestra es anómala o no. Esta estimación dependerá de las características elegidas, así como del criterio utilizado para la clasificación de los datos. Al clasificar una muestra pueden darse cuatro posibles casos:

- **Verdadero Positivo (VP):** la nueva entrada se clasifica como dato anómalo y, efectivamente, lo es.
- **Verdadero Negativo (VN):** la nueva entrada es un dato normal y se clasifica correctamente como tal.
- **Falso Positivo (FP):** la nueva entrada es un dato normal, pero el clasificador dictamina que es anómalo.
- **Falso Negativo (FN):** la nueva entrada es un dato anómalo pero se clasifica como un dato normal.

Un clasificador perfecto conseguiría porcentajes del 100 % en VN y VP y del 0 % en FP y FN. En la Figura 5.2a se observa un clasificador formado por dos variables, que debe clasificar los datos que se muestran. Puede verse una nube de puntos azules (*normales*) que se agrupan formando un cluster, y unos pocos de color rojo y verde (*anómalos*). Las anomalías físicas (rojas) parecen bastante lejos del centro de esa distribución y consecuentemente fáciles de detectar. Sin embargo, las cognitivas (verdes) se sitúan mucho más próximas, tanto que algunas se encuentran dentro del cluster de muestras normales. Para que se clasifiquen correctamente todas las anomalías (es decir, FN = 0 %), el umbral debería ser colocado como se ve en la Figura 5.2b, clasificando todas las muestras que caen dentro del elipsoide como normales y todas las que hay fuera como anómalas, lo que implicaría tener un porcentaje de Falsos Positivos altísimo. Aumentar el umbral (Figura 5.2c) disminuiría el porcentaje de FP, pero también provocaría que

muestras anómalas se clasificaran como normales. Este es un ejemplo de un mal clasificador, con el que es imposible obtener valores bajos de ambos FN y FP.

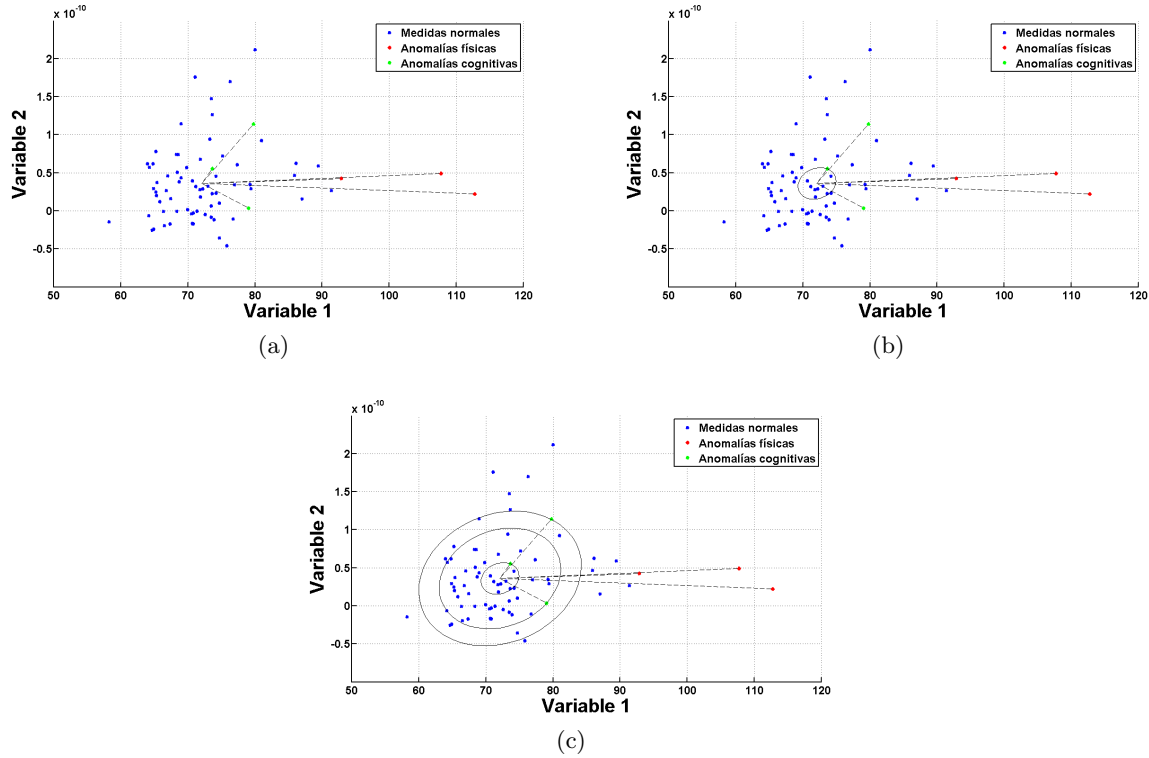


Figura 5.2: Ejemplo de mal clasificador.

El clasificador perfecto se parecería más al representado en la Figura 5.3a donde la distancia de las anomalías con el centro de la distribución es muy grande y es posible establecer una separación clara entre datos normales y anómalos. Por ejemplo, estableciendo el umbral como en la Figura 5.3b, quedan porcentajes de $FN = 0\%$ y $FP < 2\%$.

