



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



Departamento de Sistemas Informáticos y Computación
Universitat Politècnica de València

Análisis de emociones de sujetos a partir de señales de EEG, una aproximación con modelos de aprendizaje automático

TRABAJO FIN DE MÁSTER

Máster Universitario en Inteligencia Artificial, Reconocimiento de Formas e Imagen Digital

Autor: José Javier Calvo Moratilla

Tutores: Jose Antonio Gil Gomez
Carlos Monserrat Aranda

Curso 2021-2022

Agradecimientos

A todas las personas que dedicaron parte de su tiempo para ayudarme a comprender que la sociedad siempre está en continuo cambio y que tenemos que ser partícipes de ello, a mis profesores/ras del IES Salvador Gadea de Aldaia, CE. Cheste por ayudarme a crecer, a mi tía Fanny Llorens por haber creído en mí desde Barcelona, a mi padre José Vicente Calvo por transmitirme su amor por la cultura y las nuevas tecnologías, a mi familia por estar ahí en los tiempos más difíciles, a mi equipo del Hospital Oftalmológico de Valencia, Pablo, Juanjo, Carlos, Gabi, Javi, Edu, Joel, Gustavo, Patxi, Ximo por ser y haber sido parte de mi segunda familia durante estos últimos cinco años y en último lugar a mi tutores Jose Antonio Gil Gomez del Instituto de automática y informática industrial de la UPV por la oportunidad de poder investigar en el instituto y a Carlos Monserrat Aranda del DSIC por el apoyo con el área de la inteligencia artificial a la hora de desarrollar el presente proyecto.

Resum

El cervell és un dels òrgans més desconeguts pels éssers humans, sobre el que any re any s'adquireix coneixement sobre el seu funcionament real. La seva anàlisi ens ajuda a entendre el seu comportament de manera més precisa permetent descobrir com maneja el control d'elements vitals essencials del cos humà. L'objectiu d'aquest projecte és analitzar com el cervell gestiona les emocions. Per a això, s'utilitzarà un mètode d'adquisició de dades del cervell (senyors EEG) i, mitjançant la seva anàlisi, poder interpretar el senyal rebut i discriminar entre diferents situacions emocionals de cada individu. Per a l'anàlisi d'aquests senyals complexos es faran servir tècniques d'aprenentatge automàtic buscant aquells models que millor puguin predir les emocions dels subjectes. Els resultats obtinguts amb Random Forest milloren els resultats de projectes similars per a la classificació de Arousal, amb un 66% de precisió, la resta de models no aconsegueixen millorar l'estat de la qüestió.

Paraules clau: Electroencefalograma, EEG, Aprenentatge automàtic, classificació, intel·ligència artificial, xarxes neuronals, emocions, cervell, senyals

Resumen

El cerebro es uno de los órganos más desconocidos por los seres humanos, sobre el que año tras año se adquiere conocimiento sobre su funcionamiento real. Su análisis nos ayuda a entender su comportamiento de manera más precisa permitiendo descubrir cómo maneja el control de elementos vitales esenciales del cuerpo humano. El objetivo de este proyecto es analizar cómo el cerebro gestiona las emociones. Para ello, se utilizará un método de adquisición de datos del cerebro (sensores EEG) y, mediante su análisis, poder interpretar la señal recibida y discriminar entre diferentes situaciones emocionales de cada individuo. Para el análisis de estas señales complejas se utilizarán técnicas de aprendizaje automático buscando aquellos modelos que mejor puedan predecir las emociones de los sujetos. Los resultados obtenidos con Random Forest mejoran los resultados de proyectos similares para la clasificación de Arousal, con un 66 % de precisión, el resto de modelos no consiguen mejorar el estado de la cuestión.

Palabras clave: Electroencefalograma, EEG, Aprendizaje automático, clasificación, inteligencia artificial, redes neuronales, emociones, cerebro, señales

Abstract

The brain is one of the most unknown organs by human beings, about which year after year we acquire knowledge about its natural functioning. Its analysis helps us to understand its behaviour in a more precise way, allowing us to discover how it manages the control of essential vital elements of the human body. This project aims to analyze how the brain operates emotions. To do this, we will use a method of data acquisition from the brain (EEG sensors) and, through its analysis, interpret the signal received and discriminate between different emotional situations of each individual. For the analysis of these complex signals, machine learning techniques will be used to search for those models that can best predict the subjects' emotions. The results obtained with Random Forest improve the results of similar projects for the Arousal classification, with 66% accuracy, the rest of the models fail to improve the state of the art.

