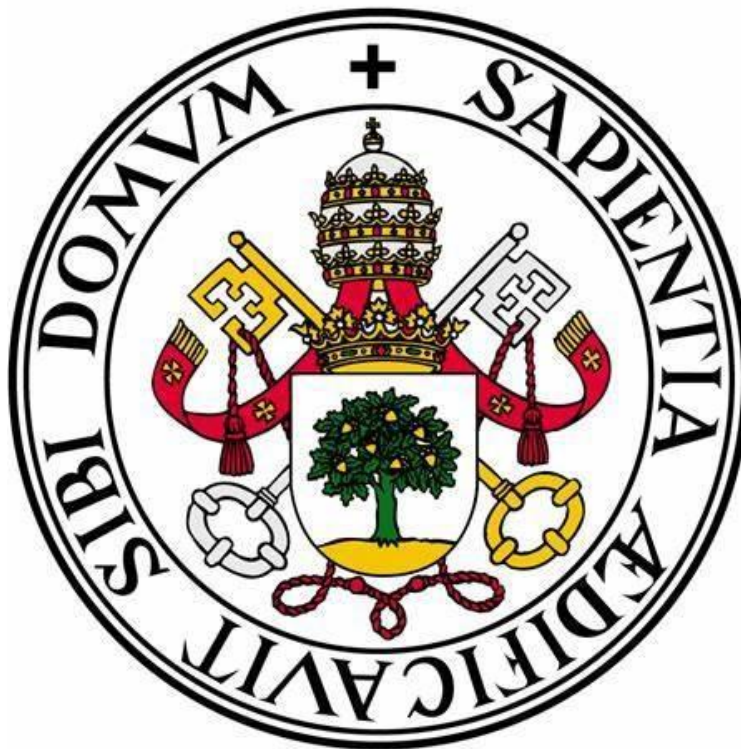


UNIVERSIDAD DE VALLADOLID

FACULTAD DE MEDICINA

ESCUELA DE INGENIERÍAS INDUSTRIALES



TRABAJO DE FIN DE GRADO DE INGENIERÍA BIOMÉDICA

2025

Estimación del Estado Emocional utilizando Señales Fisiológicas Periféricas

Autora: Alba Iglesias Pascual

Tutor: Javier Pérez Turiel

Tutora: Ana Cisnal de la Rica

Este trabajo se lo dedico a mi familia, que siempre me ha apoyado en mis decisiones y en mis metas, dándome el espacio y la comprensión que necesito en cada momento. También se lo dedico a los amigos que he hecho durante la carrera, porque sin ellos hubiese sido todo mucho más duro y aburrido. Doy gracias por todos los momentos de risas y anécdotas que nunca voy a olvidar, y que han hecho de esta etapa de mi vida algo especial.

RESUMEN

Se han hecho varios estudios acerca de la predicción del estado emocional de las personas a partir de señales fisiológicas. Una forma común de abordar la clasificación es mediante el modelo circunflejo de Russell, donde las emociones se representan en un espacio bidimensional con dos ejes: valencia y excitación. En este trabajo se ha diseñado un modelo predictivo para la clasificación emocional utilizando las señales fisiológicas provenientes del electrocardiograma (ECG), la actividad electrodérmica (EDA), y la temperatura corporal periférica (SKT). Se han planteado tres problemas de clasificación: clasificación binaria de la valencia, clasificación binaria de la excitación, y clasificación multiclase de los cuatro cuadrantes emocionales definidos por modelo circunflejo de Russell. Para cada uno de estos problemas se ha evaluado el desempeño de los modelos de clasificación Random Forest (RF), y Red Neuronal Artificial (RNA). Los mejores resultados obtenidos alcanzaron un 62% de exactitud con el modelo RF en la clasificación binaria de la valencia y la excitación. Estos resultados son comparables con los obtenidos en investigaciones previas y representan un avance en el uso de señales periféricas para la detección emocional.

ABSTRACT

Several studies have been conducted on the prediction of people's emotional states from physiological signals. A common approach to classification is through Russell's circumplex model, where emotions are represented in a two-dimensional space with two axes: valence and arousal. In this work, a predictive model for emotional classification has been designed using physiological signals from the electrocardiogram (ECG), electrodermal activity (EDA), and peripheral body temperature (SBT). Three classification problems have been posed: binary classification of valence, binary classification of arousal, and multiclass classification of the four emotional quadrants defined by Russell's circumplex model. For each of these problems, the performance of the Random Forest (RF) and Artificial Neural Network (ANN) classification models has been evaluated. The best results obtained reached 62% accuracy with the RF model in the binary classification of valence and arousal. These results are comparable with those obtained in previous research and represent an advance in the use of peripheral signals for emotional detection.

ÍNDICE

RESUMEN.....	3
ABSTRACT	3
ÍNDICE	4
ÍNDICE FIGURAS	7
ÍNDICE TABLAS	8
ABREVIATURAS Y ACRÓNIMOS.....	9
1. INTRODUCCIÓN.....	10
1.1. MARCO TEÓRICO	10
1.1.1. Las emociones.....	10
1.1.2. Señales fisiológicas	12
1.1.2.1. Señal ECG.....	12
1.1.2.2. Señal EDA.....	15
1.1.2.3. Señal SKT	16
1.2. JUSTIFICACIÓN Y PLANTEAMIENTO DEL PROBLEMA.....	17
1.3. HIPÓTESIS	18
1.4. OBJETIVOS.....	18
2. MATERIALES Y MÉTODOS.....	20
2.1. BASE DE DATOS	21
2.1.1. DEAP <i>Dataset</i>	21
2.1.2. MAHNOB-HCI.....	22
2.2. SEÑALES FISIOLÓGICAS	25
2.2.1. Señal ECG.....	25
2.2.2. Señal EDA	27
2.2.3. Señal SKT	28
2.3. EXTRACCIÓN DE CARACTERÍSTICAS.....	28
2.3.1. Características de la señal de ECG	28

2.3.2. Características de la señal de EDA	32
2.3.3. Características de la señal de SKT	34
2.4. SELECCIÓN DE CARACTERÍSTICAS	36
2.5. MODELOS DE CLASIFICACIÓN	38
2.5.1. Fundamentos generales	39
2.5.1.1. Redes Neuronales Artificiales	39
2.5.1.2. Random Forest	40
2.5.2. Modelo de clasificación binaria de la valencia y excitación.....	41
2.5.2.1. Random Forest	41
2.5.2.2. Redes Neuronales Artificiales	41
2.5.3. Modelo de clasificación multiclase de los cuatro cuadrantes emocionales	41
2.5.4. Random Forest	43
2.5.5. Redes Neuronales Artificiales.....	43
2.6. ENTRENAMIENTO Y VALIDACIÓN	43
2.7. MEDIDAS DE EVALUACIÓN	44
2.8. ADQUISICIÓN DE SEÑALES FISIOLÓGICAS	45
2.8.1. Señal ECG.....	46
2.8.2. Señal EDA	48
2.9. VALIDACIÓN CON SEÑALES FISIOLÓGICAS PROPIAS	49
3. RESULTADOS.....	50
3.1. Clasificación binaria de la valencia	50
3.1.1. Random Forest	50
3.1.2. Red Neuronal Artificial.....	51
3.2. Clasificación binaria de la excitación	51
3.2.1. Random Forest	51
3.2.2. Red Neuronal Artificial.....	52
3.3. Clasificación multiclase de los 4 cuadrantes emocionales	53

3.3.1. Random Forest	53
3.3.2. Red Neuronal Artificial.....	54
4. DISCUSIÓN.....	56
4.1. Análisis de los resultados	56
4.1.1. Clasificación binaria de la valencia	56
4.1.2. Clasificación binaria de la excitación	56
4.1.3. Clasificación multiclase de los 4 cuadrantes emocionales	57
4.2. Comparación con trabajos previos	57
4.3. Limitaciones	58
5. CONCLUSIÓN Y LÍNEAS FUTURAS	59
6. REFERENCIAS	60

ÍNDICE FIGURAS

1. INTRODUCCIÓN..... 10

Figura 1: Modelo de clasificación de emociones de Russell [5].	11
Figura 2: ECG del corazón con ritmo sinusal normal [8].	14
Figura 3: Derivaciones precordiales [10]. Figura 4: Derivaciones del plano frontal [10].	14
Figura 5: Señal SCR [14].	15
Figura 6: Señales en bruto (SC) y sus componentes tónica (SCL) y fásica (SCR)[14].	16

2. MATERIALES Y MÉTODOS..... 20

Figura 7: Esquema general del procedimiento para la clasificación emocional a partir de señales fisiológicas.	20
Figura 8: Representación del canal 47, Status.	23
Figura 9: Señal ECG de MAHNOB-HCI sin filtrar.	25
Figura 10: Señal ECG de MAHNOB-HCI sin filtrar ampliada.	26
Figura 11: Señal ECG de MAHNOB-HCI filtrada.	26
Figura 12: Señal ECG de MAHNOB-HCI filtrada ampliada.	27
Figura 13: Señal EDA MAHNOB-HCI sin filtrar.	27
Figura 14: Señal EDA MAHNOB_HCI filtrada.	28
Figura 15: Detección de picos R con el algoritmo de Pan-Tompkins.	29
Figura 16: Red Neuronal Artificial [38].	39
Figura 17: Árbol de decisión [41].	40
Figura 18: Histograma de las etiquetas feltEmo en la base de datos MAHNOB-HCI.	42
Figura 19: Histograma de los cuatro cuadrantes emocionales según las etiquetas de feltEmo en la base de datos MAHNOB-HCI.	43
Figura 20: Sensor SparkFun AD8232 [45].	46
Figura 21: Electrodo para SparkFun AD8232 [46].	46
Figura 22: Colocación de electrodos según la página web del fabricante [45].	47
Figura 23: Señal ECG con SparkFun AD8232.	48
Figura 24: Sensor Grove-GSR de Seeed Studio [48].	48
Figura 25: Señal EDA obtenido con el sensor Grove-GSR V1.2.	49

3. RESULTADOS..... 50

Figura 26: Matriz de confusión normalizada para la clasificación binaria de la valencia con RF.	50
Figura 27: Matriz de confusión normalizada para la clasificación binaria de la valencia con RNA.	51
Figura 28: Matriz de confusión normalizada para la clasificación binaria de la excitación con RF.	52
Figura 29: Matriz de confusión normalizada para la clasificación binaria de la excitación con RNA.	53
Figura 30: Matriz de confusión normalizada para la clasificación multiclase de los 4 cuadrantes emocionales con RF.	54
Figura 31: Matriz de confusión normalizada para la clasificación multiclase de los 4 cuadrantes emocionales con RNA.	55

ÍNDICE TABLAS

2. MATERIALES Y MÉTODOS..... 20

Tabla 1: Asignación y configuración de canales en DEAP Dataset.	22
Tabla 2: Asignación y configuración de canales en MAHNOB-HCI Dataset.	23
Tabla 3: Correspondencia entre etiqueta de feltEmo y emoción reportada en MAHNOB-HCI.	24
Tabla 4: Comparación de las bases de datos MAHNOB-HCI y DEAP.	24
Tabla 5: Características extraídas de ECG, EDA y SKT.....	35
Tabla 6: Características seleccionadas para el modelo de clasificación binaria de la valencia.	37
Tabla 7: Características seleccionadas para el modelo de clasificación binaria de la excitación.	37
Tabla 8: Características seleccionadas para el modelo de clasificación multiclase.	38
Tabla 9: Correspondencia de etiquetas, emociones y cuadrantes afectivos.	42
Tabla 10: Conexiones entre sensor SparkFun AD8232 y Arduino.	46
Tabla 11: Asignación de colores de electrodos de SparkFun AD8232 y su ubicación según la web del fabricante.	47
Tabla 12: Asignación de colores de electrodos de SparkFun AD8232 y su ubicación correcta.	47
Tabla 13: Conexiones entre sensor Grove-GSR V1.2 y Arduino.....	49

3. RESULTADOS..... 50

Tabla 14: Resultado del rendimiento del modelo RF de clasificación binaria de la valencia.	50
Tabla 15: Resultado del rendimiento del modelo RNA de clasificación binaria de la valencia.....	51
Tabla 16: Resultado del rendimiento del modelo RF de clasificación binaria de la excitación.	52
Tabla 17: Resultado del rendimiento del modelo RNA de clasificación binaria de la excitación.....	52
Tabla 18: Resultado del rendimiento del modelo RF de clasificación multiclase de los 4 cuadrantes emocionales.	53
Tabla 19: Resultado del rendimiento del modelo RNA de clasificación multiclase de los 4 cuadrantes emocionales.	54
Tabla 20: Tabla comparativa del estado del arte.	57

4. DISCUSIÓN..... 56

Tabla 20: Tabla comparativa del estado del arte.	57
---	----

ABREVIATURAS Y ACRÓNIMOS

FC – Frecuencia cardíaca

SNA – Sistema nervioso autónomo

SNC – Sistema nervioso central

SNP – Sistema nervioso periférico

SNS – Sistema nervioso somático

ECG – Electrocardiograma

EDA – Actividad electrodérmica

SKT – Temperatura corporal periférica

SCR – Respuestas de conductancia cutánea

SCL – Nivel de conductancia cutánea

GSR - *Galvanic skin réflex*

SC – Conductancia cutánea

SR – Resistencia cutánea

SI – Impedancia cutánea

SA – Admitancia cutánea

SVM – Máquina de vectores de soporte

EEG – Electroencefalograma

EMG – Electromiograma

EOG – Electrooculograma

HOS – Estadísticos de orden superior

HBF – Funciones de base Hermite

HRV – Variabilidad de la frecuencia cardíaca

FFT – Transformada rápida de Fourier

RFECV – *Recursive Feature Elimination with Cross-Validation*

RF – *Random forest*

RNA- Red neuronal artificial

VP – Verdaderos positivos

VN – Verdaderos negativos

FP – Falsos positivos

FN – Falsos negativos

1. INTRODUCCIÓN

1.1. MARCO TEÓRICO

1.1.1. Las emociones

El estudio de las emociones ha generado duda y confusión a lo largo de los años. Como señaló Wenger, Jones y Jones (1962), “casi todo el mundo piensa que sabe qué es una emoción hasta que intenta definirla. En ese momento casi nadie afirma poder entenderla” [1]. Es muy complicado hacer una simple definición de lo que son las emociones, y se han desarrollado diversas teorías.

Una forma común de entender las emociones es tratándolas como algo que sentimos y activa de forma simultánea tres sistemas de respuesta: cognitivo/subjetivo, conductual/expresivo y fisiológico/adaptativo. Cada uno de estos sistemas actúa de forma diferente en función de la emoción experimentada, de la persona e incluso de la situación en la que se produce el estímulo [1]. Esta perspectiva tridimensional de las emociones nos permite entender que la respuesta emocional no es algo que se observe únicamente a simple vista, sino que también afecta al organismo de las personas y, por ello, a sus respuestas fisiológicas.

Desde el punto de vista funcional, las emociones cumplen tres funciones esenciales: adaptativas, sociales y motivacionales. La función adaptativa ya fue planteada por Darwin en 1872, señalando que las emociones nos permiten responder de forma adecuada a las diferentes situaciones con el objetivo de sobrevivir. En cuanto a la función social, las emociones permiten mejorar la interacción social y la comunicación afectiva, además de controlar la conducta de los demás. Por último, a nivel motivacional, las emociones nos permiten conseguir nuevas metas actuando como fuente de energía y dirigiendo la conducta en función de la emoción percibida. Toda acción motivada genera una respuesta emocional, pero, además, una emoción puede ser un factor desencadenante de la conducta motivada, permitiendo a las personas alcanzar sus objetivos [1].

En cuanto a su clasificación, se han propuesto dos principales enfoques: el categorial o discreto y el dimensional. El enfoque discreto afirma que hay un conjunto de emociones básicas limitado, que son universales e innatas. En 1994, Paul Ekman diferenció seis emociones básicas que no están influenciadas por la cultura de las personas, es decir, que son universales: alegría, miedo, tristeza, ira, disgusto, y sorpresa. Se les consideran básicas ya que están relacionadas con la adaptación y la supervivencia del ser humano. Además de estas emociones básicas, también hay unas emociones secundarias derivadas de éstas que sí tienen influencia cultural como, por ejemplo, la culpa [2]. Autores como Izard también han apoyado la existencia de las emociones básicas, señalando que están asociadas a funciones cerebrales y a unas expresiones faciales concretas, así como a propiedades adaptativas y motivacionales propias [1].

Este enfoque discreto ha planteado dificultades en el estudio de las emociones, ya que tratar con listas de emociones extensas no siempre es óptimo ni asegura una buena comprensión de la dinámica y las relaciones entre emociones. Por esta razón, se planteó trabajar con un enfoque dimensional. Uno de los modelos más reconocidos es el propuesto por el psicólogo James Russel,

en el cual se plantea un plano bidimensional compuesto por dos variables: la valencia, representada en el eje horizontal (x), y la activación o excitación (*arousal* en inglés) representado en el eje vertical (y). Las distintas emociones se clasifican en este espacio en función de las dos dimensiones y es lo que se conoce como el modelo circunflejo de Russell. La valencia es un indicativo del grado de agrado o desagrado que genera la emoción. A la izquierda del eje se sitúan las emociones que generen sensaciones negativas, mientras que aquellas asociadas a experiencias agradables se sitúan a la derecha. Por otro lado, la excitación refleja el nivel de energía o intensidad emocional, desde el estado de calma, situado en la parte inferior del eje y, hasta uno de alta excitación en la parte superior [3]. Algunos autores han añadido una tercera dimensión, el grado de control o dominancia, que refleja la sensación de control sobre una emoción [4].

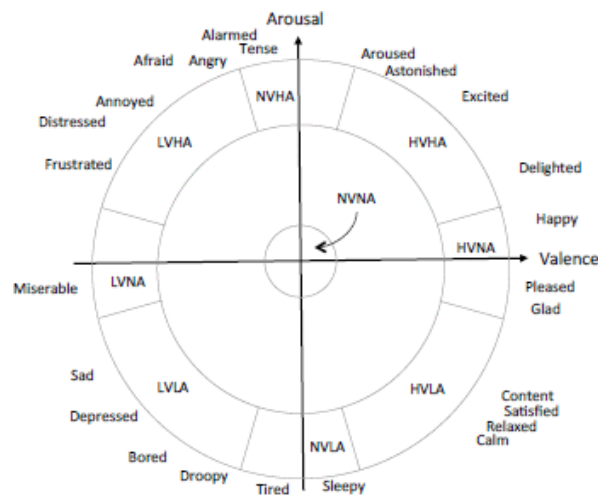


Figura 1: Modelo de clasificación de emociones de Russell [5].

El modelo circunflejo de Russell no es el único que existe, pero nos centramos en él debido a su sencillez y eficacia, especialmente en contextos de reconocimiento de emociones, donde se requiere una codificación clara. Este modelo permite representar los estados emocionales de manera continua facilitando su clasificación automática [4].