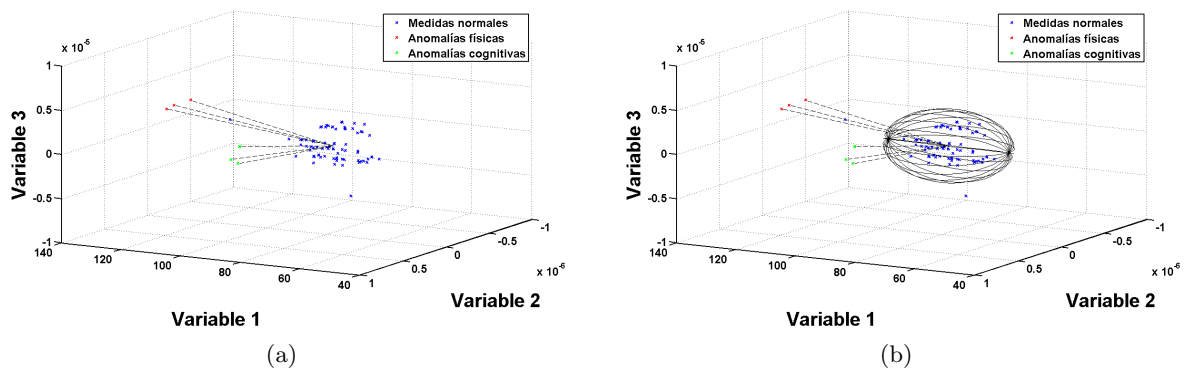


Figura 5.3: Ejemplo de buen clasificador.

Este será el objetivo de nuestro clasificador, establecer de la forma más precisa posible donde se sitúan los datos normales y dónde los anómalos. De esta forma, ante una nueva entrada, será capaz de clasificar correctamente esa muestra.

5.5.3. Selección de clasificadores

Durante las secciones anteriores se han descrito varios métodos que pueden ser utilizados para la creación de los clasificadores. Con distintas combinaciones de ellos se han generado 12 métodos diferentes agrupados en 4 tipos según el algoritmo usado que serán descritos a continuación, así como resumidos en la Tabla 5.1:

- **Clasificador 1:** Sin utilizar ningún algoritmo de reducción, se calcula la distancia robusta utilizando las 16 características totales.
- **Clasificadores 2-5:** se utiliza el algoritmo *Búsqueda Exhaustiva* de selección de características, escogiéndose las 4, 6, 8 y 10 mejores para cada sujeto.
- **Clasificadores 6-9:** Utilización del algoritmo *SFS* de selección de características, escogiéndose las 4, 6, 8 y 10 mejores para cada sujeto.
- **Clasificadores 10-12:** Se emplea el algoritmo *PCA* de transformación de características, escogiendo las 2, 4, y 6 componentes principales por sujeto.

	Reducción de características	Algoritmo	Nº de componentes
Clasificador 1	No	Ninguno	16
Clasificador 2	Sí	Búsqueda exhaustiva	4
Clasificador 3	Sí	Búsqueda exhaustiva	6
Clasificador 4	Sí	Búsqueda exhaustiva	8
Clasificador 5	Sí	Búsqueda exhaustiva	10
Clasificador 6	Sí	SFS	4
Clasificador 7	Sí	SFS	6
Clasificador 8	Sí	SFS	8
Clasificador 9	Sí	SFS	10
Clasificador 10	Sí	PCA	2
Clasificador 11	Sí	PCA	4
Clasificador 12	Sí	PCA	6

Tabla 5.1: Clasificadores desarrollados.

6. Resultados

6.1. Introducción

En este apartado se compararán los resultados obtenidos para todos los clasificadores generados en el capítulo anterior. Comenzaremos (sección 6.2) explicando el método de evaluación utilizado para después proceder a la comparación entre clasificadores en la sección 6.3, donde también se determinará cual es el mejor clasificador dependiendo de las circunstancias. Posteriormente se realizará un análisis más profundo de los resultados obtenidos con uno de ellos, buscando los valores óptimos de tiempo de análisis y de muestras normales necesarias (sección 6.4), y comparando los resultados según el tipo de anomalía (sección 6.5). Además, se presentarán los resultados para el caso de un nuevo sujeto del cual no tengamos ninguna muestra previa (sección 6.6) y para finalizar se resumirán los resultados extraídos en la sección 6.7.

6.2. Método de evaluación

6.2.1. Curva ROC

Como criterio de evaluación se ha elegido la **Curva ROC** (acrónimo de *Receiver Operating Characteristic*, *Característica Operativa del Receptor*), una representación gráfica de dos parámetros que definiremos a continuación: sensibilidad y especificidad.

$$Sensibilidad = \frac{VP}{VP + FN} = VPR \quad (6.1)$$

$$Especificidad = \frac{VN}{VN + FP} = 1 - FPR \quad (6.2)$$

La sensibilidad indica la capacidad de nuestro estimador para clasificar como positivos los casos que lo son realmente, o lo que es lo mismo, el Ratio de Verdaderos Positivos (TPR). La especificidad indica la habilidad del clasificador para detectar la normalidad en medidas normales, indicando la proporción de medidas normales correctamente identificadas. Es equivalente a $1 - FPR$ (Ratio de Falsos Positivos).

La curva ROC representa la *sensibilidad* frente a $(1 - \textit{especificidad})$ o, lo que es lo mismo, el ratio de Falsos Positivos frente al de Verdaderos Positivos. Para cada umbral se dibujará un punto en la gráfica, quedando una curva como resultado.