Key words: Electroencephalogram, EEG, machine learning, classification, artificial intelligence, neural networks, emotions, brain, signals

Índice general

Índice general	V
Índice de figuras	VII
Índice de tablas	VIII
1 Introducción	1
1.1 Motivación	1
1.2 Objetivos	2
1.3 Estructura de la memoria	3
2 Antecedentes	5
2.1 Bases neurofisiológicas del EEG	5
2.1.1 El cerebro humano	5
2.1.2 Las neuronas	6
2.1.3 Tipos de onda	6
2.1.4 Instrumentación	8
2.1.5 Técnicas de registro	9
2.1.6 Bases de interpretación del EEG	11
2.1.7 Artefactos	11
2.1.8 Técnicas especiales de análisis y registro	12
2.2 Aprendizaje automático	14
2.2.1 Fundamentos	14
2.2.2 Árboles de decisión	17
2.2.3 Multi Layer Perceptron	19
2.3 Métricas	20
2.3.1 Coeficiente de Sorensen-Dice (<i>F1-Score</i>)	20
2.4 Situación actual de la tecnología	21
2.4.1 Investigating the Use of Pretrained Convolutional Neural Network on Cross-Subject and Cross-Dataset EEG Emotion Recognition	21
2.4.2 Classification of Human Emotions from Electroencephalogram (EEG) Signal using Deep Neural Network	23
2.4.3 Emotion Recognition Based on DEAP Database using EEGTime-Frequency Features and Machine Learning Methods	25
3 Materiales y métodos	27
3.1 DEAP Dataset	27
3.1.1 Valencia	28
3.1.2 Arousal	28
3.1.3 Áreas cerebrales	29
3.2 Diseño	29
3.2.1 Especificación de requerimientos	29
3.2.2 Metodología	30
3.3 Herramientas utilizadas	31
3.3.1 Software	31
3.3.2 Lenguajes de programación	32

3.3.3	Librerías	32
3.4	Técnicas de aprendizaje automático utilizadas	33
3.4.1	Random Forest	33
3.4.2	Multi Layer Perceptron	33
3.4.3	XGBoost	34
3.5	Implementación	34
3.5.1	Búsqueda del dataset	35
3.5.2	Carga del dataset y preprocesamiento de los datos, división de ancho de banda, Método de Welch	36
3.5.3	Implementación y experimentación usando un modelo Random Forest	38
3.5.4	Implementación y experimentación usando un modelo Multi Layer Perceptron	40
3.5.5	Implementación y experimentación usando un modelo XGBoost	42
4	Resultados	45
4.1	Evaluación modelo Random Forest	45
4.2	Evaluación modelo Multi Layer Perceptron	46
4.3	Evaluación modelo XGBoost	47
5	Discusión y conclusiones	49
5.1	Crítica y discusión de los resultados	49
5.2	Principales aportaciones	50
5.3	Trabajos futuros	51
	Bibliografía	53