Desde el punto de vista fisiológico, se han desarrollado diferentes teorías para explicar la relación entre las emociones y la actividad del sistema nervioso central y autónomo. Uno de los modelos más influyentes fue la teoría de James-Lange, que afirmaba que las emociones eran el resultado de la percepción de cambios fisiológicos. Es decir, no lloramos porque estamos tristes, sino que estamos tristes porque lloramos. Esto se fundamenta en que cada emoción se identifica con un patrón fisiológico diferenciado [1].

Algunos patrones fisiológicos detectados han sido los siguientes:

- Frecuencia cardíaca (FC): aumenta con emociones como miedo, ira o tristeza.
- Sudoración: se incrementa en estados emocionales intensos, especialmente en emociones negativas de baja valencia como la tristeza, miedo, ira y asco.
- Temperatura periférica: Las emociones de baja valencia y alta excitación provocan una disminución en la temperatura periférica, mientras que las emociones positivas y calmadas

la aumentan. Por ejemplo, la ira se asocia con un aumento de la temperatura, mientras que el miedo puede hacer que se reduzca [1].

Este modelo fue posteriormente puesto en duda por autores como Walter Cannon, quien defendía que las emociones no se explican solo por las reacciones fisiológicas, sino que el cerebro es el que tiene el papel principal. Así surgió la teoría emergentista, donde es el cerebro, en particular el tálamo, el que interpreta lo que está pasando y activa tanto la emoción que sentimos como las reacciones del cuerpo. Desde esta teoría, la rama simpática del sistema nervioso autónomo (SNA) prepara al organismo para una reacción intensa ante el estímulo (por ejemplo, huida), mientras que el sistema parasimpático ayuda a restablecer el equilibrio tras la respuesta emocional. Estas respuestas no definen el tipo de emoción experimentada, pero sí su intensidad [1].

En resumen, las emociones no solo tienen una manifestación externa y conductual, sino que implican cambios fisiológicos medibles que cumplen funciones adaptativas. Estos cambios son relevantes tanto para la supervivencia como para las relaciones sociales, y han sido utilizados en ámbitos como el reconocimiento automático de emociones a través de señales fisiológicas.

1.1.2. Señales fisiológicas

Como se ha comentado anteriormente, hay una componente fisiológica muy relevante en el estudio de las emociones. Los cambios que se producen en el organismo como respuesta a un estímulo emocional son regulados por el SNA y se pueden medir objetivamente, lo que resulta especialmente útil en investigación.

Desde el punto de vista biológico, el sistema nervioso se divide en el sistema nervioso central (SNC) y sistema nervioso periférico (SNP). A su vez, el SNP se divide en sistema nervioso somático (SNS) y SNA. Este último participa en la regulación de las funciones viscerales como la actividad cardíaca, respiratoria o glandular, y está muy relacionado con la respuesta fisiológica de las emociones.

El SNA está compuesto por neuronas sensoriales y motoras que conectan el SNC con los órganos internos como el corazón, pulmones o glándulas. Estas neuronas generan una actividad eléctrica cuando se excitan en respuesta a impulsos despolarizantes, lo que da lugar a potenciales de acción. Estos impulsos eléctricos generan las señales fisiológicas que se miden mediante sensores específicos, y reflejan la actividad eléctrica generada por estímulos emocionales, en este caso.

A continuación, se describen las principales señales fisiológicas asociadas a los estados emocionales que serán analizadas en este trabajo: electrocardiograma (ECG), la actividad electrodérmica (EDA), y la temperatura corporal periférica (SKT).

1.1.2.1. Señal ECG

La señal de ECG representa la actividad eléctrica del corazón. Su utilidad va más allá del ámbito clínico, y es que es una de las señales más utilizadas para estudios de las emociones y el estrés.

El corazón es un órgano ubicado en el mediastino, entre los dos pulmones, que tiene una forma cónica. Está dividido en cuatro cavidades, dos aurículas (derecha e izquierda) en la parte superior, que actúan como cámaras que reciben la sangre, y dos ventrículos (derecho e izquierdo) en la parte inferior encargados de bombear la sangre hacia los pulmones y al resto del cuerpo. La función del corazón es actuar como una bomba, donde llega sangre desoxigenada y envía a los tejidos sangre oxigenada. Esto ocurre gracias al musculo cardíaco, llamado miocardio, que se contrae y se relaja impulsando la sangre hacia las distintas cavidades a través de las válvulas cardíacas [6].

El ciclo cardíaco es un conjunto de eventos mecánicos, eléctricos, cambios de presión, flujo y volumen de sangre que ocurre en el corazón durante cada latido. Se repite entre 60 y 100 veces por minuto en reposo y se divide en dos grandes fases: la sístole y la diástole.

La diástole corresponde a la fase donde el corazón se relaja y se llena de sangre, y tiene dos etapas. La primera es un llenado rápido donde la sangre fluye de forma pasiva desde las aurículas a los ventrículos, ya que las válvulas auriculoventriculares (mitral a la izquierda y tricúspide a la derecha) están abiertas. La segunda fase es la contracción auricular para enviar el pequeño volumen de sangre sobrante a los ventrículos.

La sístole es la fase en la que el corazón se contrae y expulsa la sangre a las arterias y también tiene dos etapas. La primera es la contracción isovolumétrica donde los ventrículos comienzan a contraerse con las válvulas aún cerradas, lo que aumenta la presión interna. La segunda es la eyección, donde la presión ventricular es mayor que la presión arterial y las válvulas semilunares (aórtica y pulmonar) se abren y se expulsa la sangre. Tras la eyección tiene lugar la relajación isovolumétrica donde los ventrículos comienzan a relajarse y las válvulas semilunares se cierran. La presión disminuye rápidamente hasta que empieza un nuevo ciclo [6].

El ECG registra, a través de electrodos ubicados en la superficie corporal, las corrientes eléctricas que se generan durante la despolarización y repolarización del miocardio. Este registro se visualiza en una serie de ondas que corresponden con diferentes fases del ciclo cardíaco: la onda P, el segmento P-R, el complejo QRS, el segmento ST y la onda T.

- La onda P representa la despolarización auricular, donde las aurículas se contraen y envían sangre a los ventrículos.
- En el segmento P-R las aurículas terminan de vaciarse antes de la contracción ventricular.
- El complejo QRS indica la despolarización ventricular.
- El segmento ST es el intervalo de tiempo entre el final del complejo QRS y el inicio de la onda T, y tiene un gran valor diagnóstico ya que su elevación o descenso con respecto a la línea basal puede indicar insuficiencia cardíaca.
- La onda T corresponde a la repolarización ventricular donde el corazón se relaja una vez expulsada la sangre de los ventrículos [7].

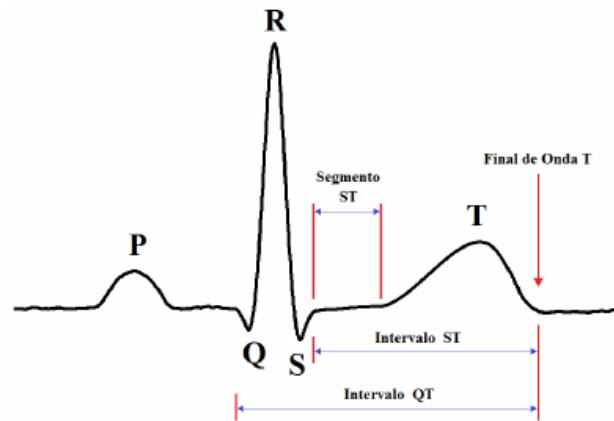


Figura 2: ECG del corazón con ritmo sinusal normal [8].

Para interpretar el ECG es necesario entender el significado de las derivaciones, su orientación y su polaridad. El ECG registra el paso de impulsos eléctricos en forma de vectores, que generan ondas positivas, si se dirigen al electrodo positivo de la derivación, o negativas si se alejan de él. El ECG estándar tiene 12 derivaciones, distribuidas en dos planos principales, el plano frontal con las seis derivaciones I, II, III, aVR, aVL y aVF, y el plano horizontal o transversal formado por las derivaciones precordiales (de V1 a V6) [9].

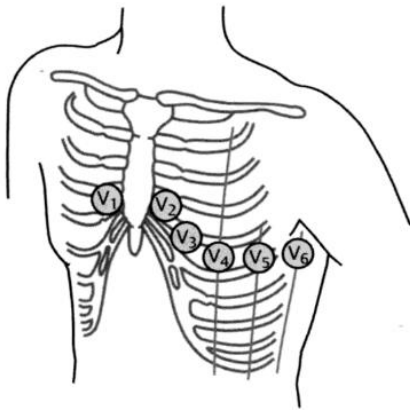


Figura 3: Derivaciones precordiales [10].

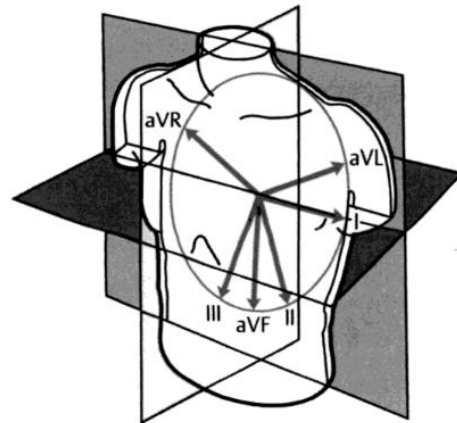


Figura 4: Derivaciones del plano frontal [10].

Para adquirir un ECG con 12 derivaciones se necesitan 10 electrodos, pero para la monitorización básica de la actividad eléctrica del corazón es suficiente con visualizar las derivaciones I, II y III utilizando 3 electrodos, usualmente en los brazos derecho e izquierdo y en la pierna izquierda. Este estudio se llevará a cabo utilizando esta última metodología.

Estas tres derivaciones son bipolares, es decir, se mide la diferencia entre dos electrodos. La derivación I mide la diferencia entre el brazo izquierdo y el brazo derecho, la derivación II mide la diferencia entre el brazo derecho y la pierna izquierda, y la III la diferencia entre en brazo y pierna izquierdos. Estas tres derivaciones forman el Triángulo de Einthoven [11]

1.1.2.2. Señal EDA

La señal EDA es una señal fisiológica que refleja los cambios en la conductancia eléctrica de la piel, los cuales están muy relacionados con la actividad del sistema autónomo, en concreto, con las glándulas sudoríparas.

La EDA se basa en el principio de que la piel del ser humano presenta una variación en su conductividad eléctrica en función de la sudoración. Cuando se experimenta una emoción intensa como estrés o ansiedad el sistema nervioso simpático activa las glándulas sudoríparas, lo que hace que la humedad de la piel aumente y así su conductancia eléctrica, siendo así un buen indicador fisiológico de las activaciones del cerebro [12]. Este cambio en la conductancia aparece aproximadamente dos segundos después de la estimulación, y alcanza el máximo después de dos a diez segundos, donde comienza a disminuir aproximadamente a la misma velocidad [13]

En cuanto a la adquisición de la señal, se suele registrar la conductancia mediante electrodos colocados en los dedos índice y corazón en contacto con la piel. Se registran las variaciones en la conductancia expresadas en términos de respuestas electrodérmicas conocidas como respuestas de conductancia cutánea (SCR, por sus siglas en inglés). La amplitud de las SCR varía en función de las activaciones del SNC que actúan sobre las glándulas sudoríparas [12].

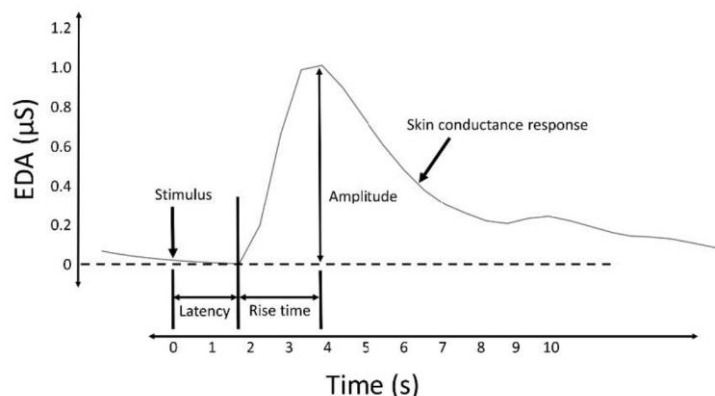


Figura 5: Señal SCR [14].

De esta forma, las SCR son un buen indicador del impacto cerebral de las emociones sobre el cuerpo. Durante la estimulación cerebral, la amplitud de las SCR aumenta con la activación emocional, independientemente de su valencia [15].

A lo largo del siglo XX se han desarrollado una gran diversidad de técnicas de registro de las variaciones eléctricas de la piel, dando lugar a una forma confusa de nombrarlas. Uno de los términos más usados en las bases de datos es “galvanic skin reflex” (GSR), pero se utilizaron distintas nomenclaturas como: “skin potential reflex”, “neuro-galvanic response”, “skin potential response”, “skin resistance response”, “skin conductance response” o “electrodermal response”. Finalmente, en 1981 el término de “Electrodermal Activity” (EDA) se designó como el estándar internacional para la actividad eléctrica cutánea espontánea e inducida, por lo que es el que se usará a lo largo de este trabajo [12].

Se pueden diferenciar diferentes mediciones fisiológicas como la conductancia cutánea (SC), la resistencia cutánea (SR), la impedancia cutánea (SI) y la admitancia cutánea (SA). Estas técnicas permiten registrar las variaciones eléctricas de la piel inducidas por un voltaje externo, explicadas por la ley de Ohm. De acuerdo con esta ley, al inducir un voltaje constante de 0,5V a 1V se pueden medir las variaciones en la intensidad de una corriente externa dependientes de la resistencia de la piel [12].

El método internacionalmente reconocido para el registro de EDA es mediante la resistencia de la piel, la SR, de donde se extraen dos componentes: *skin conductance level* (SCL), y *skin conductance response* (SCR). SCL se trata de la parte tónica de la señal de SC, tiene una baja frecuencia por lo que cambia lentamente con el tiempo. Es un reflejo del estado general de excitación, es decir, si estas tranquilo o nervioso durante un periodo largo de tiempo.

Por otro lado, SCR es la parte fásica de la SC. Tiene una alta frecuencia por lo que son aumentos breves y rápidos de la conductancia producidos por un estímulo específico que afecta al estado emocional. Esta componente es un reflejo de las reacciones inmediatas del sistema nervioso simpático [12].

La componente de interés en los estudios de excitación emocional es la SCR ya que, al estar asociada al sistema nervioso simpático no puede ser controlada de manera consciente. Además, la SCL tiene una gran dependencia de factores como la sequedad e hidratación de la piel y una gran variabilidad intersujeto.

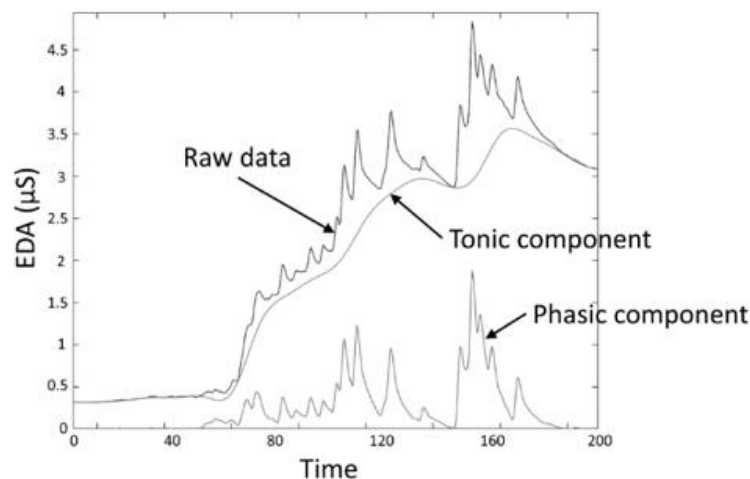


Figura 6: Señales en bruto (SC) y sus componentes tónica (SCL) y fásica (SCR)[14].

1.1.2.3. Señal SKT

La señal SKT mide la respuesta térmica de la piel del cuerpo humano. Esta señal proviene de las variaciones en la temperatura provocadas por cambios en el flujo sanguíneo debido a la resistencia de los vasos y presión arterial. La resistencia vascular está regulada por el sistema nervioso

simpático, que controla su musculatura generando vasoconstricción o, por el contrario, vasodilatación. La presión arterial también es regulada por el sistema nervioso autónomo mediante ajustes en la actividad cardiovascular. Por estas razones, se demuestra que se puede evaluar el estado emocional y la actividad del sistema nervioso mediante la señal de SKT [16].

En [17] se hizo un estudio donde 100 personas escuchaban dos tipos de música, una que generaba emociones negativas y excitantes y la otra positivas y calmadas. Al escuchar el primer tipo de música, las emociones negativas y excitantes produjeron una disminución de la temperatura corporal. Por otro lado, con la música que generaba sensación de calma y emociones positivas aumentó.

Estos cambios se explican ya que en las emociones de excitación la actividad simpática aumenta generando una vasoconstricción y disminución de la temperatura, mientras que el efecto de la actividad parasimpática en situaciones de calma favorece la vasodilatación.

1.2. JUSTIFICACIÓN Y PLANTEAMIENTO DEL PROBLEMA

La detección de emociones a partir de señales fisiológicas presenta un desafío relevante dentro del ámbito del aprendizaje automático. En este trabajo se trata de demostrar la posibilidad de predecir el estado emocional de las personas utilizando las señales fisiológicas de ECG, EDA y SKT. No obstante, se identifican múltiples dificultades en el estudio.

Una de las principales complicaciones es la alta variabilidad inter e intrasujeto de las respuestas fisiológicas ante estímulos emocionales, lo que dificulta la capacidad de los modelos predictivos para generalizar. Dos individuos pueden presentar respuestas fisiológicas diferentes ante un mismo estímulo, y un mismo individuo puede reaccionar de forma diferente a un mismo estímulo en función del contexto. Además, en algunas ocasiones, los sujetos no generan ninguna respuesta ante estímulos emocionales [18]. Estas dificultades afectan de forma notable la capacidad de etiquetar las emociones en términos de las dimensiones de valencia (agradable - desagradable) y excitación o *arousal*, ya que no existe un indicador único y universal.

Además, aunque se ha avanzado considerablemente en el uso de aprendizaje automático para el reconocimiento de emociones, muchos de estos modelos requieren de grandes cantidades de datos de entrada y con varias señales fisiológicas, en ocasiones, difíciles o incómodas de obtener, como el EEG. Por ello, es conveniente explorar alternativas con las señales periféricas empleadas en este trabajo.