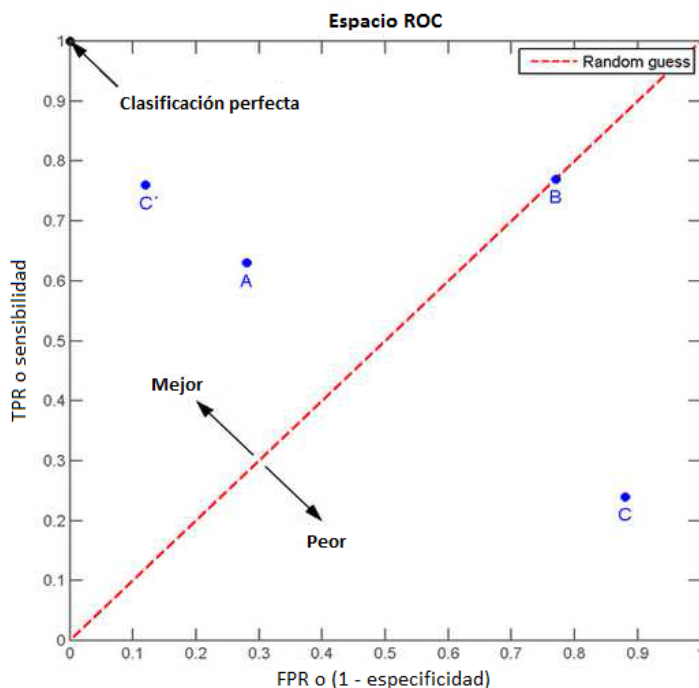


Figura 6.1: Curva ROC: la línea discontinua marca la curva obtenida por un clasificador aleatorio. Por encima de ella cualquier valor será mejor que la estimación aleatoria, situándose la clasificación perfecta en (0, 1).

El mejor método posible de predicción se situaría en la esquina superior izquierda, o coordenada (0,1) del espacio ROC, representando un 100 % de sensibilidad (ningún falso negativo) y un 100 % también de especificidad (ningún falso positivo). Dado que la esquina superior izquierda es el punto óptimo, un evaluador directo del clasificador es el **Área Bajo la Curva (AUC)**. El clasificador ideal formaría un ángulo recto con el punto (0,1) y tendría una AUC de 1. De esta forma, cuanto más próximo a 1 sea el valor de la AUC, mejor será el clasificador.

La elección del umbral del clasificador será equivalente a la selección del punto óptimo de la curva. El emplazamiento de este punto vendrá dado por el coste asignado a los dos tipos de clasificación errónea (FN y FP), lo que dependerá del contexto de aplicación.

6.2.2. Cross-validation

Para que los resultados sean fiables, no pueden ser utilizadas las mismas muestras para la generación del clasificador y su evaluación, ya que correríamos el riesgo de cometer lo que se denomina *overfitting* o sobreajuste, que es el efecto de sobreentrenar un algoritmo de aprendizaje con unos ciertos datos para los que se conoce el resultado deseado.

Para solventar este problema se utiliza una técnica denominada **cross-validation** o **validación cruzada**, utilizada para evaluar los resultados de un análisis estadístico y garantizar que

no se produce ese sobreajuste. El método que se ha utilizado se denomina *validación cruzada de K-iteraciones* o *K-fold cross validation*, donde los datos de muestra se dividen en K subconjuntos o *folds* (elegidos de forma aleatoria en nuestro sistema) de igual o similar tamaño. Uno de los subconjuntos se utiliza como datos de prueba y el resto (K-1) como datos de entrenamiento. Finalmente se realiza la media aritmética de los resultados de cada iteración para obtener un único resultado. Este método es muy preciso puesto que evaluamos a partir de K combinaciones de datos de entrenamiento y de prueba, aunque tiene la desventaja de que es lento desde el punto de vista computacional. En la práctica, la elección del número de iteraciones depende de la medida del conjunto de datos, para este sistema se ha determinado un valor de $K = 5$.

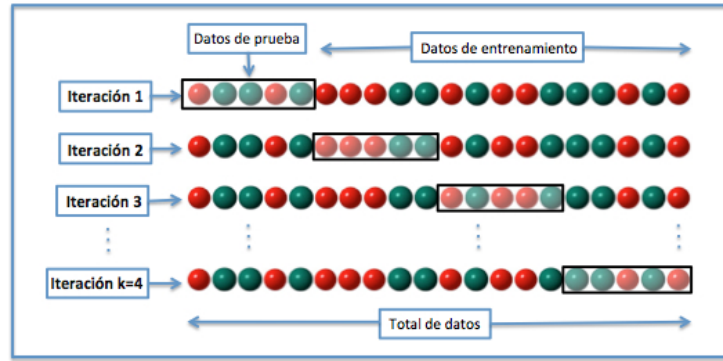


Figura 6.2: Validación cruzada de K-iteraciones, con $K = 4$. Tras las 4 repeticiones se hace una media de los resultados obtenidos con los datos de prueba. En el ejemplo de esta figura los *folds* se han elegido de forma consecutiva, mientras que en el nuestro se han elegido de forma aleatoria.

6.3. Evaluación y comparación de clasificadores

En la Figura 6.3 se representa la curva ROC media de todos los sujetos obtenida con cada clasificador, junto con los valores de AUC para cada curva.