Índice de figuras

2.1	Áreas del cerebro humano, Universidad de Córdoba, Argentina	5
2.2	Cargas eléctricas en una neurona. EEG Pocketguide	6
2.3	Tipos de onda EEG	7
2.4	Grupo de electrodos, Natus	8
2.5	Sistema 10-20, Oscar Humberto, ResearchGate	9
2.6	Sistema geodésico. EEG Pocketguide.	10
2.7	Ejemplo de artefacto fisiológico. EEG Pocketguide.	11
2.8	Proceso de discretización. Toptal	13
2.9	Comparativa FFT con Welch, Mike X Cohen	14
2.10	Artificial Intelligence. Medium	14
2.11	Ejemplo de segmentación semántica. Analyticsindiamag	15
2.12	Ejemplo de reconocimiento de caracteres (OCR). Pyimagesearch	16
2.13	Ejemplo de reconocimiento de voz, Apple (SIRI). Apple	16
2.14	Ejemplo de reconocimiento de patrones de movimiento en un partido de fútbol. Laboratory of Biological Networks, Center for Biomedical Technology (UPM)	17
2.15	Árbol de decisión, Cristobal Aguirre	17
2.16	Random Forest, towards data science	18
2.17	Multi Layer Perceptron (MLP), Mohamed Zahran	19
2.18	F1-Score, The Data Scientist	20
2.19	Cálculo de la precisión	20
2.20	Cálculo del recall	21
2.21	Métrica F1-Score	21
2.22	Arquitectura de la solución, InceptionResnetV2, Fully Connected	22
2.23	Ejemplo en el uso de ventanas deslizantes con superposición	22
2.24	Power Spectral Density (PSD)	24
2.25	Modelo de clasificación Denso Fully Connected para clasificación EGG	25
2.26	Procedimiento Artículo EEG, Modelo SVM, Random Fores, K-NN, wk-NN	25
3.1	Interpretación valor de Valencia	28
3.2	Interpretación valor de Arousal	28
3.3	Localización de los electrodos en las áreas del cerebro humano	29
3.4	Backlog del proyecto	31
3.5	División de los datos en conjuntos	34
3.6	Etiquetas originales DEAP dataset en Numpy Array	36
3.7	Formato One Hot Encoding de las etiquetas	36
3.8	Periodograma de una muestra EEG al aplicar Welch	37
3.9	Discriminación tipo de onda al aplicar Welch	37
3.10	Características más importantes Random Forest, Valencia	38
3.11	Características más importantes Random Forest, Arousal	38
3.12	Características más importantes MLP, Valencia	40
3.13	Características más importantes MLP, Arousal	41
3.14	Características más importantes XGBoost, Valencia	42

3.15	Características más importantes XGBoost, Arousal	42
4.1	Resultados Random Forest, Valencia	45
4.2	Resultados en gráfica, Random Forest, Valencia	45
4.3	Resultados Random Forest, Arousal	46
4.4	Resultados en gráfica, Random Forest, Arousal	46
4.5	Resultados MLP, Valencia	46
4.6	Resultados en gráfica, MLP, Valencia	46
4.7	Resultados MLP, Arousal	47
4.8	Resultados en gráfica, MLP, Arousal	47
4.9	Resultados XGBoost, Valencia	47
4.10	Resultados en gráfica, XGBoost, Valencia	47
4.11	Resultados XGBoost, Arousal	48
4.12	Resultados en gráfica, XGBoost, Arousal	48
4.13	Gráfica resumen resultados evaluación modelos	48

Índice de tablas

2.1	Bandas de frecuencias para EEG	7
2.2	Casos reales de uso bandas	8
2.3	Nomenclatura electrodos en sistema 10-20	9
2.4	Agentes fisiológicos	11
2.5	Agentes externos	12
3.1	DEAP Dataset	27
3.2	DEAP Dataset, Data Preprocessed	27
3.3	Duraciones de los <i>Sprints</i> para el desarrollo del proyecto	31
3.4	Nº de muestras por conjunto	35
3.5	Resultados balance de las etiquetas en los datos	36
3.6	hiperparámetros Random Forest en la selección de características más significativas	38
3.7	Selección de hiperparámetros en Random Forest	39
3.8	Mejores hiperparámetros para Valencia, Random Forest	40
3.9	Mejores hiperparámetros para Arousal, Random Forest	40
3.10	Selección de hiperparámetros en MLP	41
3.11	Mejores hiperparámetros para Valencia, MLP	41
3.12	Mejores hiperparámetros para Arousal, MLP	41
3.13	Selección de hiperparámetros en XGBoost	43
3.14	Mejores hiperparámetros para Valencia, XGBoost	43
3.15	Mejores hiperparámetros para Arousal, XGBoost	43
5.1	Comparativa resultados Random Forest en la tarea de clasificación de la etiqueta Valencia y Arousal	49
5.2	Comparativa resultados Multi Layer Perceptron (MLP) en la tarea de clasificación de la etiqueta Valencia y Arousal	49
5.3	Comparativa resultados XGBoost en la tarea de clasificación de la etiqueta Valencia y Arousal	50

CAPÍTULO 1

Introducción

El cerebro es uno de los órganos más desconocidos por los seres humanos, dónde año tras año se adquiere conocimiento sobre su funcionamiento real, que nos ayuda a leer su comportamiento de manera más precisa. Está siempre en funcionamiento, coordinando y realizando procesos vitales esenciales para el funcionamiento del cuerpo humano.