En estudios previos como [19] se trató de predecir el estado afectivo de las personas a partir de su valencia y excitación mediante las señales fisiológicas de ECG, respiración, SKT y EDA. Utilizaron la base de datos MAHNOB-HCI y una Máquina de Vectores de Soporte (SVM) como modelo de clasificación. Los resultados obtenidos fueron un *accuracy* del 68,75% en la clasificación de la valencia y un 64,23 % en la clasificación de la excitación. En [20] se realizó un estudio similar utilizando únicamente las señales de respiración y ECG, con la misma base de datos MAHNOB-HCI y utilizando un SVM multiclase como clasificador. Se obtuvo un *accuracy*

de 60,41% para excitación y un 59,57% para valencia. Por último, en el artículo original de la base de datos MAHNOB-HCI [21] utilizaron registros de ECG, EDA, SKT y respiración y un SVM para alcanzar un *accuracy* de 46,2% para la excitación y un 45,5% para la valencia.

En este trabajo se propone un modelo predictivo alternativo, centrado exclusivamente en las señales fisiológicas de ECG, EDA y SKT utilizando modelos de clasificación binaria para las dimensiones emocionales de valencia y excitación. Además, se aplican técnicas específicas de extracción y selección de características para cada dimensión emocional, con el fin de mejorar la exactitud y capacidad de generalización.

Las principales contribuciones de este trabajo son:

- La validación del uso exclusivo de señales fisiológicas periféricas como ECG, EDA y SKT para la clasificación binaria de valencia y excitación utilizando la base de datos MAHNOB-HCI.
- Diseñar un flujo de procesamiento de señales que permita la extracción y selección de las características más relevantes para la predicción de valencia y excitación de forma eficiente.
- Demostrar que un modelo Random Forest puede alcanzar una precisión competitiva con la de los estudios actuales.

Este trabajo crea una base para futuras aplicaciones en las que la predicción de emociones basada en señales fisiológicas pueda suponer una aportación positiva. Un ejemplo es su aplicación en juegos serios empleados en rehabilitación, donde se ha visto que el *biofeedback* visual, es decir, proporcionar al sujeto una representación visual en tiempo real de sus respuestas fisiológicas, puede mejorar la terapia [22].

1.3. HIPÓTESIS

En este trabajo se presenta la hipótesis de que es posible desarrollar un modelo computacional eficaz de detección emocional a partir de las señales fisiológicas de ECG, EDA y SKT. Se espera que este modelo permita clasificar los estados emocionales de forma correcta en términos de valencia y excitación.

1.4. OBJETIVOS

El objetivo principal de este trabajo es el desarrollo de un modelo de clasificación capaz de detectar de forma objetiva los estados emocionales en términos de valencia y excitación, utilizando exclusivamente las señales fisiológicas de ECG, EDA y SKT.

Los objetivos específicos de este trabajo son los siguientes:

- Seleccionar y analizar una base de datos adecuada para este estudio, en este caso, MAHNOB-HCI, comprendiendo su estructura y características.
- Diseñar y aplicar un proceso de preprocesamiento de señales fisiológicas para eliminar ruido y artefactos que afectan a la calidad de los datos.

- Extraer de las características de las señales fisiológicas que serán la entrada de los modelos predictivos.
- Seleccionar las características más relevantes de cada señal para cada dimensión emocional. De esta forma se reduce la dimensionalidad del conjunto de datos y se elimina información redundante o irrelevante. El objetivo es mejorar la eficiencia del modelo y reducir el riesgo de sobreajuste.
- Evaluar y comparar el desempeño de varios modelos predictivos, en concreto Random Forest y Redes Neuronales.
- Validar los modelos utilizando la base de datos MAHNOB-HCI y analizar su capacidad de generalización.
- El registro y procesamiento de señales propias de ECG, EDA y SKT para evaluar el rendimiento del modelo con datos reales obtenidos por mí utilizando los sensores correspondientes.

2. MATERIALES Y MÉTODOS

Una vez explicados los conocimientos necesarios para abordar este trabajo, se ha desarrollado el modelo predictivo. El planteamiento metodológico que se ha seguido se puede observar en la Figura 7, donde se esquematiza un flujo de procesamiento de datos desde la adquisición de las señales fisiológicas hasta la etapa final de clasificación mediante modelos de aprendizaje automático.

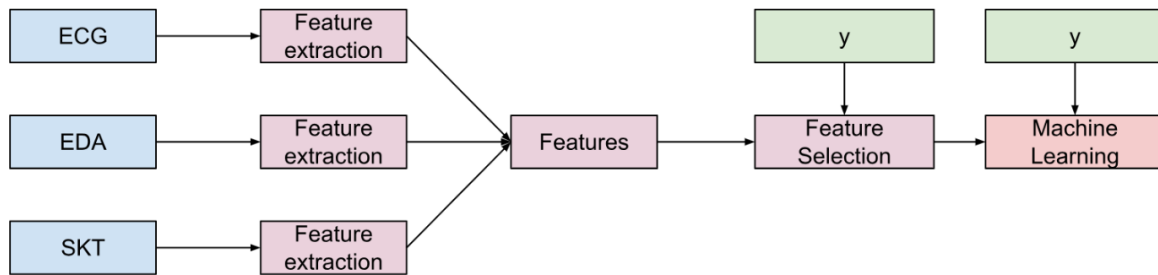


Figura 7: Esquema general del procedimiento para la clasificación emocional a partir de señales fisiológicas.

Como se muestra en la Figura 7, el procedimiento se divide en las siguientes etapas:

1. **Adquisición y procesamiento de las señales fisiológicas** desde la base de datos MAHNOB-HCI.
2. **Extracción de características:** a partir de cada una de las señales fisiológicas (ECG, EDA y SKT) se extraen las características que describen a cada señal y se fusionan todas en un único conjunto de datos, que sirve como entrada de las siguientes etapas.
3. **Selección de características:** Se aplica una técnica de selección de características con el fin de obtener el mejor rendimiento posible del modelo. Estas características son las que aportan una mayor cantidad de información para optimizar el proceso de entrenamiento del modelo.
4. **Implementación del modelo predictivo con aprendizaje automático:** Se entrenan modelos predictivos utilizando las características seleccionadas como entrada, y las etiquetas como salida. Este proceso se realiza individualmente para cada dimensión emocional.

Se han desarrollado tres modelos predictivos independientes:

1. Modelo de clasificación binaria de la valencia: este modelo se encarga de predecir si el sujeto presenta valencia positiva o negativa.
2. Modelo de clasificación binaria de la excitación: en este caso se predice si la excitación es alta o baja.
3. Modelo de clasificación multiclase de los cuadrantes emocionales del modelo circunflejo: Se implementó este tercer modelo para evaluar si proporciona mejor resultado utilizar

directamente las emociones específicas (alegría, ira, tristeza...) en lugar del valor de valencia y excitación.

2.1. BASE DE DATOS

En la detección de emociones a partir de señales fisiológicas es necesario el uso de bases de datos para entrenar y validar los modelos predictivos. Estas bases de datos se componen de registros de las señales obtenidos bajo condiciones controladas, que reflejan las respuestas fisiológicas de sujetos ante estímulos emocionales. De esta forma es posible comparar resultados de diferentes estudios, reproducir los experimentos y mejorar la generalización de los modelos desarrollados.

Existen varias bases de datos en las que sujetos voluntarios son sometidos a distintos estímulos con el fin de generar una respuesta emocional. Estas emociones pueden ser etiquetadas con adjetivos específicos (como alegría, miedo, etc.) o mediante valores numéricos en función de si se usa un modelo categorial o dimensional, respectivamente.

Según la bibliografía las bases de datos más utilizadas en estudios de detección de emociones con anotación de valencia y excitación y utilizando las señales de ECG, EDA y SKT son DEAP *dataset* y MAHNOB-HCI, ya que presentan los mejores resultados en investigaciones previas y tienen una buena calidad en los registros [23].

2.1.1. DEAP Dataset

La base de datos DEAP (*Database for Emotion Analysis using Physiological signals*) se trata de un repositorio de datos multimodales para el estudio de la identificación del estado emocional. En ella se pueden ver señales de electroencefalograma (EEG), ECG, EDA, electromiograma (EMG), electrooculograma (EOG), respiración, SKT y vídeos faciales de 32 participantes. Estas señales fueron registradas mientras veían 40 vídeos musicales de un minuto de duración cada uno. Tras visualizar estos vídeos, los participantes calificaron cada video en función de los niveles de excitación, valencia, grado de dominancia, familiaridad y agrado en una escala de 1 a 9. Además, se grabaron las expresiones faciales de 22 participantes [24].

Los videos musicales utilizados en este experimento fueron seleccionados previamente a través de un proceso de etiquetado afectivo automático mediante una web llamada Last.fm, que asigna etiquetas a las canciones en función la emoción que genera [24].

Esta base de datos es de acceso público y contiene los datos preprocesados en formatos compatibles con MATLAB y Python. Además de las señales crudas, incluye las señales filtradas, alineadas temporalmente y con una frecuencia de muestreo reducida de 512 Hz a 128 Hz para facilitar su análisis [24].

Cada archivo de los participantes contiene dos matrices:

- data: se trata de una matriz de dimensiones (40 x 40 x 8064), donde cada eje representa, respectivamente, el video, el canal y los datos de la señal.

- labels: en este caso es una matriz de 2 dimensiones (40 x 4), donde las filas corresponden a los vídeos y cada columna representa una de las cuatro dimensiones evaluadas (valencia, excitación, dominancia y agrado)

La forma en la que se organizan los canales en DEAP es la siguiente:

Canal	Señal
1-32	EEG
33, 34	EOG horizontal, EOG vertical
35	EMG
36	ECG
37	EDA
38	Respiración
39	Fotopletismografía (PPG)
40	SKT

Tabla 1: Asignación y configuración de canales en DEAP Dataset.

2.1.2. MAHNOB-HCI

MAHNOB-HCI se trata de una base de datos multimodal creada con el fin de investigar acerca del reconocimiento de emociones y el etiquetado implícito. Se trata de una base de datos pública disponible para la comunidad académica a través de su página web. Consta de varios registros multimodales de respuestas fisiológicas ante estímulos afectivos, incluyendo la grabación simultánea de vídeos faciales, señales de audio, rastreo ocular y señales fisiológicas del sistema nervioso periférico y central (ECG, EEG, EDA, temperatura corporal y respiración) [21].

Para crear esta base de datos se realizaron dos experimentos con 30 participantes de ambos sexos (17 mujeres y 13 hombres), de diferentes orígenes culturales y entre los 19 y los 40 años con distintos niveles educativos. En el primer experimento, los sujetos visualizaron 20 videos que generaban una respuesta emocional divididos en varios fragmentos de diferente duración. Posteriormente, rellenaron un informe donde reflejaban las emociones percibidas mediante la excitación, la valencia, la dominancia y la predictibilidad, así como adjetivos que definen la emoción. En el segundo experimento los voluntarios vieron videos cortos e imágenes una vez sin etiquetas y otra con etiquetas correctas o incorrectas, y debían evaluar cómo de acuerdo estaban con las etiquetas mostradas [21]. Para este trabajo se han utilizado los datos obtenidos con el primer experimento.

Las señales fisiológicas se almacenan en archivos con formato .bdf de 47 canales, los 32 primeros recogen las señales de EEG en uV (de 32 canales), y el resto son los siguientes:

Canal	Nombre del canal	Posición	Unidades
33	EXG1	ECG1 (pecho lado derecho)	uV
34	EXG2	ECG1 (pecho lado izquierdo)	uV
35	EXG3	ECG1 (lado izquierdo abdomen)	uV
36-40	-	No usado	-
41	GSR1	GSR (dedo corazón y anular mano izquierda)	Ohm
40-44	-	No usado	
45	Resp	Sensores de respiración	uV
46	Temp	SKT (dedo meñique, mano izquierda)	Celsius
47	Status	Marcadores	Boolean

Tabla 2: Asignación y configuración de canales en MAHNOB-HCI Dataset.

Todos los archivos tienen 30 segundos antes y después del estímulo, y si no hay señal se rellena con ceros en todos los canales. El canal 47 (*Status*) indica el estado del experimento, señalando el inicio y el fin de los estímulos. Esto es para diferenciar entre la señal de respuesta y los intervalos de los 30 segundos de antes y después del estímulo. Se trata de una codificación binaria donde 1 indica el inicio y el fin del estímulo y 0 el resto del tiempo.

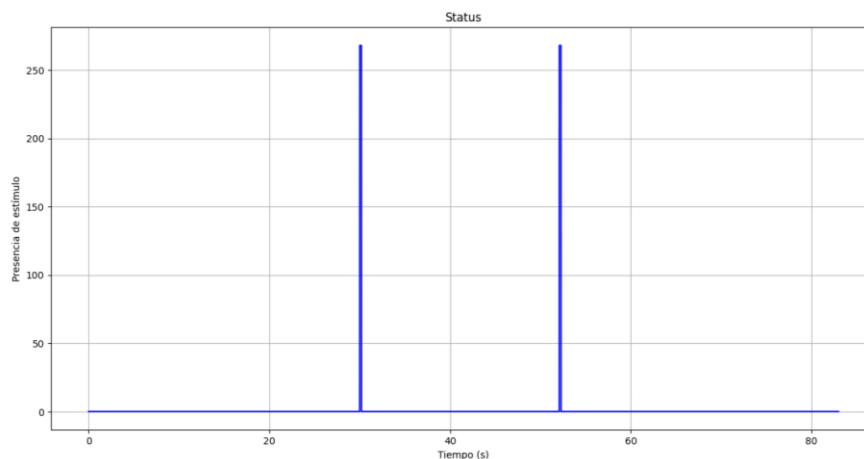


Figura 8: Representación del canal 47, *Status*.

Además de los archivos .bdf, la base de datos incluye unos archivos llamados “session.xml” que contienen los metadatos de cada sesión, es decir, las etiquetas. Incluye varias etiquetas sobre el aspecto físico del sujeto, como si tiene barba, bigote o gafas. Para el experimento 1, las etiquetas más relevantes fueron las respuestas del cuestionario realizado por cada participante tras cada vídeo:

- feltArsl (1-9): Indica el valor de excitación percibida por el sujeto, donde 1 representa ninguna excitación y 9 la excitación máxima.
- feltVlnc (1-9): Representa la valencia emocional percibida, donde 1 indica una emoción negativa, 9 una positiva y 5 corresponde a una emoción neutra.
- feltEmo (1-12): Corresponde al adjetivo que describe la emoción sentida:

feltEmo	Emoción
0	Neutral
1	Ira
2	Asco
3	Miedo
4	Alegría
5	Tristeza
6	Sorpresa
7	-
8	-
9	-
10	-
11	Diversión
12	Ansiedad

Tabla 3: Correspondencia entre etiqueta de feltEmo y emoción reportada en MAHNOB-HCI.

A continuación, se muestra una tabla comparativa de ambas bases de datos:

MAHNOB-HCI	DEAP
30 participantes.	32 participantes.
Señales crudas.	Señales crudas y preprocesadas.
Se mide la valencia y la excitación con valores entre 1 y 9 mediante autorreporte.	Se mide la valencia y la excitación con valores entre 1 y 9 mediante autorreporte.
Duración de prueba variable (34,9-117 segundos).	Duración de prueba fija (60 segundos).
Frecuencia de muestreo = 256 Hz.	Frecuencia de muestreo = 128 Hz.
Varios vídeos, duración total de 40 minutos.	40 vídeos de 1 minuto cada uno.
EEG (32 canales), ECG, EDA, SKT, Respiración, Vídeos de la cara y rastreador de mirada.	EEG (32 canales), ECG, EDA, EMG, EOG, Respiración, SKT, Vídeos de la cara.

Tabla 4: Comparación de las bases de datos MAHNOB-HCI y DEAP.

MAHNOB-HCI fue la seleccionada para este trabajo, ya que, según la bibliografía, es la que ofrece mejores resultados para la predicción del estado emocional con las señales utilizadas en este trabajo [19].

La adquisición de las señales fisiológicas en este repositorio ha consistido en la lectura de los archivos .bdf proporcionados en la carpeta “Sessions”, utilizando la librería MNE de Python, la cual permite acceder fácilmente a los canales del registro. Como se muestra en la Tabla 2, cada señal se encuentra registrada en un canal diferente. Para la señal de ECG, se utilizó el canal llamado “EXG1”, para EDA el canal “GSR1” y para la temperatura “Temp”.

Es importante tener en cuenta que los registros originales tienen una duración variable, pero todos tienen en común que los 30 primeros y últimos segundos no corresponden a la exposición de un estímulo emocional, por lo que estos segmentos fueron eliminados sistemáticamente en cada archivo. Posteriormente, se hizo un filtrado de las señales válidas para este estudio. Para garantizar una correcta extracción de características, especialmente aquellas relacionadas con la variabilidad, se estableció como criterio que la duración mínima de la señal debía ser de 60 segundos. De esta forma, de 1201 archivos disponibles en el repositorio, se seleccionaron 463 registros válidos para el análisis.

2.2. SEÑALES FISIOLÓGICAS

2.2.1. Señal ECG

Para el análisis de la señal de ECG se utilizó el canal correspondiente a EXG1 en la base de datos MAHNOB-HCI.

Como se ha comentado anteriormente, se han eliminado los primeros y los últimos 30 segundos de los registros ya que no pertenecían a una respuesta emocional, y se ha establecido como criterio una duración mínima de la señal de 60 segundos. La frecuencia de muestreo utilizada para la adquisición de esta señal fueron 256 Hz, por lo que trabajaremos con segmentos de mínimo 15360 muestras.

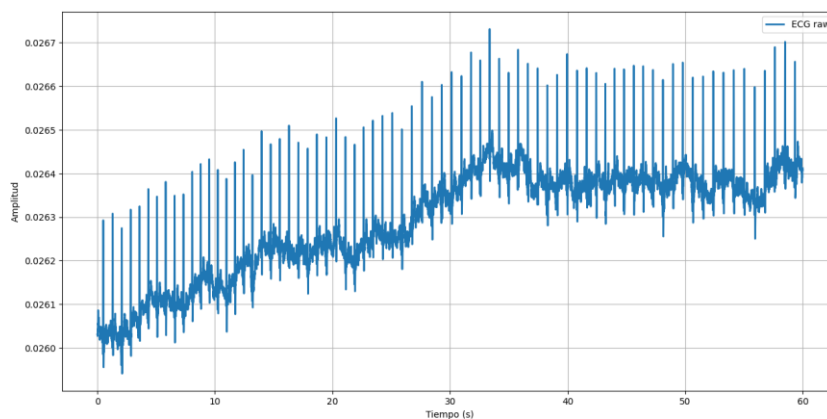


Figura 9: Señal ECG de MAHNOB-HCI sin filtrar.

Como se puede apreciar en la Figura 9, que muestra una señal de ECG original (sin procesado) de 60 segundos de duración, es necesario aplicar técnicas de filtrado para obtener una señal limpia y adecuada para su posterior análisis. Se puede ver presencia de deriva de la línea base a lo largo del tiempo, que es característica de movimientos del sujeto, problemas en el contacto de los electrodos, o ruido de baja frecuencia. Además, se puede apreciar ruido de alta frecuencia de forma clara si ampliamos la señal.