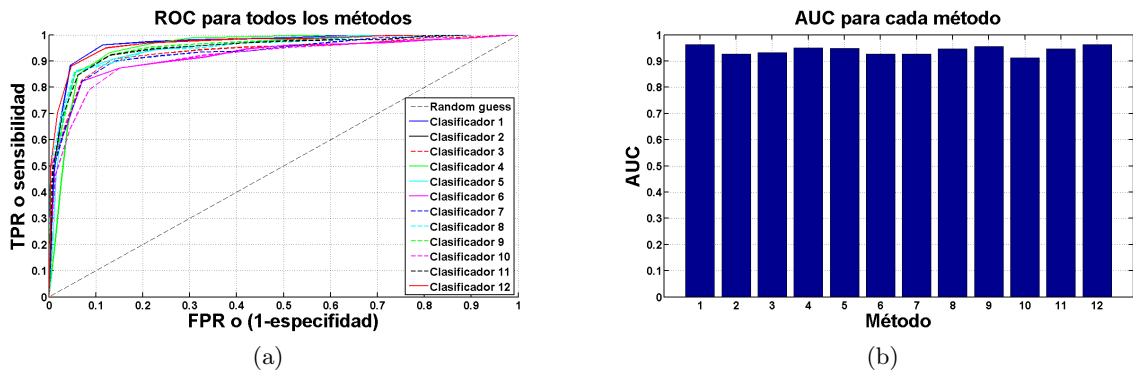


Figura 6.3: Curvas ROC para todos los clasificadores. Cada curva representa la media de todos los sujetos con cada clasificador.

En la sección 5.5.3 se describieron los 12 clasificadores, divididos en 4 grupos según el algoritmo usado. Cabe destacar que para cada grupo se estudiaron también clasificadores con un número mayor de características, pero no fueron incluidos ya que no mejoraban los resultados o incluso los empeoraban respecto a los mejores métodos de cada grupo. Observando las curvas de la Figura 6.3 pueden observarse los clasificadores de cada grupo que obtienen mejores resultados, siendo éstos el 1, 4, 9 y 12. Sus curvas ROC se vuelven a exponer en la Figura 6.4, esta vez representando tanto la media de todos los sujetos como la desviación típica obtenida.

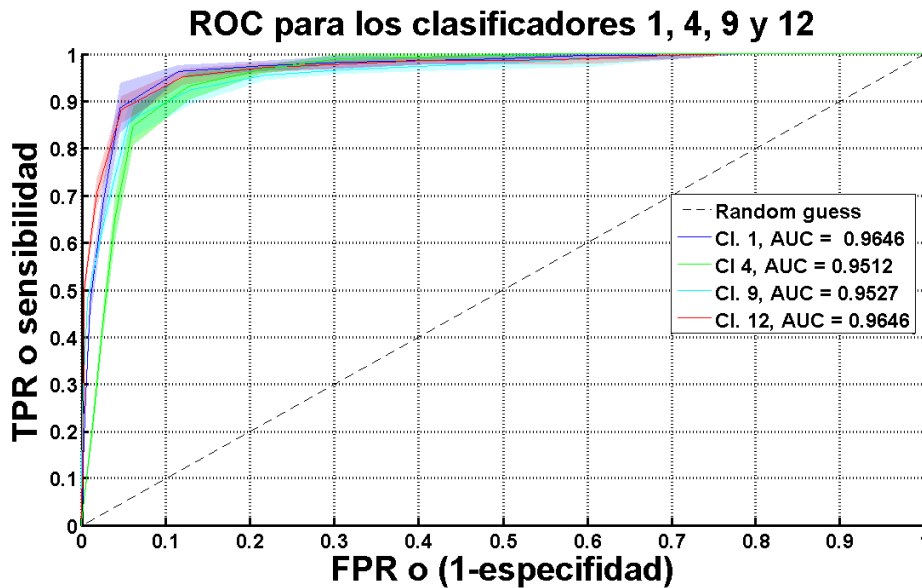


Figura 6.4: Curvas ROC para los mejores clasificadores. La línea continua representa la media de todos los sujetos y la zona sombreada la desviación típica, para cada clasificador.

Cabe destacar que la selección del mejor clasificador no es inmediata. Existe un compromiso entre el porcentaje de Falsos Positivos permitidos y el de Verdaderos Positivos obtenidos. Por lo tanto, la elección de un clasificador u otro dependerá del coste que asignemos a estimar erróneamente un dato normal o una anomalía. Se pueden considerar tres situaciones:

1. **Coste de un FP \gg Coste de un FN.** El objetivo es tener el menor número de FP posible sin considerar tan importantes los FN, es decir, se le da mucha importancia a que ningún dato normal sea detectado anómalo. En este caso, el mejor clasificador sería el número 12, y el punto óptimo de la curva podría ser el de coordenadas (FPR = 0.02, TPR = 0.5) (Figura 6.5a).
2. **Coste de un FP = Coste de un FN.** Ambos errores de clasificación tienen el mismo coste. El clasificador elegido sería el 1 en este caso, con un punto óptimo situado en (FPR = 0.07, TPR = 0.93), donde el porcentaje de FN y FP sería el mismo (Figura 6.5b).
3. **Coste de un FP \ll Coste de un FN.** En este caso el sistema se ve penalizado en mayor medida por la incorrecta clasificación de un dato anómalo. El mejor clasificador sería el 4, con punto óptimo en (FPR = 0.6, TPR = 1), donde se clasifican correctamente todas las anomalías (Figura 6.5c).