El presente trabajo utiliza un método de adquisición de datos del cerebro, utilizando las señales recibidas para identificar la situación emocional de cada individuo. Se identifica la *Valencia y el Arousal* para detectar las emociones.

En la tarea de adquisición de datos se utiliza el método EEG (Electroencefalograma), que es clave para obtener información de los campos eléctricos generados por las neuronas, dado que se leen directamente en el área cerebral. Los electrodos están separados del cerebro solo por el cuero cabelludo y la superficie ósea del cráneo. Es utilizado ampliamente para la obtención de diagnósticos de patologías como la esquizofrenia, o el insomnio y actualmente se utiliza para clasificar las emociones de los seres humanos.

Otra de las ventajas del EEG es que presenta un presupuesto reducido y una mayor portabilidad para obtener los resultados, en comparación con otras áreas de la medicina, como la adquisición de datos WSI.[2] Únicamente se precisa de un conjunto de electrodos y un medidor al extremo para obtener el valor de potencia leído, en el caso de la obtención de datos WSI se precisan equipos de electromedicina complejos y excesivamente caros.

La adquisición de datos elegida no es invasiva pero está limitada a unas condiciones estipuladas para que la lectura adquirida sea lo más limpia posible. Dichas limitaciones presentan problemas a la hora de normalizar los datos y la generalización de los modelos de clasificación, dificultando más la tarea.

1.1 Motivación

La clasificación de sentimientos mediante el uso de técnicas basadas en aprendizaje automático es un área que precisa de un continuo esfuerzo en horas investigación, donde se observan grandes retos que deben de superarse, mejorando la tecnología de captación de las señales cerebrales y la construcción de modelos que sean capaces de generalizar las diferencias existentes entre sujetos.

La utilización de las nuevas tecnologías aplicadas al ámbito sanitario psicológico reducen la carga del personal médico dentro de los hospitales y agilizan la tarea de detección de patologías en los pacientes. De este modo, se incrementa la probabilidad de supervivencia, mejorando enormemente la calidad del servicio sanitario y reduciendo la lista de espera en los hospitales.

Las posibilidades a la hora de clasificar los campos eléctricos que generan las neuronas en el cerebro ofrecen múltiples aplicaciones dentro del campo de la medicina, siendo la clasificación de patologías médicas una de las áreas dónde más se ha extendido su uso para la detección de esquizofrenia o falta de sueño.

Existen nuevas áreas innovadoras como la codificación de señales cerebrales para realizar acciones, dónde puede ayudar en la tarea de creación y control de exoesqueletos en personas con parálisis cerebral en extremidades, pudiendo controlar el movimiento de las mismas mediante las señales cerebrales, algo que sería muy positivo en la rehabilitación en hospitales.

Poder identificar los sentimientos de una persona mediante el uso de señales de electroencefalograma (EEG) es una oportunidad para seguir explorando áreas relacionadas con la medicina, en concreto con la psicología médica.

Se abarca el reto por la necesidad personal de querer ayudar al campo de la medicina, concretamente en el área de la psicología médica.

1.2 Objetivos

El presente trabajo final de máster se enmarca en la detección de emociones dentro del campo de la psicología médica dónde, mediante el uso técnicas de aprendizaje automático es posible clasificar emociones de sujetos, mediante el análisis de señales como la *Valencia y Arousal*, obtenidas a partir de señales de electroencefalograma (EEG).

El principal objetivo del trabajo es probar el rendimiento de diferentes modelos basados en aprendizaje automático para la tarea en cuestión y tratar de mejorar los resultados obtenidos con modelos similares por otros equipos de investigación.

De este modo, se han planteado los siguientes objetivos secundarios:

- Búsqueda de dataset
- Carga del dataset y preprocesamiento de los datos, división de ancho de banda, Método de Welch
- Implementación y experimentación usando un modelo Random Forest
- Implementación y experimentación usando un modelo Multi Layer Perceptron
- Implementación y experimentación usando un modelo XGBoost
- Evaluación modelo Random Forest
- Evaluación modelo Multi Layer Perceptron
- Evaluación modelo XGBoost

En primera instancia se realiza un estudio de los datasets existentes de señales de electroencefalograma (EEG) para la tarea de clasificación de las clases *Valencia y Arousal*. De este modo, se podrán comparar los resultados obtenidos con estudios previos.