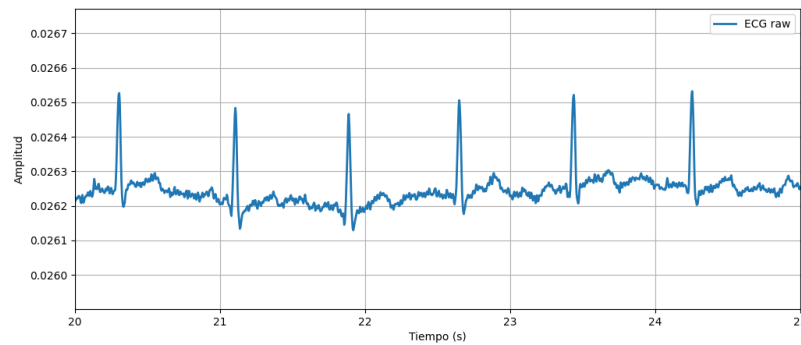


Figura 10: Señal ECG de MAHNOB-HCI sin filtrar ampliada.

Con el fin de mejorar la calidad de la señal y facilitar la posterior extracción de características, se aplicaron las técnicas de preprocesado requeridas. Para ello, se utilizaron principalmente las funciones del módulo *scipy.signal*, de la librería SciPy, que es muy utilizada en el procesamiento de señales biomédicas. Se diseñó un filtro paso banda de tipo Butterworth de cuarto orden, con frecuencias de corte de 0,5 y 40 Hz. De esta forma se consiguen eliminar tanto las componentes de baja frecuencia, como la deriva de la línea base, como las de alta frecuencia; sin distorsionar las componentes clave de la señal de ECG [25]. Se ha elegido un filtro Butterworth por su respuesta en frecuencia suave, lo que minimiza las distorsiones de la señal.

Para aplicar este filtrado se utilizó la función *filtfilt*, que realiza un filtrado con fase cero. Esto significa que no se producen desplazamientos de fase, lo que es fundamental en el análisis del ECG ya que para la extracción de características se analizan eventos temporales como los intervalos entre los picos R.

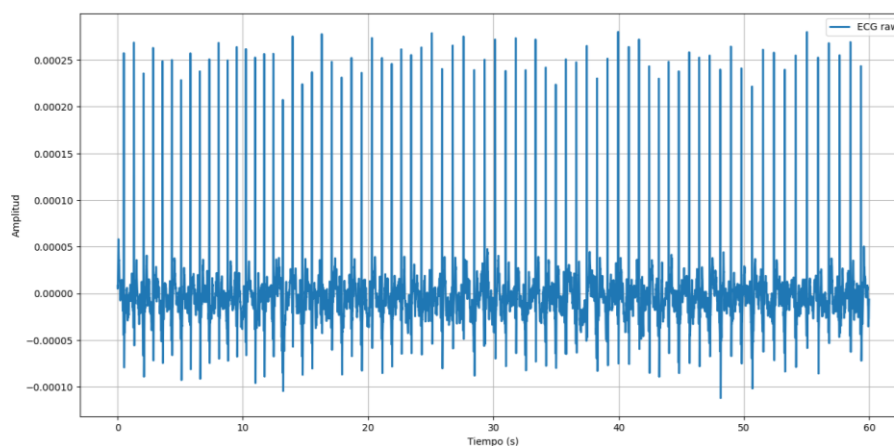


Figura 11: Señal ECG de MAHNOB-HCI filtrada.

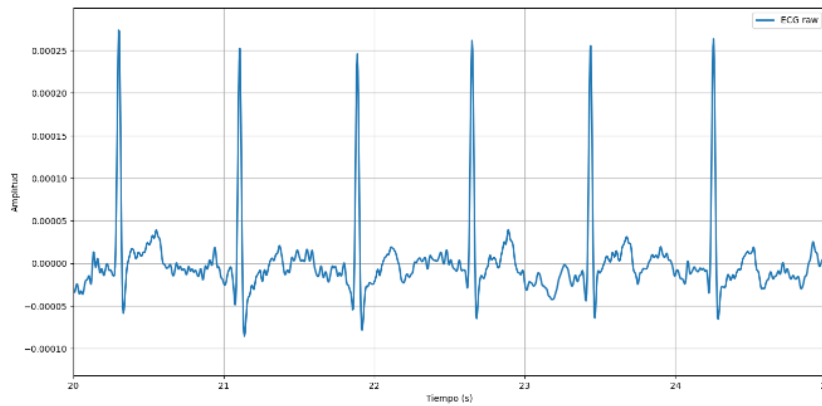


Figura 12: Señal ECG de MAHNOB-HCI filtrada ampliada.

Como se ve en la Figura 11 el filtrado ha eliminado la deriva de la línea base con éxito. Aunque aún persiste algo de ruido en la señal, se considera lo suficientemente limpia para su análisis.

2.2.2. Señal EDA

Para el análisis de la señal EDA, se cargaron las señales del canal GSR1 de la base de datos. Del mismo modo que para la señal de ECG, se eliminaron los primeros y los últimos 30 segundos del registro, y se estableció el criterio de duración mínima de 60 segundos. La frecuencia de muestreo también es de 256 Hz, por lo que se requiere un mínimo de 15360 muestras de la señal.

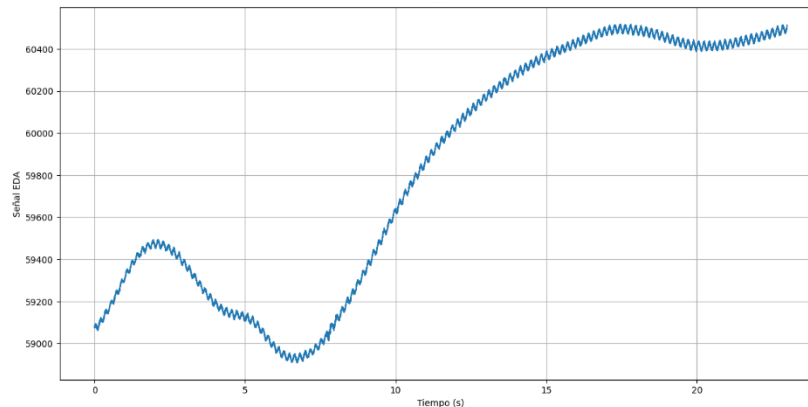


Figura 13: Señal EDA MAHNOB-HCI sin filtrar.

Como se puede apreciar hay presencia de ruido de alta frecuencia que debe ser eliminado. La banda principal de la señal EDA se encuentra entre los 0 y los 2.4 Hz [26], por lo que se ha aplicado un filtro Butterworth paso bajo con una frecuencia de corte de 2.4 Hz. Para evitar el desfase de la señal filtrada, se utilizó *filtfilt*.

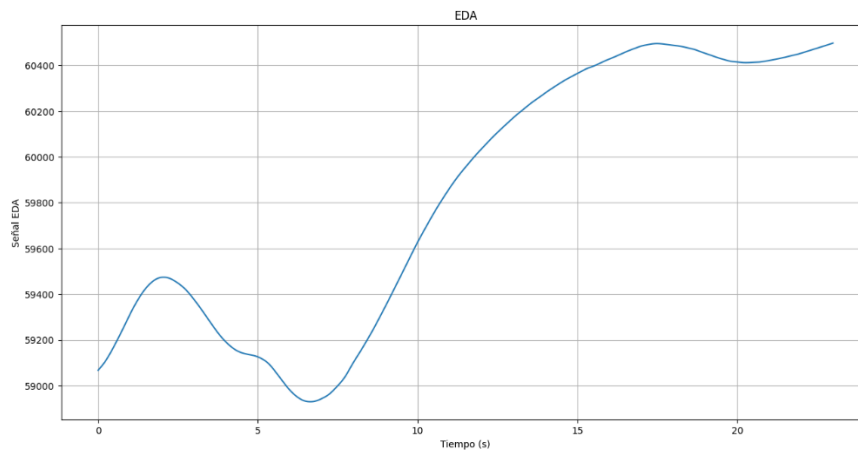


Figura 14: Señal EDA MAHNOB_HCI filtrada.

2.2.3. Señal SKT

Para la señal de SKT no ha sido necesario aplicar ningún preprocesado, ya que se pueden extraer directamente las características a partir de la señal original.

2.3. EXTRACCIÓN DE CARACTERÍSTICAS

Una vez preprocesadas las señales, se procede a la extracción de características. En este trabajo se ha implementado un sistema robusto de extracción de características de las señales de ECG, EDA, y SKT. El objetivo de esto es obtener, a partir de las señales, un conjunto de variables numéricas que representen la actividad fisiológica de cada sujeto ante el estímulo emocional.

Para asegurar la correcta obtención de las características se ha implementado una ventana temporal de los últimos 60 segundos de cada grabación para la extracción de características, partiendo de la hipótesis de que las reacciones emocionales más relevantes ocurren al final del estímulo. Esta duración es suficiente para poder calcular de forma correcta las características relacionadas con la variabilidad de la señal.

2.3.1. Características de la señal de ECG

1. Intervalos RR

El primer paso para la extracción de características de la señal de ECG fue la detección de picos R mediante el algoritmo Pan-Tompkins.

El pico R en el ECG corresponde con un máximo local de la señal, es decir, una zona donde la pendiente de la señal cambia de positiva a negativa. Por ello, el primer paso para la detección de este pico es calcular la derivada de la señal. Si la derivada es positiva, la señal está subiendo, y si es negativa bajando. Esto se comprueba buscando los índices donde la derivada pasa por cero de positivo a negativo, indicando un máximo local y posible pico R.

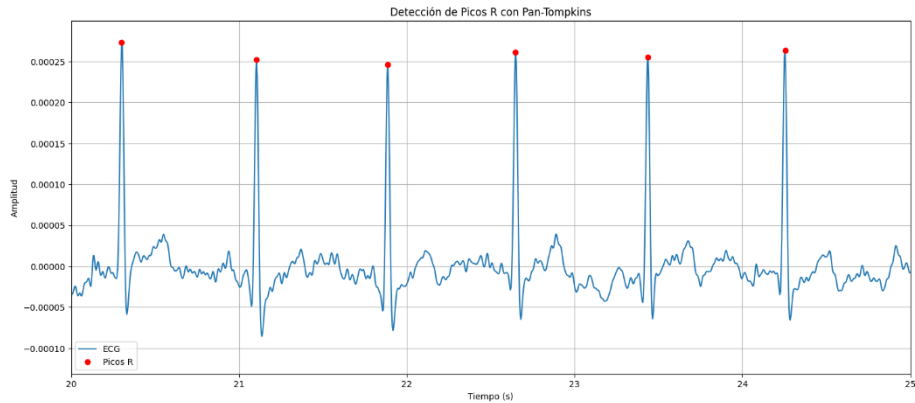


Figura 15: Detección de picos R con el algoritmo de Pan-Tompkins.

Posteriormente, se han calculado los intervalos RR a partir de estos picos R detectados, que representan el tiempo entre latidos consecutivos. Para cada pico R detectado, se calcularon cuatro intervalos:

- **pre_R**: Distancia temporal entre el pico R actual y el anterior.
- **post_R**: Distancia temporal entre el pico R actual y el siguiente.
- **local_R**: Promedio de los últimos 10 intervalos RR.
- **global_R**: Promedio de todos los intervalos RR de la ventana de 60 segundos de la señal.

Para verificar que se ha detectado un intervalo RR real y no es ruido se realizó un filtrado donde sólo se conservaron los latidos con intervalo RR dentro del rango de frecuencias entre 47,5 y 220 bpm.

2. Características morfológicas

Se implementaron cuatro grupos de características morfológicas con el objetivo de representar las variaciones en la forma del complejo QRS. Estas características aportan información en los diferentes dominios de tiempo y frecuencia, y sobre la forma de la señal.

1. **Características Wavelet**: Se ha utilizado la transformada Wavelet discreta para analizar las frecuencias del complejo QRS. La transformada de Wavelet es una técnica que descompone la señal de ECG en diferentes frecuencias, lo que permite analizarla a diferentes escalas. Esta herramienta es ampliamente utilizada en señales no estacionarias como ECG. Esta transformada devuelve coeficientes de aproximación de la parte general de la señal, y coeficientes de detalle a varios niveles; cuantos más niveles, se analizan detalles más finos de la señal [27]. En este caso se han extraído los coeficientes de aproximación de la señal.

2. **Características de estadísticos de orden superior (HOS):** Se han calculado 10 características basadas en los momentos estadísticos de orden superior, como asimetría (*skewness*) y curtosis (*kurtosis*) que permiten obtener información más compleja acerca de la forma del complejo QRS.
 - **Skewness:** mide qué tan simétrica es la distribución de amplitudes del complejo QRS.
 - **Kurtosis:** mide como de “picuda” es la señal. Mide el grado de concentración de los valores alrededor de la media.
3. **Características de funciones de base Hermite (HBF):** El objetivo también es obtener información acerca de la forma del complejo QRS. Se trata de representar esta forma como la combinación de los polinomios de Hermite. Cada polinomio tiene un grado distinto, y se ha ajustado la señal a los polinomios de grado 3, 4 y 5; en este caso. Cada uno de estos ajustes devuelve una serie de coeficientes, para un polinomio de grado n se obtienen $n+1$ coeficientes; por esta razón, en este caso se extraen 15 coeficientes para cada complejo QRS. Se puede ver el uso de estas características para el análisis del complejo QRS en [28].
4. **Características basadas en el algoritmo Pan-Tompkins:** En primer lugar, se identificaron eventos clave utilizando el algoritmo Pan-Tompkins [29].
 - Pico de la onda P: indica la despolarización auricular.
 - Pico de la onda Q: inicio de la despolarización ventricular.
 - Pico de la onda R: despolarización ventricular.
 - Pico de la onda S: final de la despolarización ventricular

A partir de esto, se han calculado características como:

- Amplitudes: **Ppeak, Qpeak, Rpeak, Speak.**
 - Intervalos temporales: **PRw** (duración de PR), **QSw** (duración de QS), **QRSw** (duración de QRS)
 - Pendientes: **QRSs**, la pendiente máxima del complejo QRS.
 - Diferencias de amplitud: **PQa, RQa y RSa**
3. **Características de la variabilidad de la frecuencia cardíaca (HRV):** HRV mide cómo varían los intervalos RR consecutivos en el tiempo. Para calcular estas características se utilizó la librería Neurokit2. Se pueden ver definidas en [30], [31], [32].
 1. **Características temporales:** Se calculan directamente a partir de los intervalos RR.
 - **HRV_MeanNN:** Se trata del promedio de todos los intervalos RR, e indica la frecuencia cardíaca media durante la ventana de 60 segundos analizada.
 - **HRV_SDNN:** Calcula la desviación estándar de los intervalos RR, midiendo

la variabilidad global.

- **HRV_RMSSD:** Es la raíz cuadrada de la media de las diferencias al cuadrado entre intervalos RR consecutivos. En este caso se mide la variabilidad a corto plazo y es un buen indicador de la actividad parasimpática.
- **HRV_pNN50:** Se trata del porcentaje de pares de intervalos RR consecutivos que se diferencian en más de 50 milisegundos de duración. Un valor alto indica una alta variabilidad y es señal de un sistema nervioso saludable.
- **HRV_pNN20:** Es igual que pNN50 pero con un umbral más pequeño, de 20 ms, detectando variaciones más pequeñas entre latidos.

2. Características frecuenciales: Estas se obtienen transformando los intervalos RR al dominio de frecuencia, y dan información de qué bandas de frecuencia predominan en la señal HRV.

- **HRV_ULF:** Corresponde con la banda de frecuencia ultra baja ($<0,003$ Hz). Representa cambios a muy largo plazo y están relacionadas con ritmos circadianos y metabólicos. En este estudio no es relevante ya que se requiere un registro muy largo para calcularla de forma correcta.
- **HRV_VLF:** Banda de frecuencia muy baja (0,003-0,04). Se trata de la banda de frecuencia predominante ya que corresponde con procesos de regulación.
- **HRV_LF:** Banda de frecuencia baja (0,04-0,15 Hz). Refleja la actividad del sistema nervioso simpático.
- **HRV_HF:** Banda de alta frecuencia (0,15-0,4 Hz). Refleja la actividad del sistema nervioso parasimpático.
- **HRV_LFHF:** Se trata del cociente entre LF y HF y representa el balance entre la actividad simpática y parasimpática.

3. Características no lineales: Permiten ver patrones complejos que no son capaces de obtener los métodos lineales.

- **HRV_SD1, SD2:** Representan la correlación entre intervalos RR consecutivos. SD1 mide la variabilidad a corto plazo (rápida), y la SD2 a largo plazo.
- **HRV_ApEn:** Es la entropía aproximada. Mide la irregularidad de la serie de intervalos RR.
- **HRV_SampEn:** Se trata de la entropía muestral. Es similar a la anterior, pero es más robusta al ruido y tamaño de muestra.
- **HRV_DFA_alpha1:** Mide la autocorrelación de la señal a corto plazo. Permite analizar como varía la señal en diferentes escalas de tiempo eliminando tendencias.

Se han extraído 200 características de cada señal de ECG en total mediante el análisis de HRV y la morfología del complejo QRS.

2.3.2. Características de la señal de EDA

El primer paso para la extracción de características de la señal EDA es la descomposición en sus componentes tónica y fásica (SCL y SCR respectivamente). Para ello se ha utilizado la biblioteca Neurokit2. Como se ha comentado anteriormente, esta descomposición es fundamental ya que la componente fásica es la que está directamente relacionada con las respuestas a estímulos y la tónica es indicativa del estado del sistema nervioso basal.

1. Características en el dominio del tiempo

▪ Características estadísticas básicas

1. Características de la señal cruda: de la señal sin procesar se extrajeron las siguientes características:

- **avg_raw:** Se trata del promedio de la conductancia en la ventana de 60 segundos. Un valor elevado puede indicar mayor actividad simpática.

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i \quad (1)$$

- **sd_raw:** Es la desviación estándar, y mide la dispersión de los valores alrededor de la media.