Dado que el sistema de este proyecto tiene como objeto detectar las anomalías, se considera la situación 3 como la que más se adapta a estos objetivos, y por lo tanto, se seleccionará el clasificador 4 para realizar un análisis más detallado en los siguientes apartados.

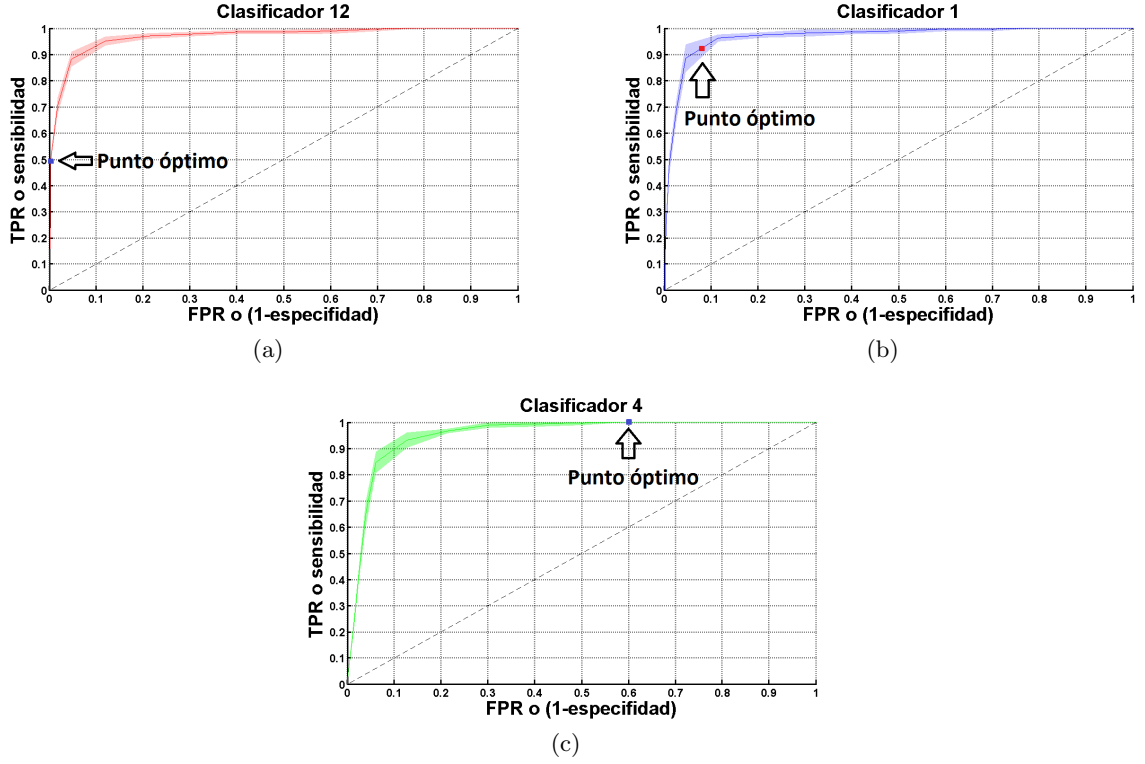


Figura 6.5: Clasificadores y puntos óptimos según el coste de FN y FP. 6.5a: Coste de un FP \gg Coste de un FN, 6.5b: Coste de un FP = Coste de un FN y 6.5c: Coste de un FP \ll Coste de un FN.

6.4. Mejoras

Durante el desarrollo de los clasificadores hubo dos parámetros que se establecieron, pero no era posible en ese momento saber si se trataban de los valores óptimos o no. Se trata del tiempo de grabación analizado, que se tomó de 30 segundos y del número de muestras normales usadas para la creación del clasificador. Después de la toma de datos se tenían a disposición 70 muestras normales. Sin embargo, tras su paso por los algoritmos de reducción de características y la validación cruzada, solo 40 pueden ser utilizadas como máximo para la generación de los clasificadores. Resulta interesante estudiar ahora si estos valores son óptimos o no.

6.4.1. Tiempo de grabación

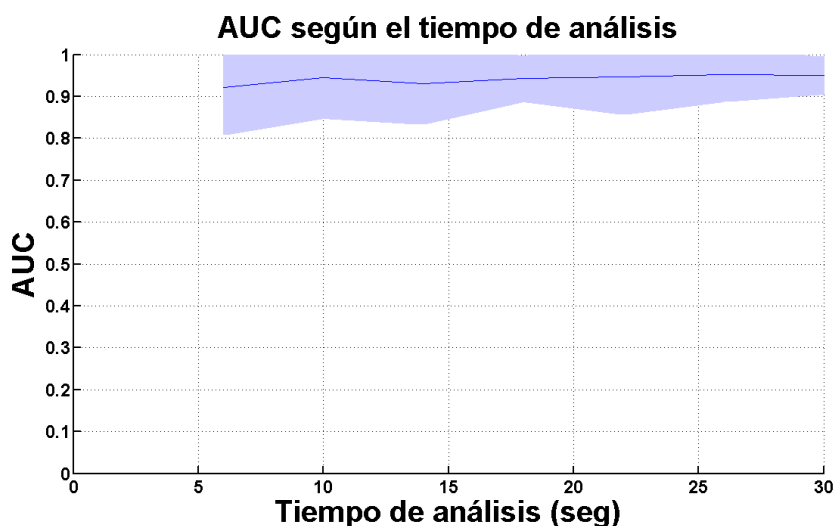


Figura 6.6: Valor del área bajo la curva dependiendo del tiempo de análisis utilizado. La línea continua representa la media de todos los sujetos y la zona sombreada la desviación típica.