Posteriormente, con el dataset seleccionado, se organiza la estructura de carpetas en los servidores de la plataforma *Google Colab*, donde se almacenan los datos para el desarrollo del proyecto, instalando las librerías necesarias para la implementación del código para los experimentos y la evaluación de los mismos.

Seguidamente, el proceso de entrenamiento de los modelos planteados precisa de un preprocesado de datos, dónde primero se transforma el dominio temporal de los datos al dominio de frecuencia, utilizando el método de *Welch* y por último se filtran cada una de las señales por tipo de onda y por electrodo.

Una vez realizado el preprocesado de datos se procede a la implementación del código necesario para la definición e implementación de los modelos basados en aprendizaje automático. Posteriormente, se realiza experimentación con los diferentes modelos, seleccionando los hiperparámetros más significativos y la cantidad de características relevantes para obtener el mayor rendimiento.

Para terminar el proyecto, se procede a realizar las pruebas de evaluación de los modelos de aprendizaje automático generados y, así, medir su rendimiento y comparar los resultados con modelos similares, publicados en otros artículos científicos.

1.3 Estructura de la memoria

El trabajo se estructura en cinco capítulos:

En el primer capítulo se introduce el trabajo realizado, abordando las motivaciones más significativas para la elección del tema, los objetivos marcados para la consecución del proyecto y la estructura de la memoria.

En el segundo capítulo se aborda los antecedentes, dónde se recogen las bases neurofisiológicas del EEG, las bases teóricas más relevantes que afectan al campo del aprendizaje automático, las métricas utilizadas para la evaluación de los modelos y por último se enumera la situación actual de la tecnología. En el último punto se explican los artículos más relevantes relacionados con la temática del presente trabajo, con el fin de tener datos relevantes para interpretar el rendimiento de nuestro modelo.

En el tercer capítulo, se abordan los métodos seguidos en el proceso de investigación junto a los materiales dónde se se presenta el Dataset utilizado, los criterios de diseño con la especificación de requerimientos, la metodología utilizada para desarrollar el proyecto, seguido por las herramientas, librerías para realizar los experimentos correspondientes, todo seguido se enumeran las técnicas de aprendizaje automático seleccionadas en el presente trabajo y por último se muestra el trabajo de investigación realizado.

En el cuarto capítulo, se muestran los resultados obtenidos por la evaluación de los modelos utilizados en la investigación.

En el quinto y último capítulo, se discuten los resultados de la investigación realizada, siendo comparados con trabajos similares. También se enumeran las principales aportaciones realizadas y por último se tratan los posibles trabajos futuros para lograr un mayor rendimiento en los modelos de clasificación de emociones basados en aprendizaje automático.

CAPÍTULO 2

Antecedentes

2.1 Bases neurofisiológicas del EEG

2.1.1. El cerebro humano

El cerebro humano está dividido en diferentes áreas, donde cada una de ellas está especializada en una tarea esencial para el sistema nervioso central.

En primer lugar se identifica el tronco del encéfalo, área de forma cilíndrica, que conecta el cerebro con la médula espinal. Está especializada en el control de procesos automáticos como el latido del corazón, la respiración, la función de la vejiga y control del equilibrio.

En segundo lugar se identifica el sistema límbico, área localizada en el centro del cerebro, que incluye el tálamo, el hipotálamo y la amígdala, encargada de las tareas emocionales del cuerpo, la memorización o el aprendizaje.

En tercer lugar se identifica el cerebelo, está localizado en la parte posterior del cerebro en el área del tronco encefálico, tiene una forma redondeada como el cerebro y es la encargada de regular y controlar los movimientos correctos de los seres humanos.

En último lugar se identifica la corteza cerebral, que ocupa la mayor parte del órgano y está formada por dos hemisferios, el izquierdo y el derecho. Se encarga de controlar el pensamiento consciente.

Si se analiza la corteza cerebral en mayor profundidad, ésta se divide en cuatro áreas llamadas lóbulos, que corresponden con el occipital, temporal, parietal y frontal.