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2} \quad (2)$$

- **maxm_raw:** Se trata del máximo valor de la señal EDA en la ventana temporal.
 - **minm_raw:** Es el mínimo valor de conductancia, que representa el estado de menor activación simpática.
 - **kurtEDA:** Es la *kurtosis*, el cuarto momento estadístico.
 - **skewEDA:** *Skewness*, el tercer momento estadístico que mide su asimetría.
2. Características de la componente tónica: De la misma forma que se hizo con la señal cruda, se han extraído las mismas características estadísticas de la componente tónica de la señal.
 - **avg_tonic:** Media de la componente tónica.
 - **sd_tonic:** Desviación estándar de la componente tónica.
 - **maxm_tonic:** Valor máximo.
 - **minm_tonic:** Valor mínimo

3. Características de la componente fásica: Se han extraído las mismas características estadísticas en la componente fásica.
 - **avg_phasic**
 - **sd_phasic**
 - **maxm_phasic**
 - **minm_phasic**
- **Características de las derivadas:** Estas características aportan información acerca de la dinámica temporal y los cambios en la señal EDA.
 1. Primera derivada
 - **deriv_avg:** Se trata de la media de la derivada de la señal EDA y se utiliza para ver la tendencia promedio de cambio de la señal.
 - **deriv_max:** Es el máximo de la derivada e indica la mayor tasa de incremento.
 - **deriv_min:** El mínimo valor de la derivada, es decir, la mayor tasa de decremento.
 2. Segunda derivada
 - **deriv2_avg:** La aceleración promedio de cambio.
 - **deriv2_max:** La máxima aceleración positiva.
 - **deriv2_min:** La máxima aceleración negativa.
 - **Análisis de SCR:** En este caso se analiza la morfología de la señal SCR
 - **peakCount:** Es el número de picos de la señal, es decir, la cantidad de respuestas ante el estímulo detectadas.
 - **meanpeakAmplitude:** Amplitud promedio de picos.
 - **meanRiseTime:** Tiempo promedio desde el inicio hasta el pico.
 - **sumPeakAmplitude:** Suma de las amplitudes.
 - **sumRiseTime:** Suma de tiempos de subida (tiempo de inicio hasta el pico).
2. **Características en el dominio frecuencial:** Para analizar la señal en el dominio de la frecuencia se utiliza la Transformada de Fourier.
 - **Características estadísticas frecuenciales**
 - **SMA_f:** Cuantifica la energía total distribuida a lo largo de todas las frecuencias. Se calcula sumando los valores absolutos de los coeficientes obtenidos en la Transformada Rápida de Fourier (FFT).
 - **meanEDA_f:** Es la media de la señal en el dominio de frecuencia. Si es alta, puede indicar una mayor activación basal del SNA.
 - **stdEDA_f:** Es la desviación estándar frecuencial y representa la distribución de la potencia.
 - **signalRange_f:** Es la diferencia entre el máximo y el mínimo de la frecuencia.

- **KurtEDA_f:** *Kurtosis* frecuencial. Indicativo de la distribución de frecuencias.
 - **skewEDA_f:** *Skewness* frecuencial. Mide la asimetría en el dominio frecuencial.
- **Características de potencia por bandas:** Se dividió la señal en frecuencia en cinco bandas de igual longitud y en cada una se calculó:
- **SpectralPower_band1-5:** La potencia en cada banda específica.
 - **minSpectralPower:** La menor potencia entre las cinco bandas.
 - **maxSpectralPower:** La mayor potencia entre las cinco bandas.
 - **varSpectralPower:** La varianza de las potencias, que mide cuánto varían las potencias de las bandas entre sí.
3. **Características Mel-Frequency Cepstral Coefficients:** Para ello se utilizan los coeficientes MFCC. La escala Mel se trata de una forma de representar las frecuencias donde se comprimen las frecuencias altas y se expanden las bajas. Esta escala ha sido usada para el análisis de señales fisiológicas como EDA ya que puede ayudar a detectar patrones ocultos en la frecuencia [33].

Se han extraído 20 coeficientes y para cada uno de ellos se ha calculado:

- **meanMFCC[i]:** Valor promedio del coeficiente i.
- **stdMFCC[i]:** Variabilidad del coeficiente i.
- **medianMFCC[i]:** Valor central del coeficiente i.
- **kurtMFCC[i]:** Forma de la distribución del coeficiente i.
- **skewMFCC[i]:** Asimetría del coeficiente i.

En total se han extraído 140 características de la señal EDA.

2.3.3. Características de la señal de SKT

Para la señal SKT también se utilizó una ventana de los últimos 60 segundos de la señal para la extracción de características, con una frecuencia de muestreo de 256 Hz. Para cada ventana se calcularon las siguientes características:

- **Value_mean:** Se trata del valor medio de la temperatura en la ventana temporal.
- **Value_first_deriv_1min:** Es la primera derivada de la temperatura, que indica su tasa de cambio. Refleja si la señal de temperatura sube o baja, y a qué velocidad lo hace.

Se optó por extraer estas dos características ya que se analiza una ventana de duración limitada (60 segundos), lo que limita el número de métricas que se pueden calcular. Además, la señal SKT tiene baja variabilidad en ventanas cortas, por lo que se consideró suficiente extraer la media y la primera derivada como medidas más representativas.

Señal	Categoría	Características
ECG	Intervalos RR	pre_R, post_R, local_R y global_R
	Morfología QRS	Wavelet, HOS, Hermite, amplitudes, intervalos, pendientes
	HRV en dominio temporal	HRV_MeanNN, HRV_SDNN, HRV_RMSSD, HRV_pNN50, HRV_pNN20
	HRV en dominio frecuencial	HRV_ULF, HRV_VLF, HRV_LF, HRV_HF, HRV_LFHF
	HRV características no lineales	HRV_SD1, HRV_SD2, HRVApEn, HRV_SampEn, HRV_DFA_alpha1
EDA	Estadísticas básicas señal cruda	avg_raw, sd_raw, maxm_raw, minm_raw, kurtEDA, skewEDA
	Estadísticas básicas señal tónica	avg_tonic, sd_tonic, maxm_tonic, minm_tonic
	Estadísticas básicas señal fásica	avg_phasic, sd_phasic, maxm_phasic, minm_phasic
	Estadísticas de las derivadas	deriv_avg, deriv_max, deriv_min, deriv2_avg, deriv2_max, deriv2_min
	Análisis SCR	peakCount, meanpeakAmplitude, meanRiseTime, sumPeakAmplitude, sumRiseTime
	Dominio frecuencial	SMA_f, meanEDA_f, stdEDA_f, signalRange_f, KurtEDA_f, skewEDA_f, SpectralPower_band1-5, minSpectralPower, maxSpectralPower, varSpectralPower
	MFCC	meanMFCC[i], stdMFCC[i], medianMFCC[i], kurtMFCC[i], skewMFCC[i] (i=20)
SKT	Estadísticas básicas	value_mean
	Derivadas	Value_first_deriv_1min

Tabla 5: Características extraídas de ECG, EDA y SKT.

2.4. SELECCIÓN DE CARACTERÍSTICAS

El proceso de extracción de características dio como resultado un conjunto de 342 datos. No es apropiado utilizar un conjunto de datos con una dimensionalidad tan alta como entrada de un modelo predictivo, ya que incrementa el riesgo de sobreajuste y dificulta la interpretación de los resultados. Por esta razón, se hizo una selección de las características utilizando el método de *Recursive Feature Elimination with Cross-Validation* (RFECV), que ha sido eficaz en estudios como [34].

RFECV combina dos enfoques: la eliminación de características menos útiles de forma recursiva, y por otro, la validación cruzada, es decir, evaluar de forma iterativa el modelo con diferentes partes de los datos. De esta forma se consigue una selección de características que evita el sobreajuste y que tiene buen rendimiento con datos nuevos [34]. Este algoritmo utiliza un modelo de aprendizaje automático de clasificación para puntuar cada característica y eliminar de forma iterativa las que no mejoren la precisión en la clasificación. El proceso de búsqueda de las mejores características comienza con el conjunto completo de datos, y se eliminan progresivamente las que no contribuyen a la precisión de la clasificación [35]. En este trabajo se ha utilizado un árbol de decisión como modelo de clasificación y con validación cruzada de 5, es decir, se dividió el conjunto en 5 partes iguales donde en cada iteración se utilizaron 4 para entrenar el modelo y 1 para evaluarlo.

En este trabajo se ha implementado una estrategia de modelado donde se desarrollan tres modelos predictivos independientes: uno para la predicción de la valencia, otro para la excitación, y el último donde se predice el cuadrante emocional basado en el modelo circunflejo. Para cada modelo predictivo se aplicó RFECV de forma independiente, de esta forma cada modelo identifica el subconjunto de características más relevantes para su tarea específica.

1. Modelo de clasificación binaria de la valencia

Para el modelo de clasificación binaria de la valencia fueron seleccionadas 16 características: 14 procedentes de la señal de ECG y 2 de la señal EDA.

Señal	Categoría	Características
ECG	Intervalos RR	pre_R, post_R
	Morfología QRS	QSw, PQa, RQa QSa, QRSw, QRSs, HBF4
	HRV en el dominio frecuencial	HRV_LF, HRV_HF, HRV_LFHF
	HRV características no lineales	HRV_ApEn, HRV_DFA_alpha1

EDA	Estadísticas básicas de la señal física	maxm_phasic
	Estadísticas de las derivadas	deriv2_avg

Tabla 6: Características seleccionadas para el modelo de clasificación binaria de la valencia.

2. Modelo de clasificación binaria de la excitación

Para este modelo fueron seleccionadas 29 características. 16 procedentes de la señal de ECG y 13 de EDA.

Señal	Categoría	Características
ECG	Intervalos RR	pre_R, post_R, global_R
	Morfología QRS	QSw, PQa, QRSw, QRSs, HBF4
	HRV en el dominio frecuencia	HRV_LF, HRV_HF, HRV_LFHF
	HRV en el dominio temporal	HRV_MeanNN, HRV_pNN50, HRV_SDNN
	HRV características no lineales	HRV_ApEn, HRV_DFA_alpha1
EDA	Estadísticas básicas señal cruda	avg_raw, sd_raw
	Estadísticas básicas señal física	avg_phasic, sd_phasic, maxm_phasic
	Estadísticas de las derivadas	deriv_avg, deriv2_avg
	Análisis SCR	peakCount, meanpeakAmplitude, meanRiseTime, sumPeakAmplitude
	Dominio frecuencial	meanEDA_f, stdEDA_f

Tabla 7: Características seleccionadas para el modelo de clasificación binaria de la excitación.

3. Modelo para la clasificación multiclase de los cuatro cuadrantes emocionales

Por último, para este tercer modelo fueron seleccionadas 8 características, 4 procedentes de ECG y 4 de EDA.

Señal	Categoría	Características
ECG	Intervalos RR	post_R, global_R
	HRV en dominio temporal	HRV_MeanNN
	HRV en dominio frecuencial	HRV_LF
EDA	Estadísticas básicas señal cruda	kurtEDA
	Estadísticas de las derivadas	deriv_avg, deriv2_avg, deriv2_max

Tabla 8: Características seleccionadas para el modelo de clasificación multiclase.

2.5. MODELOS DE CLASIFICACIÓN

Los modelos de clasificación son un tipo de modelo de aprendizaje automático utilizado para predecir la categoría o clase a la que pertenece un dato. Toma una entrada y asigna esa entrada a varias clases predefinidas. La clasificación consiste en predecir etiquetas para patrones que no están etiquetados basándose en observaciones. De este modo, se utiliza un conjunto de datos que será utilizado para extraer patrones que tiene etiquetas conocidas, llamado conjunto de entrenamiento. El objetivo de los modelos de clasificación es obtener una correcta predicción de la etiqueta de clase para patrones desconocidos, generalmente evaluados con un conjunto de validación [36].

Es común en la literatura sobre la detección emocional abordar la predicción automática como un problema de clasificación binaria, simplificando así el análisis y facilitando el desarrollo de los modelos [18]. Por esta razón, como ya se ha descrito anteriormente, en este trabajo se han desarrollado tres modelos de clasificación emocional: 2 modelos de clasificación binaria de valencia y excitación, y un modelo de clasificación multiclase de los 4 cuadrantes emocionales del modelo circunflejo, para evaluar si tiene un mejor rendimiento que los modelos binarios individuales. Para cada uno de estos tres problemas de clasificación, se han evaluado dos tipos de modelos: Random Forest (RF) y Red Neuronal Artificial (RNA). Se han elegido estos dos modelos ya que son los que mejores resultados ofrecen según la bibliografía [23]. Aunque SVM es el algoritmo más empleado en la detección emocional utilizando señales fisiológicas, en muchos

casos no alcanza el mismo nivel de exactitud que los modelos mencionados [19], [20], [21], [23]. Antes de implementar los modelos de clasificación se hizo un preprocesamiento de los datos, donde se estandarizaron las características y se hizo un balanceo de clases. La estandarización se hizo utilizando *StandardScaler*, que hace que los datos se transformen para que tengan media 0 y una desviación estándar 1. Esto garantiza que los datos tengan una escala equilibrada. Por otro lado, el balanceo de clases consiste en utilizar pesos balanceados en el modelo, penalizando más los errores cometidos con la clase minoritaria. Esto es muy útil en este trabajo ya que la base de datos MAHNOB-HCI tiene clases desbalanceadas.

2.5.1. Fundamentos generales

2.5.1.1. Redes Neuronales Artificiales

Las RNA son modelos computacionales que pertenecen al aprendizaje automático, y su nombre hace referencia a la similitud que tienen con las estructuras funcionales del sistema nervioso humano. Están formadas por unidades básicas llamadas nodos o neuronas, y están organizadas en capas interconectadas entre sí [37].

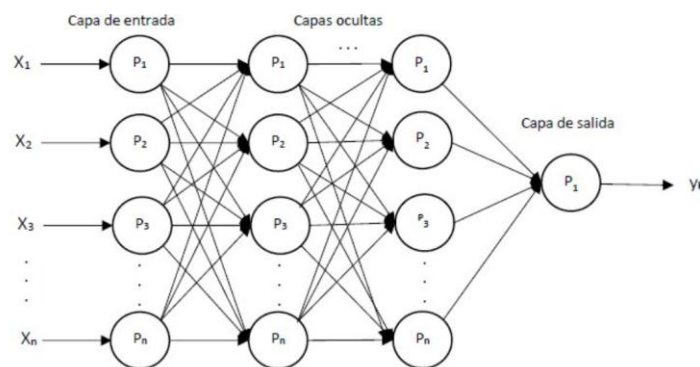


Figura 16: Red Neuronal Artificial [38].

Estas redes están compuestas por una capa de entrada, una o varias capas intermedias u ocultas, que serán las responsables de aprender los diferentes patrones ocultos en los datos, y una capa de salida que devuelve el resultado de la predicción.

Cada neurona artificial funciona de forma similar a una neurona biológica, de modo que se puede activar o no utilizando valores de 1 si está activa o 0 si está inactiva, en función de la entrada que reciba [37].

La forma en la que la red aprende puede ser mediante aprendizaje supervisado o no supervisado. El aprendizaje supervisado se basa en el entrenamiento de la red a partir de datos que están previamente etiquetados, por lo que su salida es conocida. Se ajustan los pesos de manera que se obtenga el menor error posible entre la salida obtenida y la esperada. Por otro lado, el aprendizaje no supervisado consiste en entrenar la red con datos que no han sido etiquetados previamente, de modo que la red trata de encontrar patrones y relaciones en los datos para obtener la salida correcta [39].

2.5.1.2. Random Forest

Para entender el algoritmo de RF, antes hay que explicar los modelos basados en árboles de decisión. Estos se basan en la partición del conjunto de datos de forma recursiva en dos conjuntos según un criterio definido, hasta alcanzar una condición de parada [40].

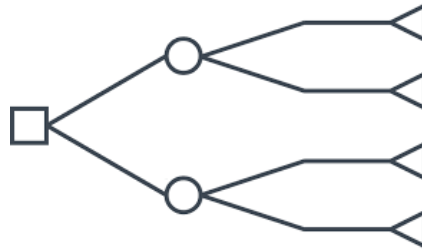


Figura 17: Árbol de decisión [41].

Se trata de un modelo de aprendizaje automático que, en función de cómo se establezcan los criterios de partición y parada pueden llevar a cabo tareas de clasificación o regresión. En ambos tipos de problema es importante elegir qué variable usar para la división de los conjuntos. Para elegir la mejor variable se usan criterios de división que tratan de resolver un problema de optimización, donde se busca obtener la mejor forma posible de dividir los datos. Uno de los criterios de división más comunes en clasificación es mediante la entropía de Shannon, que mide cómo de mezclados están los datos. Si la entropía es alta, significa que el conjunto tiene elementos de muchas clases diferentes, y si es baja es que pertenecen todos a una sola clase. Se calcula la entropía en cada nodo con la siguiente ecuación:

$$E = - \sum_{i=1}^c p_i \times \log(p_i) \quad (3)$$

c es el número de clases únicas y p_i es la probabilidad a priori de cada clase. El valor de la entropía debe ser menor en cada división, ya que el objetivo del árbol es separar bien los datos de forma que estén lo más ordenados posible [40].

Los árboles de decisión son muy propensos al sobreajuste, es decir, aprende demasiado bien los datos del conjunto de entrenamiento de forma que memoriza las respuestas exactas y no es capaz de generalizar de forma correcta. Una forma de evitar este problema es construir muchos árboles individuales utilizando sólo un subconjunto de los datos. Se hace un promedio de las predicciones de muchos árboles de decisión individuales en caso de regresión, o se elige la clase más predicha en caso de clasificación. Cada árbol individual se entrena con una muestra aleatoria de los datos con reemplazo, a esto se le llama Bootstrap [40].

2.5.2. Modelo de clasificación binaria de la valencia y excitación

En primer lugar, se desarrollaron dos modelos de clasificación binaria de valencia y excitación. Las etiquetas de estas dimensiones son valores discretos del 1 al 9 en la base de datos MAHNOB-HCI, por lo que se establecieron valores de valencia negativa y excitación baja de 1 a 4, y valencia positiva y excitación alta de 5 a 9. De las 463 muestras utilizadas como entrada de los modelos de clasificación, 214 tienen etiqueta de valencia positiva y 249 de valencia negativa. En cuanto a la excitación, 226 tienen excitación baja y 237 alta.

Se probaron los mismos modelos de RF y RNA para cada problema de clasificación ya que, tanto la valencia como la excitación son variables discretas en el mismo rango, y utilizar el mismo modelo facilita la comparación directa del desempeño de las dos dimensiones.

2.5.2.1. Random Forest

Para el modelo de clasificación RF se ajustaron los siguientes hiperparámetros de forma empírica:

- 500 árboles de decisión. Cuantos más árboles, mejor precisión y estabilidad, pero esto ocurre hasta un punto donde se incrementa excesivamente el tiempo de entrenamiento con una mejora marginal de la precisión [42].
- Profundidad máxima de 20 niveles. Esto se hace para prevenir el sobreajuste, ya que si el árbol crece mucho aprende patrones de los datos y no generaliza de forma correcta.
- Mínimo de 4 muestras para definir un nodo. De esta forma se evitan ramas inútiles que no son representativas.
- Al menos 2 muestras por hoja o nodo final: Cada hoja representa la predicción final. Esto hace que se evite el sobreajuste ya que no permite que haya hojas con una sola muestra de forma que memorice esa muestra.

2.5.2.2. Redes Neuronales Artificiales

Se diseñó una red neuronal con tres capas densas ocultas con 256, 128 y 64 neuronas respectivamente, seguidas por la capa de salida de 1 neurona que devuelve el valor de la clasificación.

Para las capas ocultas se utilizó la función de activación *ReLU* y para la capa final la función *sigmoid*, que devuelve una salida entre 0 y 1, es decir, una probabilidad de que la muestra pertenezca a una clase.