En la Figura 6.6 se observa el valor de la AUC obtenido según el tiempo de análisis tomado, en un intervalo que va de 6 a 30 segundos. Se puede observar como el funcionamiento del sistema va mejorando según aumenta el tiempo de análisis, y se estabiliza a partir de 25 segundos más o menos. Cabe destacar que con tiempos pequeños de análisis el valor de la AUC empeora pero sigue siendo bastante alto, lo que significa que el tiempo de grabación podría ser reducido en aplicaciones futuras si fuera necesario, sin acusarse excesivamente en los resultados.

6.4.2. Número de muestras normales

En la Figura 6.7 se representa la AUC obtenida según el número de muestras normales utilizadas para el análisis, estableciéndose un mínimo de 10 y un máximo de 40. Se observa un empeoramiento acusado de los resultados para un número de muestras inferior a 20, y a partir de 30 apenas se observan mejoras. Esto nos lleva a concluir que el número de muestras normales a utilizar debería ser como mínimo de 30 para obtener los mejores resultados posibles. Cabe

destacar que sin embargo no existe una cota superior para el número de muestras normales, dado que cuantas más tengamos mejor funcionará el clasificador.

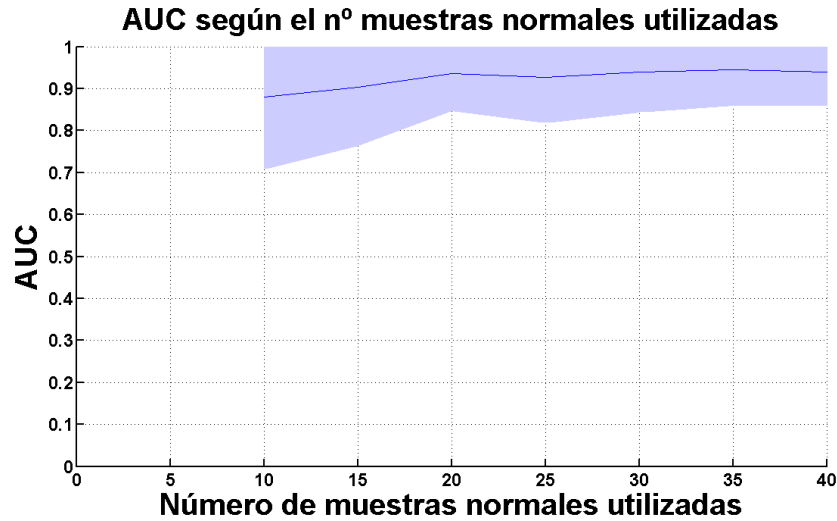


Figura 6.7: Valor del área bajo la curva dependiendo del número de muestras normales utilizado.

6.5. Comparación por tipo de anomalía

Dado que las anomalías que se han intentado detectar han sido de dos tipos, físicas y cognitivas, resulta de interés estudiar si ambas se han detectado con eficacia similar o, por el contrario, un tipo resulta más difícil de detectar que el otro.

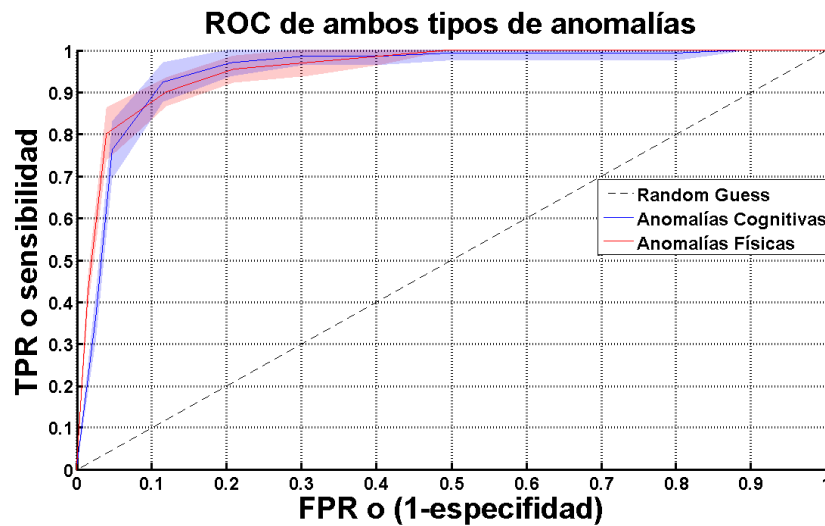


Figura 6.8: Media y desviación estándar de los sujetos para los dos tipos de anomalías.

En la Figura 6.8 se representa la media y desviación típica de los dos tipos de anomalías. No se observan grandes diferencias entre la clasificación de los dos tipos, ni se puede asegurar que uno sea mucho más fácil de clasificar que el otro. Cabe destacar que las anomalías físicas alcanzan el valor de $\text{TPR} = 1$ para un valor bastante menor de FPR , en concreto para $\text{FPR} = 0.5$ todas las anomalías físicas han sido ya detectadas mientras que esa situación no se alcanza para las cognitivas hasta $\text{FPR} = 0.85$ aproximadamente.