Para evitar el sobreajuste se aplicó regularización mediante capas de *dropout* con tasas del 40%, 30% y 20% en cada capa oculta respectivamente. Esto hace que durante el entrenamiento se desactiven aleatoriamente este porcentaje de neuronas para que el modelo no cree dependencias a nodos específicos.

2.5.3. Modelo de clasificación multiclase de los cuatro cuadrantes emocionales

Para comparar el rendimiento de varios problemas de clasificación, se implementó un modelo de

clasificación de emociones en los cuatro cuadrantes del modelo circunflejo:

- Q1: Alta valencia y alto arousal.
- Q2: Baja valencia y alto arousal.
- Q3: Baja valencia y bajo arousal.
- Q4: Alta valencia y bajo arousal.

A continuación, en la Figura 18 se muestra un histograma de la frecuencia de aparición de las etiquetas felEmo, y en la Tabla 9 qué emociones corresponden a cada cuadrante. La Figura 19 corresponde con el histograma de la frecuencia de las emociones agrupadas en los cuatro cuadrantes (Q1-Q4).

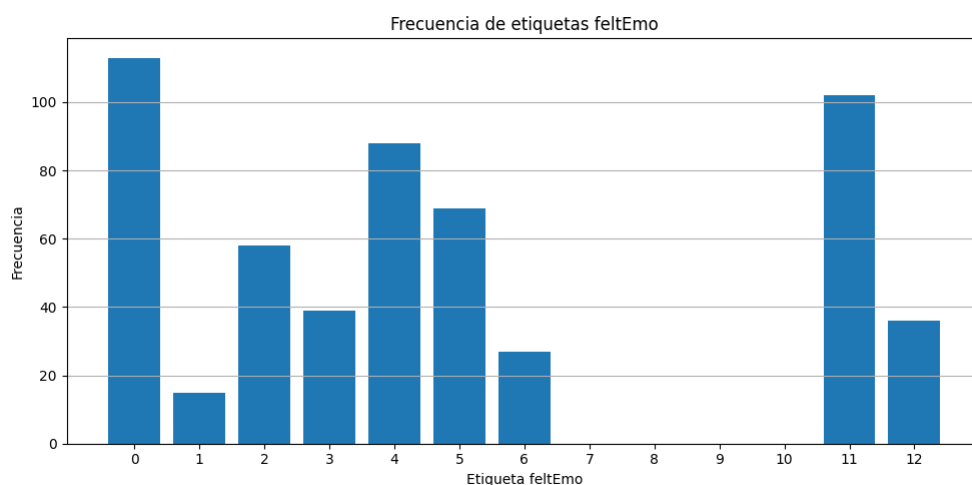


Figura 18: Histograma de las etiquetas felEmo en la base de datos MAHNOB-HCI.

Etiqueta	Emoción	Cuadrante
0	Neutral	Q4
1	Ira	Q2
2	Disgusto	Q2
3	Miedo	Q2
4	Alegría, Felicidad	Q1
5	Tristeza	Q3
6	Sorpresa	Q1
11	Diversión	Q1
12	Ansiedad	Q2

Tabla 9: Correspondencia de etiquetas, emociones y cuadrantes afectivos.

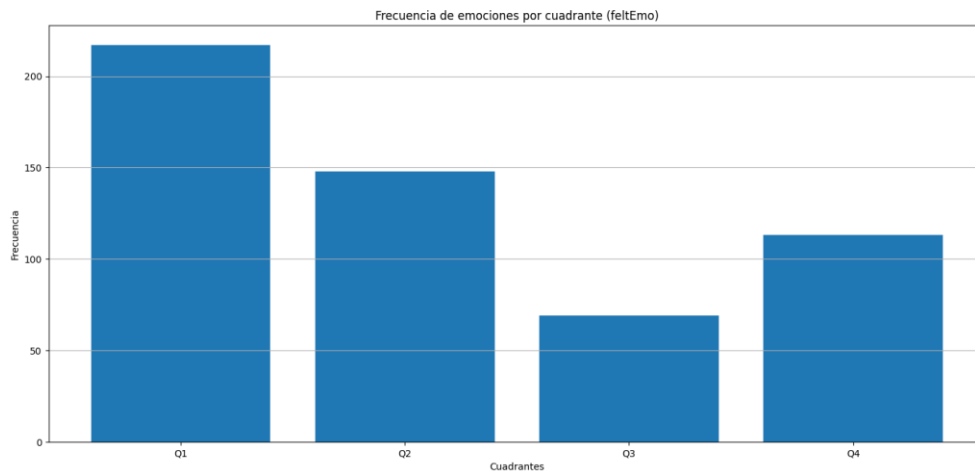


Figura 19: Histograma de los cuatro cuadrantes emocionales según las etiquetas de *feltEmo* en la base de datos MAHNOB-HCI

Para este problema de clasificación se tuvieron que adaptar los modelos para una clasificación multiclase.

2.5.4. Random Forest

En el modelo RF la configuración es idéntica al caso binario ya que este algoritmo ya está preparado para clasificar en más de dos categorías.

2.5.5. Redes Neuronales Artificiales

En este caso, la RNA se ha modificado para clasificar 4 categorías diferentes. En primer lugar, se ha utilizado una capa de salida con 4 neuronas y con una función de activación *softmax*, que permite que la salida sea una probabilidad para cada clase, de forma que todas sumen 1. De esta forma, la clase que tenga mayor probabilidad será la predicción final.

Además, como la red requiere que la salida tenga valores numéricos, se ha realizado un proceso de codificación de etiquetas. Se ha utilizado *LabelEncoder* para transformar las etiquetas categóricas Q1, Q2, Q3 y Q4 en los valores numéricos: 0, 1, 2 y 3.

2.6. ENTRENAMIENTO Y VALIDACIÓN

En el entrenamiento de los modelos de clasificación de RNA, se actualizan los parámetros de modo que se minimice una función de pérdida. Esta función de pérdida fue *Binary Crossentropy* para los modelos de clasificación primaria y *Sparse Categorical Crossentropy* para la clasificación multiclase.

Cross-entropy, o entropía cruzada, se trata de una función de pérdida que evalúa si la predicción del modelo es correcta o no. Si la probabilidad predicha por el modelo de clasificación es baja

para la etiqueta real, la pérdida será muy alta, y el modelo ajustará los pesos para reducir esta función de pérdida [43]. *Binary Crossentropy* se utiliza cuando el objetivo de la predicción es 1 o 0 y viene de una función *sigmoid*. Por otro lado, *Sparse Categorical Crossentropy* se utiliza en problemas multiclase utilizando la función *softmax* como salida de la red. Esta función mide la diferencia entre la probabilidad que el modelo ha dado a la clase real y el valor ideal, que es 1 para la clase correcta y 0 para las demás.

En cambio, los modelos RF no se entrenan minimizando funciones de pérdida, sino que, durante el entrenamiento, el modelo se va construyendo de forma que se minimice la entropía en cada nodo.

En cuanto al proceso de validación, que sirve para evaluar la capacidad de generalización del modelo, se utilizó el método de validación cruzada estratificada (*Stratified k-Fold Cross Validation*) con 5 divisiones, es decir, se divide el conjunto de datos en 5 subconjuntos de forma que en cada iteración 4 se utilicen para el entrenamiento y 1 para la validación. Este proceso se repite 5 veces utilizando cada vez un subconjunto distinto para la validación. Que la validación sea estratificada significa que en cada subconjunto hay la misma proporción de clases que en el conjunto de datos original, muy útil en problemas con clases desbalanceadas como este.

2.7. MEDIDAS DE EVALUACIÓN

La evaluación del modelo de predicción es la última etapa del proceso de reconocimiento de emociones. Para ello se utilizan métricas que permiten analizar el rendimiento del modelo como: precision, recall, F1-score y accuracy. Además, es común usar herramientas como la matriz de confusión.

- **Matriz de confusión:** Se trata de una herramienta estadística utilizada para interpretar los resultados de un clasificador[23]. En ella se describen 4 valores:
 - Verdaderos positivos (VP): número de predicciones correctas de la clase positiva.
 - Verdaderos negativos (VN): número de predicciones correctas de la clase negativa.
 - Falsos positivos (FP): número de predicciones incorrectas de la clase positiva.
 - Falsos negativos (FN): número de predicciones incorrectas de la clase negativa.
- **Precision:** Se trata de una métrica que indica el número de predicciones positivas correctas realizadas [23]. Permite saber cómo de fiable es un modelo cuando predice una clase positiva.

$$Precision = \frac{VP}{VP + FP} \quad (4)$$

- **Recall:** Esta métrica también es conocida como sensibilidad o tasa de verdaderos positivos. Cuantifica el número de casos realmente positivos que puede clasificar el modelo, es decir, la proporción de los positivos reales que ha sido detectada correctamente

por el modelo [23].

$$Recall = \frac{VP}{VP + FN} \quad (5)$$

- **F1-score:** Se trata de la media armónica entre *precision* y *recall*. Es una métrica que combina estos dos valores y es especialmente útil cuando las clases están desbalanceadas y se quiere evaluar el balance general del modelo [23].

$$F1 - score = 2 \times \frac{precision \times recall}{precision + recall} \quad (6)$$

- **Accuracy:** Indica la frecuencia con la que el modelo clasifica un conjunto de datos de forma correcta. Es el número de predicciones correctamente clasificadas entre todas las predicciones [23].

$$Accuracy = \frac{VP + VN}{VP + VN + FP + FN} \quad (7)$$

2.8. ADQUISICIÓN DE SEÑALES FISIOLÓGICAS

Una vez se han desarrollado y validado los modelos predictivos utilizando la base de datos MAHNOB-HCI, se realizó una prueba con señales fisiológicas propias para poner estos modelos en práctica.

Para registrar las señales fisiológicas se utilizan sensores específicos que permiten adquirir de forma precisa, no invasiva y objetiva las señales de interés. Estos sensores actúan como interfaz entre el cuerpo humano y los sistemas eléctricos, midiendo los parámetros relacionados con el estado de la persona, como el ECG, la EDA, o la SKT.

Los biosensores que se han utilizado en este trabajo permiten la obtención continua de datos fisiológicos con una alta resolución temporal, lo que es especialmente necesario en estudios de detección de emociones en tiempo real para respaldar de forma adecuada la toma de decisiones [44].

Es importante aclarar tres puntos necesarios para el uso correcto de los biosensores. En primer lugar, es necesario utilizar una frecuencia de muestreo adecuada para representar de forma correcta la señal. En segundo lugar, hay que asegurarse de colocar de forma correcta los sensores, ya que su mala colocación puede dar lugar a artefactos indeseados. Por último, es importante saber que las señales fisiológicas sin procesar pueden presentar ruido debido a la variabilidad fisiológica natural, que hace que las señales presenten fluctuaciones naturales como la respiración, latidos del corazón o movimientos musculares. Este ruido debe ser eliminado con el filtro adecuado [44].

2.8.1. Señal ECG

El sensor empleado en este trabajo para medir la señal eléctrica del corazón, representada mediante un ECG, es el SparkFun Single Lead Heart Rate Monitor – AD8232.

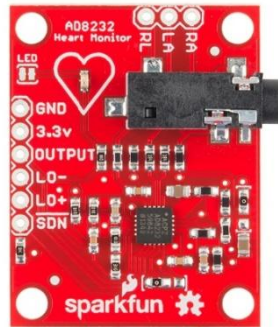


Figura 20: Sensor SparkFun AD8232 [45].

La forma en la que se obtiene la señal ECG es mediante un microcontrolador Arduino que se conecta al sensor, de manera que se obtiene un monitor cardíaco capaz de medir la actividad eléctrica del corazón en tiempo real.

Este sensor tiene nueve conexiones denominadas “pines”, de los que se utilizan cinco para conectarlos a la placa de Arduino: GND, 3.3V, OUTPUT, LO- y LO+ [45].

PIN	ARDUINO
GND	GND
3.3 V	3.3 V
OUTPUT	A0
LO-	11
LO+	10

Tabla 10: Conexiones entre sensor SparkFun AD8232 y Arduino.

Para adquirir la señal de ECG de la persona, se utilizan tres electrodos y parches desechables para electrodos.



Figura 21: Electrodos para SparkFun AD8232 [46].

Los cables están codificados por colores, y se pueden colocar en los antebrazos y la pierna derecha, o bien en ambos lados del pecho y en la parte baja y derecha del abdomen, como se ve en la Figura 22. En este trabajo hemos utilizado la primera opción. Según la página web del fabricante, el código de colores es el siguiente [45]:

COLOR	UBICACIÓN
Negro	Brazo derecho
Azul	Brazo izquierdo
Rojo	Pierna derecha

Tabla 11: Asignación de colores de electrodos de SparkFun AD8232 y su ubicación según la web del fabricante.

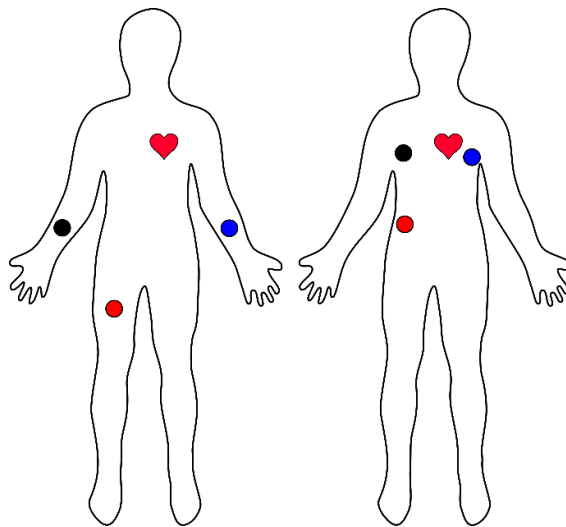


Figura 22: Colocación de electrodos según la página web del fabricante [45].

En la práctica se ha visto que con esa colocación de los electrodos no se obtenían buenos resultados, y como se observa también en [47] la colocación correcta es la siguiente:

COLOR	UBICACIÓN
Negro	Brazo derecho
Rojo	Brazo izquierdo
Azul	Pierna derecha

Tabla 12: Asignación de colores de electrodos de SparkFun AD8232 y su ubicación correcta.

De esta forma, se puede observar la señal ECG mediante Arduino:

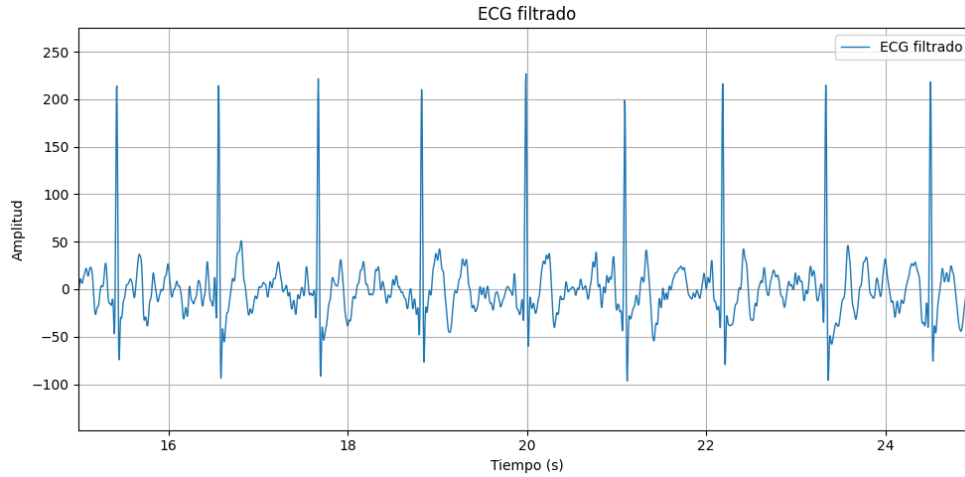


Figura 23: Señal ECG con SparkFun AD8232.

2.8.2. Señal EDA

Para medir la señal EDA se ha utilizado el sensor Grove-GSR V1.2 diseñado por Seeed Studio. Consta de dos electrodos que se colocan en los dedos y detectan la variación en la conductancia de la piel relacionadas con la sudoración, la cual varía ante estímulos emocionales por acción del sistema nervioso simpático.



Figura 24: Sensor Grove-GSR de Seeed Studio [48].

Es importante tener en cuenta que este sensor mide la resistencia de la piel, no su conductancia. La relación entre ambos valores es la siguiente:

$$\text{Conductancia } (G) = \frac{1}{\text{Resistencia } (R)} \quad (8)$$

Grove GSR V1.2 mide la resistencia eléctrica entre dos electrodos colocados en los dedos de la mano con una débil corriente eléctrica circulando entre ellos. Tras la estimulación simpática, se produce un pequeño aumento en la sudoración de la piel, lo que hace que la resistencia disminuya, ya que el sudor tiene agua y electrolitos que hacen que la conductividad de la piel aumente. De esta forma, un aumento de la actividad simpática genera una disminución en la señal recibida por el sensor [13].

La diferencia en la resistencia se convierte en un voltaje analógico y es leído por un microcontrolador, en este caso, Arduino. Dado que las variaciones en la resistencia de la piel son muy pequeñas, el sensor GSR V1.2 amplifica esta señal para que el microcontrolador lo pueda interpretar de forma correcta [48].

Las conexiones de los pines del sensor y el Arduino son las siguientes:

Cable sensor GSR V1.2	ARDUINO
Negro	GND
Rojo	5 V
Amarillo	A0

Tabla 13: Conexiones entre sensor Grove-GSR V1.2 y Arduino

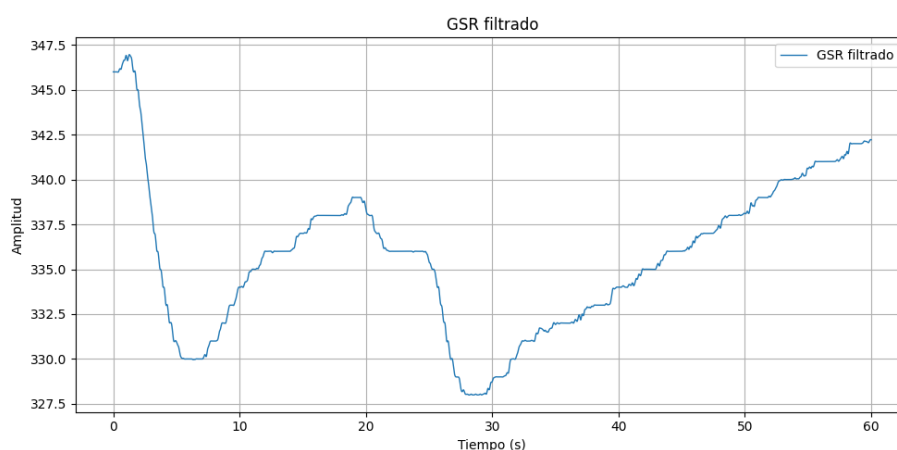


Figura 25: Señal EDA obtenido con el sensor Grove-GSR V1.2

2.9. VALIDACIÓN CON SEÑALES FISIOLÓGICAS PROPIAS

Para comprobar si los modelos realizados son aplicables en un entorno real, se recogieron las señales utilizando los sensores descritos anteriormente y se procesaron siguiendo los mismos pasos que para los datos de MAHNOB-HCI, es decir, preprocesamiento, y extracción y selección de características. Después, estas señales fueron introducidas a los modelos que mejores resultados proporcionaron. Como resultado, los modelos de clasificación binaria predijeron un estado de alta excitación y baja valencia, lo que según el modelo circunflejo está relacionado con estrés, miedo o ira.

Esta prueba no es útil para validar la predicción de forma objetiva ya que en el momento de la prueba no se sometió al participante a ningún estímulo emocional. Por el contrario, sí que sirve para demostrar que es posible utilizar los modelos desarrollados con señales registradas fuera de un entorno controlado, lo que acerca a este trabajo a la práctica en escenarios reales.

3. RESULTADOS

En esta sección se presentan los resultados obtenidos para los tres problemas de clasificación: clasificación binaria de la valencia y excitación, y clasificación multiclase de los 4 cuadrantes emocionales del modelo circunflejo. De cada uno de los tres problemas se muestra el resultado para el modelo de RF y RNA, con sus métricas de rendimiento y matriz de confusión normalizada.

3.1. Clasificación binaria de la valencia

3.1.1. Random Forest

En la Tabla 14 se muestran los resultados obtenidos para la clasificación binaria de la valencia con el modelo RF. Se muestran las métricas de *precision*, *recall* y *F1-score* por cada clase, además de la exactitud global y los promedios macro y ponderado.

Clase	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Soporte
0	0.61	0.52	0.56	214
1	0.63	0.71	0.67	249
<i>Accuracy global</i>			0.62	463
<i>Macro avg</i>	0.62	0.62	0.62	463
<i>Weighted avg</i>	0.62	0.62	0.62	463

Tabla 14: Resultado del rendimiento del modelo RF de clasificación binaria de la valencia.

La matriz de confusión normalizada para RF se puede ver en la Figura 26. En ella se observa la distribución de aciertos y errores durante la clasificación.

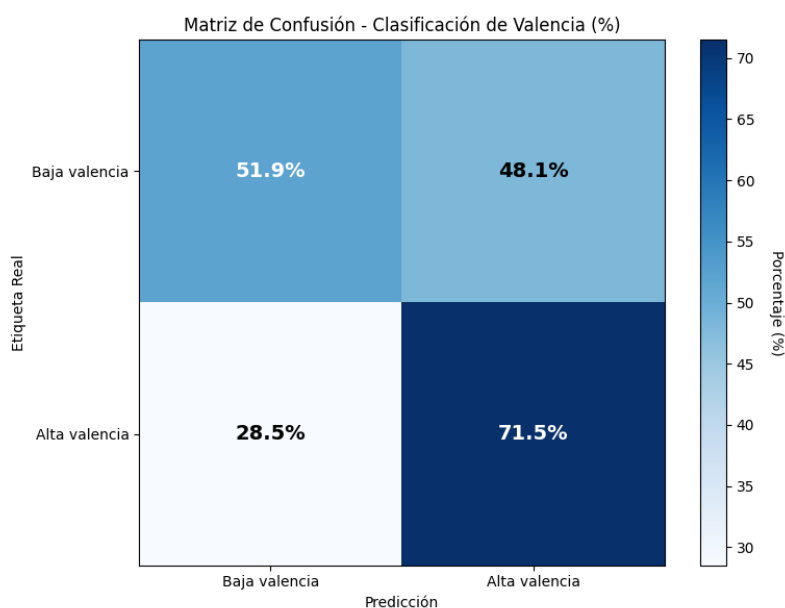


Figura 26: Matriz de confusión normalizada para la clasificación binaria de la valencia con RF.

3.1.2. Red Neuronal Artificial

Los resultados del modelo RNA para la clasificación binaria de la valencia se pueden ver en la Tabla 15.

Clase	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Soporte
0	0.612	0.486	0.542	214
1	0.625	0.735	0.675	249
<i>Accuracy global</i>			0.620	463
<i>Macro avg</i>	0.618	0.610	0.608	463
<i>Weighted avg</i>	0.619	0.620	0.614	463

Tabla 15: Resultado del rendimiento del modelo RNA de clasificación binaria de la valencia.

La Figura 27 muestra la matriz de confusión normalizada correspondiente al modelo RNA para la clasificación de la valencia.

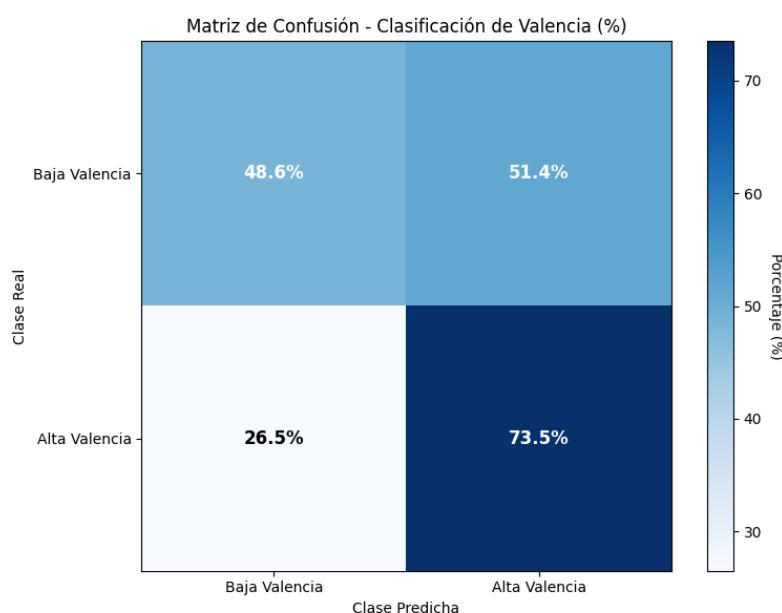


Figura 27: Matriz de confusión normalizada para la clasificación binaria de la valencia con RNA.

3.2. Clasificación binaria de la excitación

3.2.1. Random Forest

La Tabla 16 muestra el desempeño del modelo RF para la clasificación binaria de la excitación.

Clase	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Soporte
0	0.63	0.63	0.63	237
1	0.61	0.61	0.61	226
<i>Accuracy global</i>			0.62	463
<i>Macro avg</i>	0.62	0.62	0.62	463
<i>Weighted avg</i>	0.62	0.62	0.62	463

Tabla 16: Resultado del rendimiento del modelo RF de clasificación binaria de la excitación.

La matriz de confusión normalizada para este mismo problema de clasificación con RF se observa en la Figura 28.

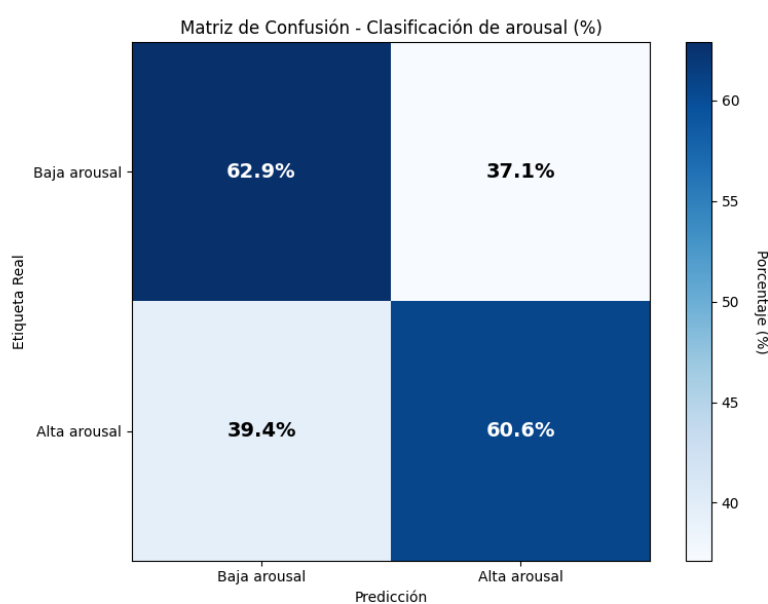


Figura 28: Matriz de confusión normalizada para la clasificación binaria de la excitación con RF.

3.2.2. Red Neuronal Artificial

La Tabla 17 muestra los resultados de RNA aplicada a la clasificación binaria de la excitación.

Clase	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Soporte
0	0.535	0.734	0.619	237
1	0.543	0.332	0.412	226
<i>Accuracy global</i>			0.538	463
<i>Macro avg</i>	0.539	0.533	0.516	463
<i>Weighted avg</i>	0.539	0.538	0.518	463

Tabla 17: Resultado del rendimiento del modelo RNA de clasificación binaria de la excitación.

En la Figura 29 se observa la matriz de confusión normalizada del modelo RNA para la clasificación de la excitación.

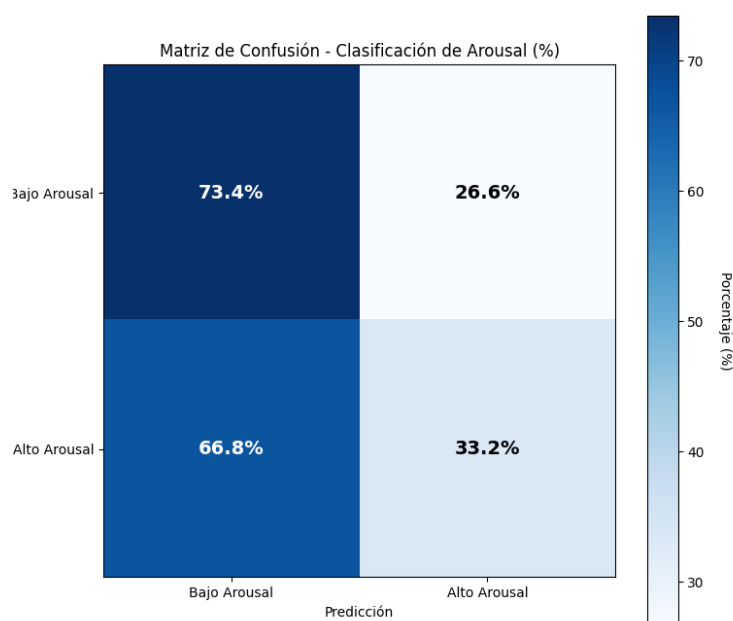


Figura 29: Matriz de confusión normalizada para la clasificación binaria de la excitación con RNA.

3.3. Clasificación multiclase de los 4 cuadrantes emocionales

3.3.1. Random Forest

Los resultados de la clasificación multiclase de los 4 cuadrantes emocionales con el modelo RF se recogen en la Tabla 18.

Clase	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Soporte
Q1	0.52	0.61	0.56	179
Q2	0.57	0.61	0.59	114
Q3	0.53	0.36	0.43	69
Q4	0.38	0.33	0.35	101
<i>Accuracy global</i>			0.51	463
<i>Macro avg</i>	0.50	0.48	0.48	463
<i>Weighted avg</i>	0.51	0.51	0.50	463

Tabla 18: Resultado del rendimiento del modelo RF de clasificación multiclase de los 4 cuadrantes emocionales.

La Figura 30 muestra la matriz de confusión normalizada obtenida con RF para el mismo problema de clasificación.

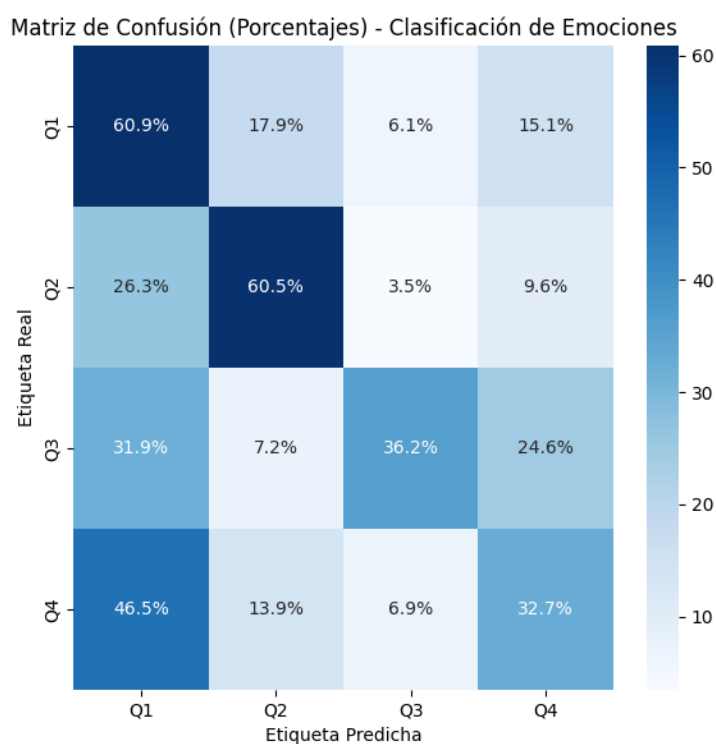


Figura 30: Matriz de confusión normalizada para la clasificación multiclase de los 4 cuadrantes emocionales con RF.

3.3.2. Red Neuronal Artificial

En la Tabla 19 se observan las métricas del desempeño del modelo RNA para la clasificación multiclase de los 4 cuadrantes emocionales.

Clase	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>	Soporte
Q1	0.411	0.374	0.392	179
Q2	0.450	0.439	0.444	114
Q3	0.256	0.435	0.323	69
Q4	0.292	0.208	0.243	101
<i>Accuracy global</i>			0.363	463
<i>Macro avg</i>	0.352	0.364	0.350	463
<i>Weighted avg</i>	0.372	0.363	0.362	463

Tabla 19: Resultado del rendimiento del modelo RNA de clasificación multiclase de los 4 cuadrantes emocionales.

Por último, la matriz de confusión normalizada del modelo RNA para la clasificación multiclase se muestra en la Figura 31.

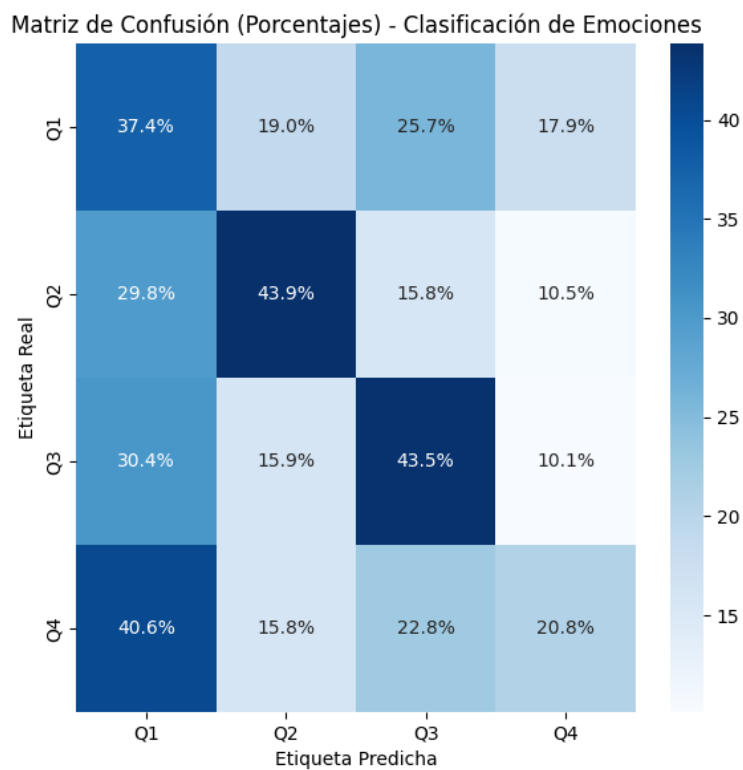


Figura 31: Matriz de confusión normalizada para la clasificación multiclase de los 4 cuadrantes emocionales con RNA.

4. DISCUSIÓN

4.1. Análisis de los resultados

En el apartado anterior se han mostrado los resultados obtenidos para cada modelo en cada uno de los tres problemas de clasificación. El mejor resultado ha sido el obtenido en los modelos de clasificación binaria de valencia y excitación utilizando el algoritmo RF, con un *accuracy* del 62%. A continuación, se discuten los resultados obtenidos para cada modelo.

4.1.1. Clasificación binaria de la valencia

Para la clasificación binaria de la valencia, los resultados muestran que tanto RF como la RNA alcanzan un rendimiento similar, con un *accuracy* global del 62% en ambos modelos.

El modelo RF consigue un balance más uniforme entre clases, ya que tiene valores similares de *precision* y *recall* en ambas clases (0,61 y 0,63 para *precision* y 0,52 y 0,71 para *recall*). Por el contrario, RNA presenta un sesgo evidente hacia la clase 1, es decir, valencia alta. Esto se aprecia claramente con el valor de *recall*, que es 0,73 para la clase 1 o valencia alta, y 0,49 para la clase 0 o valencia baja. Esta diferencia tan considerable sugiere que RNA tiende a clasificar más muestras como valencia positiva, lo que indica un desbalance en el aprendizaje o una mayor sensibilidad del modelo a características asociadas a emociones de alta valencia.

Las matrices de confusión (Figura 26 y 27) muestran lo mencionado anteriormente, RF presenta una distribución más equilibrada de los errores, y RNA presenta un sesgo considerable hacia la alta valencia, por lo que hay más falsos positivos de alta valencia (51,4%) que verdaderos positivos de baja valencia (48,6%).

Por estas razones, RF es mejor modelo para la clasificación binaria de la valencia que RNA, debido a su mayor equilibrio entre clases.

4.1.2. Clasificación binaria de la excitación

Para la clasificación binaria de la excitación, RF obtuvo un mayor rendimiento, con un *accuracy* global de 62% frente a 53,8% de RNA. Además, RF ha mostrado métricas equilibradas entre las clases de *precision* y *recall*: 0,61 para la clase 1 y 0,63 para la clase 0.

En cuanto a las métricas de RNA, revelan un problema significativo para identificar de forma correcta estados de alta excitación, ya que mientras el *recall* para la clase de baja excitación es alto (0,73), el *recall* para la alta excitación es extremadamente bajo (0,33). Esto indica que se clasifican muchas muestras como de baja excitación, como se ve en la matriz de confusión de la Figura 29. Este problema podría estar causado o bien por la arquitectura de la red, o por las características de alta excitación de los datos que no son analizadas de forma correcta por el modelo RNA.

4.1.3. Clasificación multiclase de los 4 cuadrantes emocionales

La clasificación multiclase ha revelado peores resultados que las dos clasificaciones anteriores para ambos algoritmos. RF logra una exactitud global del 51% y RNA solo 36,3%, por lo que es superior RF; a pesar de ello, es un rendimiento insuficiente.