6.6. Resultados para un nuevo sujeto

Los clasificadores desarrollados hasta ahora se han basado en los datos previos del sujeto para clasificar una nueva entrada como normal o anómala. Sin embargo, ahora se busca estudiar que resultados se obtendrían en el caso de que no hubiera ninguna muestra previa del sujeto cuya medida se intenta clasificar. Para ello se intentan clasificar los datos de cada sujeto a partir de los demás sujetos, excluyendo los suyos propios. En la Figura 6.9 se representa la media y desviación típica obtenida.

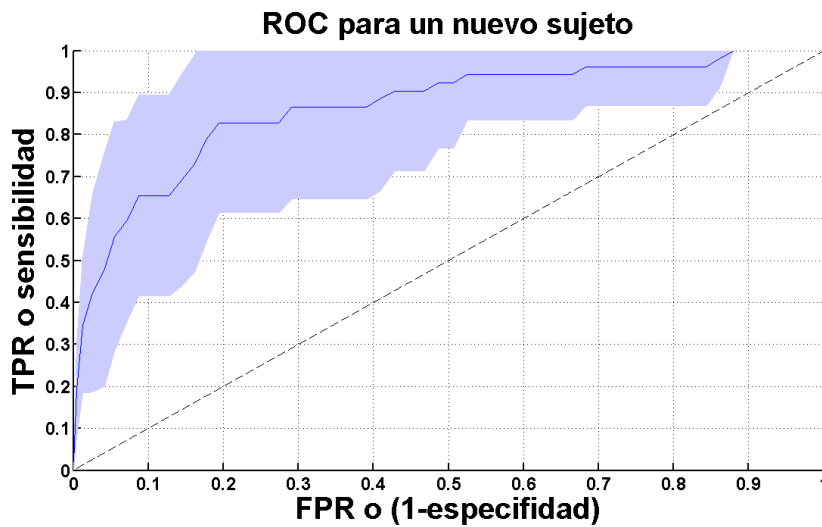


Figura 6.9: ROC para un nuevo sujeto.

Se observa un gran aumento de la variabilidad, lo cual es lógico teniendo en cuenta que no tenemos datos previos de ese sujeto. Estos resultados remarcan la necesidad de tener muestras previas del sujeto cuya medida se pretende clasificar para poder realizar una clasificación en la que se pueda confiar.

6.7. Resumen de resultados

A continuación se recapitulan los resultados más importantes obtenidos durante este capítulo:

- De los doce clasificadores desarrollados, cuatro obtienen mejores resultados que los demás al observar las curvas ROC de cada uno.
- La elección del mejor clasificador entre esos cuatro dependerá del coste que asignemos a tener un Falso Positivo y un Falso Negativo.
- El tiempo de grabación podría ser reducido a 20 segundos sin pérdida de rendimiento. Si fuera necesario, podría bajar hasta 6 segundos con un ligero detrimento de los resultados, pero con un funcionamiento todavía bueno.
- El número de muestras normales utilizadas debería ser como mínimo de 25-30 para obtener los mejores resultados posibles. En este aspecto, cuantas más muestras normales se tengan más precisa será la clasificación.
- No existen grandes diferencias entre la clasificación de los dos tipos de anomalías.
- La utilización del sistema en un nuevo sujeto del que no se tengan muestras normales previas no es fiable, ya que la variabilidad aumenta drásticamente.

7. Conclusiones y líneas futuras

En el presente proyecto se ha desarrollado un sistema de detección de datos espurios, siendo estos datos anomalías fisiológicas y psicológicas inducidas en sujetos de los que se había creado una base de datos de medidas sin estrés inducido. El estudio de patentes presentado en la sección 2.2 y ampliado en el Anexo D deja constancia de la escasez de dispositivos de este tipo en la actualidad, lo que dificulta su comparación con sistemas de similares características.

Dentro de la realización de este proyecto, y considerando los objetivos descritos en el capítulo 1, cabe destacar dos hitos principales conseguidos a partir del trabajo realizado: **(a)** la comprobación de que los sensores desarrollados por BitBrain pueden utilizarse con las mismas garantías que unos de tipo médico, lo cual resulta de gran utilidad para la empresa en su trabajo presente y futuro con estos sensores y **(b)** la creación de un sistema completo de detección de espurios. Los resultados obtenidos en la detección de anomalías fisiológicas y psicológicas son esperanzadores en cuanto a su utilización en ámbitos de aplicación más específicos. Además, el desarrollo de varios clasificadores diferentes dentro de este sistema permite que pueda ser adaptado según los requisitos y objetivos de la aplicación, ampliando así sus posibilidades futuras.

Todo lo expuesto anteriormente deja abiertos varios campos de aplicación futuros, así como alguna mejora posible a realizar si el producto se llevará al ámbito comercial, como son:

- Realización de un artículo científico similar a [6] presentando los resultados obtenidos en la comparación de sensores de BitBrain y TMSI. La propia empresa ya ha mostrado interés en el desarrollo de este artículo.
- Utilización del sistema de detección de espurios desarrollado en ámbitos de aplicación más concretos, como podrían ser: detección de picos de estrés en los trabajadores de una fábrica, identificación de infractores en el transporte público, etc. Cabe destacar que una empresa externa ya se ha puesto en contacto con BitBrain para estudiar la implantación de este sistema en alguna de sus fábricas, lo que empezará a realizarse en los próximos meses.
- Aumento de la base de datos de sujetos, así como del número de muestras tomadas por sujeto.
- Creación de un filtro para la detección automática de artefactos para las señales de GSR y BVP.
- Desarrollo de un nuevo prototipo con los mismos requisitos que el creado en este proyecto pero con un diseño más adecuado para su aplicación en el ámbito comercial.