Como se puede ver en las Tablas 18 y 19, en la columna de “soporte”, las clases están bastante desbalanceada. El cuadrante Q1 es el que más muestras reales tiene como observaciones con 179, y el cuadrante Q3 el que menos con 69. Por esta razón es conveniente utilizar *F1-score* para ver cómo está clasificando el modelo los positivos. El cuadrante Q1 y Q2 son los mejores clasificados por el modelo RF, con un *F1-score* de 0,56 y 0,59, respectivamente. Por otro lado, Q3 y Q4 presentan el peor rendimiento, lo que sugiere que las emociones de baja excitación son más difíciles de clasificar correctamente. Esto puede ser debido a lo desbalanceadas que están las clases, ya que Q3 y Q4 tienen un menor soporte.

RNA no ha mostrado buenos resultados para la clasificación de ningún cuadrante, por lo que ninguno de los dos modelos ha obtenido un rendimiento aceptable.

4.2. Comparación con trabajos previos

El mejor resultado obtenido para la clasificación binaria de valencia y excitación ha sido 62% de *accuracy*, con el algoritmo de RF. Este rendimiento puede considerarse competitivo con otros estudios previos similares, aunque todavía presenta margen de mejora.

Se han seleccionado tres trabajos que han utilizado únicamente señales periféricas y la base de datos MAHNOB-HCI para comparar resultados. A continuación, se muestra una tabla comparativa con los resultados obtenidos en esos trabajos:

Trabajo	Señales utilizadas	Algoritmo	<i>Accuracy</i>
[19]	ECG, SKT, Respiración, EDA	SVM	Valencia: 64,23% Excitación: 68,75%
[20]	Respiración, ECG	SVM	Valencia: 54,57% Excitación: 60,41%
[21]	ECG, SKT, Respiración, EDA	SVM	Valencia: 45,5% Excitación: 46,2%

Tabla 20: Tabla comparativa del estado del arte.

En [19] se han utilizado las señales fisiológicas de ECG, respiración, SKT y EDA para la clasificación de emociones en función de la valencia y la excitación. En ese trabajo se utilizó un SVM para la clasificación binaria de la valencia y la excitación y se obtuvo un *accuracy* del 64,23% y 68,75%, respectivamente. Estos resultados son ligeramente superiores a los obtenidos

en el presente estudio, pero muy similares. Esta diferencia puede deberse a que en [19] se utiliza la señal de respiración.

En [20] se utilizan únicamente las señales de respiración y ECG, y con un SVM se obtiene un *accuracy* de 54,57% para la valencia y 60,41% para la excitación. Aunque los resultados son ligeramente inferiores a los de este estudio, cabe destacar que se emplearon menos señales fisiológicas, lo que puede explicar esa diferencia.

Por último, en [21] se hizo un estudio muy similar a [19] utilizando las mismas señales fisiológicas y el mismo modelo de clasificación, pero se obtuvieron resultados inferiores debido a que se llevó a cabo un proceso distinto de extracción de características y se configuró el SVM de forma diferente.

Los resultados obtenidos en trabajos previos utilizando señales periféricas están alrededor del 50% y el 60% de *accuracy*, valores que son comparables a los alcanzados en el presente estudio.

4.3. Limitaciones

Este trabajo presenta desafíos intrínsecos del propio problema que se aborda, como la naturaleza subjetiva de las emociones. Esta subjetividad hace que haya mucha variabilidad entre individuos e incluso en un mismo individuo en diferentes contextos. Esta variabilidad dificulta la clasificación de emociones de forma precisa utilizando modelos de aprendizaje automático, ya que no existen límites definidos entre estados emocionales y es complicado crear etiquetas discretas. Aunque el modelo circunflejo bidimensional es ampliamente usado, es una simplificación de las emociones reales del ser humano, y en la práctica, hay una gran variedad de emociones que no encajan en los cuadrantes del modelo o se encuentran en el límite.

En cuanto a las limitaciones metodológicas, la base de datos utilizada tenía un gran número de muestras (1201), pero sólo fueron seleccionados finalmente 463 ya que se estableció como criterio una duración mínima de 60 segundos. Este tamaño muestral puede resultar insuficiente para captar de forma correcta la variabilidad emocional, especialmente en el algoritmo de RNA que requiere mayor cantidad de datos de entrada, como se ha visto reflejado en los resultados. Además, en la clasificación multiclase se ha visto un desbalance entre clases notorio que dificulta el rendimiento de las clases minoritarias; en este caso, las de baja excitación.

Por último, en la metodología de este trabajo se planteó una hipótesis inicial de que las emociones se ven reflejadas al final de la señal fisiológica, por lo que se ha analizado la ventana temporal de los últimos 60 segundos. En caso de no ser correcta esta hipótesis, pueden no haberse obtenido todas las características asociadas al estado emocional y se ha visto afectada la calidad del conjunto de datos.

5. CONCLUSIÓN Y LÍNEAS FUTURAS

Se ha desarrollado un modelo de clasificación de emociones en términos de valencia y excitación de forma exitosa. Los resultados han sido competitivos con el estado del arte, por lo que el objetivo principal de este trabajo ha sido cumplido.

Los objetivos específicos también se han cumplido, ya que han sido necesarios completarlos con éxito para alcanzar el objetivo principal. Se ha seleccionado la base de datos MAHNOB-HCI y se han preprocesado las señales de ECG, EDA y SKT de este repositorio. Además, se ha diseñado un proceso automático de extracción y selección de las características más relevantes de dichas señales. Esta selección de características ha sido la entrada de los modelos predictivos desarrollados. Se han comparado los algoritmos RF y RNA para 3 problemas de clasificación diferentes: clasificación binaria de valencia, clasificación binaria de excitación, y clasificación multiclase de los 4 cuadrantes emocionales.

El mejor resultado obtenido ha sido 62% de exactitud con el modelo RF para la clasificación binaria de la valencia y la excitación. Al ser un resultado aceptable y similar a los del estado del arte, se han probado los modelos con señales fisiológicas propias registradas con los sensores específicos. De esta forma, se ha comprobado que se puede poner el modelo en práctica en un entorno real.

La validación de los modelos con señales propias sirvió para demostrar que es posible utilizar este modelo fuera de un entorno controlado, pero no fue útil para validar el resultado de la predicción, ya que en el momento en el que se registraron las señales no se sometió al sujeto a ningún estímulo emocional.

En futuros estudios se puede incluir un protocolo experimental para inducir respuestas emocionales en el momento en el que se recojan las señales. De esta forma, sería posible contrastar las predicciones del modelo y validar su aplicabilidad en un entorno real.

Además, aunque el modelo RF ha mostrado un rendimiento aceptable, todavía hay margen de mejora. En investigaciones futuras, sería interesante estudiar variantes del modelo, ajustando su configuración o utilizando otras técnicas.

Por último, en este trabajo se planteó la hipótesis de que las respuestas emocionales se reflejan de una forma más clara en los últimos 60 segundos de la señal fisiológica, y se extrajeron las características en esa ventana temporal. De cara a futuros trabajos, sería conveniente contrastar esta hipótesis. En lugar de centrarse en esta ventana temporal, se podrían explorar otras opciones como una ventana en la mitad de la señal o incluso analizar toda la señal con una ventana móvil. Esto podría mejorar el rendimiento obtenido en este trabajo y comprender mejor la dinámica temporal de las emociones.

6. REFERENCIAS

- [1] M. Chóliz, "Psicología de la Emoción: El Proceso Emocional," Universidad de Valencia, Valencia, 2005.
- [2] A. Olivera-La Rosa, O. E. Arango, and J. J. Martí, "¿Estrellas o constelaciones? Implicaciones de los estudios cognitivos para el modelo dimensional de la emoción," *Escritos de Psicología / Psychological Writings*, vol. 9, no. 2, 2016, doi: 10.5231/psy.writ.2016.1503.
- [3] J. A. Russell, "A circumplex model of affect," *J Pers Soc Psychol*, vol. 39, no. 6, 1980, doi: 10.1037/h0077714.
- [4] F. Elkfury and J. Ierache, "Clasificación y representación de emociones en el discurso hablado en español empleando Deep Learning," *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, no. 42, 2021, doi: 10.17013/risti.42.78-92.
- [5] O. Alemi, W. Li, and P. Pasquier, "Affect-expressive movement generation with factored conditional Restricted Boltzmann Machines," in *2015 International Conference on Affective Computing and Intelligent Interaction, ACII*, 2015. doi: 10.1109/ACII.2015.7344608.
- [6] J. Tamargo and Eva Delpón, "Función de bomba del corazón: el ciclo cardíaco," in *Fisiología humana*, 3rd ed., Tresguerres JA, Benitez EA, Cachofeiro MV, Cardinali DP, Gil Loyzaga P, and Lahera JV, Eds., Madrid: McGraw Hill Interamericana, 2012, pp. 485–506.
- [7] L. Azcona, "El electrocardiograma," in *Libro de la salud cardiovascular del Hospital Clínico San Carlos y la fundación BBVA.*, 1ª ed., López Farré A, Macaya Miguel C, and directores, Eds., Bilbao: Fundación BBVA, 2009, pp. 49–56.
- [8] N. Dugarte J, A. Alvarez, E. Dugarte, and G. Álvarez, "Técnicas de procesamiento de la señal ECGAR aplicadas en el prototipo DIGICARDIAC," *Revista del Instituto Nacional de Higiene Rafael Rangel*, vol. 46, no. 1–2, 2015.
- [9] A. Ortigado Matamala, "Lectura del ECG," *Pediatría Integral*, vol. 16, no. 9, 2012.
- [10] D. C. W. Hamm and S. Willems, *El electrocardiograma: su interpretación práctica; 32 cuadros*. Ed. Médica Panamericana, 2010.
- [11] C. V. Silva and V. G. Rojas, "Diseño e implementación de un sistema electrocardiográfico digital Design and implementation of a digital electrocardiographic system," *Rev. Fac. Ing. Univ. Antioquia*, vol. 55, 2010.
- [12] M. E. Moncada and J. de la Cruz, "La Actividad Electrodermica - Revisión," *Ingeniería e Investigación*, vol. 31, no. 2, 2011.
- [13] M. Sharma, S. Kacker, and M. Sharma, "A Brief Introduction and Review on Galvanic Skin Response," *International Journal of Medical Research Professionals*, vol. 2, no. 6, 2016, doi: 10.21276/ijmrp.2016.2.6.003.
- [14] J. D. Romero-Ante, J. M. Vicente-Samper, J. Manrique-Cordoba, V. Esteve, M. A. De la Casa-Lillo, and J. M. Sabater-Navarro, "Actividad electrodermica como indicador preventivo de complicaciones por neuropatía periférica," 2024.
- [15] P. J. LANG, M. K. GREENWALD, M. M. BRADLEY, and A. O. HAMM, "Looking at pictures: Affective, facial, visceral, and behavioral reactions," *Psychophysiology*, vol. 30, no. 3, 1993, doi: 10.1111/j.1469-8986.1993.tb03352.x.
- [16] E. H. Jang, B. J. Park, M. S. Park, S. H. Kim, and J. H. Sohn, "Analysis of physiological signals for recognition of boredom, pain, and surprise emotions," *J Physiol Anthropol*, vol. 34, no. 1, 2015, doi: 10.1186/s40101-015-0063-5.
- [17] R. A. McFarland, "Relationship of skin temperature changes to the emotions accompanying music," *Biofeedback Self Regul*, vol. 10, no. 3, 1985, doi: 10.1007/BF00999346.
- [18] H. Gunes and B. Schuller, "Categorical and dimensional affect analysis in continuous input: Current trends and future directions," *Image Vis Comput*, vol. 31, no. 2, 2013, doi: 10.1016/j.imavis.2012.06.016.
- [19] M. Ben and Z. Lachiri, "Emotion Classification in Arousal Valence Model using MAHNOB-HCI Database," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 3,

- 2017, doi: 10.14569/ijacsa.2017.080344.
- [20] W. Mimoun Ben Henia and Z. Lachiri, "Multiclass SVM for affect recognition with hardware implementation," in *2018 15th International Multi-Conference on Systems, Signals and Devices, SSD*, 2018. doi: 10.1109/SSD.2018.8570398.
 - [21] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Trans Affect Comput*, vol. 3, no. 1, 2012, doi: 10.1109/T-AFFC.2011.25.
 - [22] J. J. V. Mayor, N. J. V. Jiménez, G. P. A. Hoyos, and E. F. C. Bravo, "Sistema de Biofeedback para Rehabilitación de Marcha Asistida por un Exoesqueleto," *Rev Ing Biomed*, vol. 12, no. 24, 2019, doi: 10.24050/19099762.n24.2018.713.
 - [23] José Alberto Morales Morante, "Metodología para el preprocesamiento y clasificación de datos fisiológicos multimodales basado en el modelo Valencia-Excitación," Centro Nacional de Investigación y Desarrollo Tecnológico, Cuernavaca, Morelos, México, 2021.
 - [24] S. Koelstra *et al.*, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans Affect Comput*, vol. 3, no. 1, 2012, doi: 10.1109/T-AFFC.2011.15.
 - [25] D. Dobrev, T. Neycheva, V. Krasteva, and I. Jekiva, "Design of High-Pass and Low-Pass Active Inverse Filters to Compensate for Distortions in RC-Filtered Electrocardiograms," *Technologies (Basel)*, vol. 13, no. 4, Apr. 2025.
 - [26] A. R. Aguinaga, M. Lopez, A. Alanis, R. Baltazar, and V. M. Zamudio, "Emotion analysis through physiological measurements," *Workshop Proceedings of the 9th International Conference on Intelligent Environments*, vol. 17, pp. 97–106, 2013.
 - [27] C. Li, C. Zheng, and C. Tai, "Detection of ECG Characteristic Points Using Wavelet Transforms," *IEEE Trans Biomed Eng*, vol. 42, no. 1, 1995, doi: 10.1109/10.362922.
 - [28] Z. Vulaj, M. Brajovic, A. Draganic, and I. Orovic, "Detection of irregular QRS complexes using Hermite transform and support vector machine," in *Proceedings Elmar - International Symposium Electronics in Marine*, 2017. doi: 10.23919/ELMAR.2017.8124435.
 - [29] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans Biomed Eng*, vol. BME-32, no. 3, 1985, doi: 10.1109/TBME.1985.325532.
 - [30] G. Valenza, A. Lanata, and E. P. Scilingo, "The role of nonlinear dynamics in affective valence and arousal recognition," *IEEE Trans Affect Comput*, vol. 3, no. 2, 2012, doi: 10.1109/T-AFFC.2011.30.
 - [31] P. J. Bota, C. Wang, A. L. N. Fred, and H. Placido Da Silva, "A Review, Current Challenges, and Future Possibilities on Emotion Recognition Using Machine Learning and Physiological Signals," 2019. doi: 10.1109/ACCESS.2019.2944001.
 - [32] M. Malik *et al.*, "Heart rate variability. Standards of measurement, physiological interpretation, and clinical use," 1996. doi: 10.1093/oxfordjournals.eurheartj.a014868.
 - [33] J. Shukla, M. Barrada-Angeles, J. Oliver, G. C. Nandi, and D. Puig, "Feature Extraction and Selection for Emotion Recognition from Electrodermal Activity," *IEEE Trans Affect Comput*, vol. 12, no. 4, 2021, doi: 10.1109/TAFFC.2019.2901673.
 - [34] S. Kumar P, S. Kumar, and J. F. Agastinose Ronickom, "Investigating the Effects of Two-Class Categorical Emotion Classification Through Electrodermal Activity and Machine Learning," in *5th IEEE International Conference on Cybernetics, Cognition and Machine Learning Applications, ICCCMILA*, 2023. doi: 10.1109/ICCCMLA58983.2023.10346868.
 - [35] M. Awad and S. Fraihat, "Recursive Feature Elimination with Cross-Validation with Decision Tree: Feature Selection Method for Machine Learning-Based Intrusion Detection Systems," *Journal of Sensor and Actuator Networks*, vol. 12, no. 5, 2023, doi: 10.3390/jsan12050067.
 - [36] O. Kramer, "K-Nearest Neighbors," in *Dimensionality Reduction with Unsupervised Nearest Neighbors*, Berlin, Heidelberg, 2013, pp. 13-23. doi: 10.1007/978-3-642-38652-7_2.
 - [37] R. García, "El perceptrón: una red neuronal artificial para clasificar datos," *Revista de investigación en modelos matemáticos aplicados a la gestión y la economía*, pp. 1–14, 2021.
 - [38] D. A. Pérez-Aguilar, R. H. Risco-Ramos, and L. Casaverde-Pacherrez, "Transfer learning en la clasificación binaria de imágenes térmicas," *Ingenius*, no. 26, 2021, doi: 10.17163/ings.n26.2021.07.
 - [39] R. Salas, "Redes Neuronales Artificiales," Departamento de Computación, Univ. De Valparaíso,

- Valparaíso, Chile, vol. 1, pp. 1–7, 2004.
- [40] M. Schonlau and R. Y. Zou, “The random forest algorithm for statistical learning,” *Stata Journal*, vol. 20, no. 1, 2020, doi: 10.1177/1536867X20909688.
 - [41] Lucidchart, “Qué es un diagrama de árbol de decisión,” Lucidchart.com. <https://www.lucidchart.com/pages/es/que-es-un-diagrama-de-arbol-de-decision> (accessed Aug. 4, 2025).
 - [42] L. Breiman, “Random forests,” *Mach Learn*, vol. 45, no. 1, 2001, doi: 10.1023/A:1010933404324.
 - [43]. Usha Ruby Dr.A, “Binary cross entropy with deep learning technique for Image classification,” *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, 2020, doi: 10.30534/ijatcse/2020/175942020.
 - [44] K. Kyriakou *et al.*, “Detecting moments of stress from measurements of wearable physiological sensors,” *Sensors (Switzerland)*, vol. 19, no. 17, 2019, doi: 10.3390/s19173805.
 - [45] SparkFun Electronics, “AD8232 Heart Rate Monitor Hookup Guide,” SparkFun Electronics. <https://learn.sparkfun.com/tutorials/ad8232-heart-rate-monitor-hookup-guide/all> (accessed Aug. 4, 2025).
 - [46] Electrónica Embajadores, “SparkFun CAB-12970 – Cable con 3 electrodos ECG de sensores biométricos a jack 3.5 mm,” Electrónica Embajadores. <https://www.electronicaembajadores.com/es/Productos/Detalle/SSBIECGCX1/sensores/sensore-s-biometricos/sparkfun-cab-12970-cable-con-3-electrodos-ecg-de-sensores-biometricos-a-jack-3-5-mm/> (accessed Aug. 4, 2025).
 - [47] A. Álvarez Urueña, “Implementación de un entorno de comunicación Bluetooth, basado en el módulo CC2650, para la transmisión del ritmo cardiaco,” Trabajo Fin de Grado, Universidad de Valladolid. Escuela de Ingenierías Industriales, Valladolid, 2021. [Online]. Available: <https://uvadoc.uva.es/bitstream/handle/10324/47901/TFG-I-1903.pdf;jsessionid=B82152BC109A33B31386C111BA5D55D9?sequence=1>
 - [48] Seeed Studio, “Grove - GSR Sensor,” Seeed Studio Wiki. https://wiki.seeedstudio.com/Grove-GSR_Sensor/ (accessed Aug. 4, 2025).