7. Conclusiones y líneas futuras

A nivel personal, cabe destacar mi satisfacción por la posibilidad de desarrollar este PFC dentro de una empresa, lo que me ha permitido tener una primera toma de contacto con el mundo y ritmo laboral, a la vez que obtener una ganancia bastante importante en cuanto a conocimientos prácticos y experiencia profesional se refiere. Así mismo, considero altamente gratificante el hecho de que el trabajo expuesto en esta memoria pueda ser continuado en el futuro por parte de la empresa en la que he realizado el proyecto.

Bibliografía

- [1] Rosalind W. Picard. Affective computing. *The MIT Press*, 1997.
- [2] Rosalind W. Picard, E Vyzas, and J Healey. Toward machine emotional intelligence: analysis of affective physiological state. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23:1175–1191, 2001.
- [3] Mathieu Lajante, Olivier Droulers, Thibaut Dondaine, and David Amarantini. Opening the black box of electrodermal activity in consumer neuroscience research. *Journal of Neuroscience, Psychology and Economics*, 5:238–249, 2012.
- [4] W. Boucsein. *Electrodermal activity*. Plenum Press, 1992.
- [5] G. G. Berntson, J. T. Jr. Bigger, D. L. Eckberg, P. Grossman, P. G. Kaufmann, M. Malik, H. N. Nagaraja, S. W. Porges, J. P. Saul, P. H. Stone, and M. W. van Der Molen. Heart rate variability: origin, methods, and interpretive caveats. *Psychophysiology*, 34:623–648, 1997.
- [6] Poh Ming-Zher, Nicholas C. Swenson, and Rosalind W. Picard. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering*, 57:1243–1252, 2010.
- [7] E.T. McAdams, J Jossinet, Lacknermeier A., and F. Risacher. Factors affecting electrode-gel-skin interface impedance in electrical impedance tomography. *Medical, Biological Engineering and Computing*, 34:397–408, 1996.
- [8] D. C. Fowles, M. J. Christie, R. Edelberg, W. W. Grings, D. T. Lykken, and P.H. Venables. Committee report. publication recommendations for electrodermal measurements. *Psychophysiology*, 18:232–239, 1981.
- [9] S. Fisher. Body image and asymmetry of body reactivity. *J. Abnorm. Psychol.*, 57:292–298, 1958.
- [10] Caroline Di Bernardi Luft, Emilio Takase, and David Darby. Heart rate variability and cognitive function: Effects of physical effort. *Biological Psychology*, 82:186–191, 2009.
- [11] Patrice Renaud and Jean-Pierre Blondin. The stress of stroop performance: physiological and emotional responses to color-word interference, task pacing, and pacing speed. *International Journal of Psychophysiology*, 27:87–97, 1997.

- [12] Sue E. Kase and Frank E. Ritter. Serial subtraction errors revealed. *Cognitive Science Society*, pages 1551–1556, 2009.
- [13] Mark H. Ashcraft and Elizabeth P. Kirk. The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology*, 130:224–237, 2001.
- [14] K. H. Kim, S. W. Bang, and Kim S. R. Emotion recognition system using short-term monitoring of physiological signals. *Medical, Biological Engineering and Computing*, 42:419–427, 2004.
- [15] Jing Zhai and Armando Barreto. Stress detection in computer users based on digital signal processing of noninvasive physiological variables. *IEEE*, pages 1355–1358, 2006.
- [16] Isabelle Guyon and Andre. Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003.
- [17] Jonathon Shlens. A tutorial on principal component analysis. *Google Research*, 2014.
- [18] Victoria J. Hodge and Jim Austin. A survey of outlier detection methodologies. *Artificial Intelligence Review*, pages 85–126, 2004.
- [19] D. Hawkins. *Identification of Outliers*. Chapman and Hall, 1980.
- [20] Edwin M. Knorr, Raymond T. Ng, and Vladimir. Tucakov. Distance-based outliers: algorithms and applications. *The VLDB Journal*, 8:237–253, 2000.
- [21] Maria Teresa Escobedo Portillo and Jorge A. Salas Plata Mendoza. Mahalanobis y las aplicaciones de su distancia estadística. *CULCyT*, 5:13–20, 2008.
- [22] P. Filzmoser, R. Maronna, and M. Werner. Outlier identification in high dimensions. *Computational Statistics and Data Analysis*, 52:1694–1711, 2008.
- [23] Sabine. Verboven and Mia. Hubert. Libra: a matlab library for robust analysis. *Chemometrics and Intelligent Laboratory Systems*, 75:127–136, 2005.
- [24] Peter J. Rousseeuw and Katrien Van Driessen. A fast algorithm for the minimum covariance determinant estimator. *Technometrics*, 41:212–223, 1999.
- [25] S. Theodoridis and K. Koutroumbas. *Pattern Recognition*. Academic Press, 2009.
- [26] S. Theodoridis and K. Koutroumbas. *An introduction to pattern recognition: a MATLAB approach*. Academic Press, 2010.