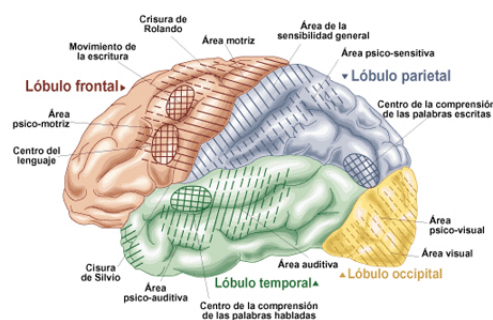


Figura 2.1: Áreas del cerebro humano, Universidad de Córdoba, Argentina

Primero se observa el lóbulo occipital, área encargada del sentido de la visión, correspondiente con la interpretación del espacio, la orientación, el color o la percepción del movimiento.

En segundo lugar se identifica el lóbulo temporal, encargado de la memoria a largo plazo, obtenida del reconocimiento de la visión, la interpretación del lenguaje natural o de las emociones.

En tercer lugar se observa el lóbulo parietal, cuya función principal es adquirir información de los músculos, extremidades, cabeza, ojos, etc. El lóbulo adquiere dicha información y se obtiene una representación coherente para la interacción de nuestro cuerpo con el entorno.

Por último, se examina el lóbulo frontal, región donde se toman las decisiones para realizar acciones voluntarias en el cuerpo humano y que contiene la mayoría de las neuronas sensibles a la dopamina.

2.1.2. Las neuronas

En cada una de las áreas descritas las neuronas se excitan y comunican con otras mediante el uso de señales eléctricas. La comunicación es posible gracias a las cargas negativas de los axones y la carga positiva del cuerpo de una célula piramidal.

El potencial post sináptico de una sola neurona es pobre, por ello el proceso lo desarrollan un gran número de neuronas de manera conjunta, provocando un campo eléctrico más significativo.

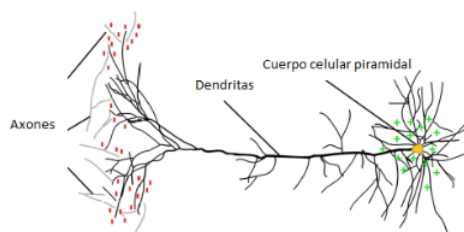


Figura 2.2: Cargas eléctricas en una neurona. EEG Pocketguide

Las células piramidales se localizan en las áreas corticales, orientadas perpendicularmente con la superficie, alejados de la superficie, mientras que las dendritas sí que están más próximas con la superficie. La orientación de las neuronas generan un campo eléctrico estable, en contraposición las neuronas del tronco del encéfalo o del cerebelo no tienen dicha orientación.

Uno de los problemas es que existen altas probabilidades de que los campos eléctricos sean propagados en múltiples direcciones, provocando así su cancelación, evitando ser proyectados de manera estable a la superficie del cuero cabelludo, fuente de obtención de datos mediante el uso de EEG.

2.1.3. Tipos de onda

Para la lectura de los campos eléctricos generados por las neuronas primero se debe de amplificar la señal, ya que los campos eléctricos son reducidos. Una vez amplificada

la señal se identifica la señal de potencia mediante las variables de frecuencia, potencia y fase.

Los avances en dicho campo ha permitido identificar diferentes rangos de frecuencias que son significativas para la discriminación entre tareas o procesos cognitivos del cerebro humano.

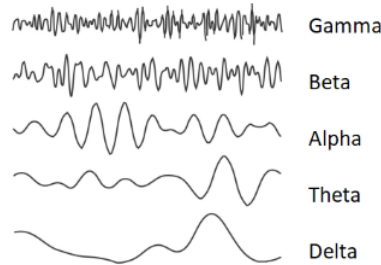


Figura 2.3: Tipos de onda EEG

En la siguiente tabla, se identifican las bandas de frecuencias definidas para la lectura de campos eléctricos mediante electroencefalograma EEG:

Banda	Rango de frecuencias	Referencia
Delta	[1-4] Hz	Niedermeyer & da Silva, 2012
Theta	[4-8] Hz	Niedermeyer & da Silva, 2012
Alpha	[8-12] Hz	Hans Berger, 1929
Beta	[12-25] Hz	Niedermeyer & da Silva, 2012
Gamma	>25 Hz	Dimigen, 2009

Tabla 2.1: Bandas de frecuencias para EEG

Haciendo un análisis detallado de cada una de las bandas utilizadas en EEG, en primer lugar se observa la banda *delta*, dónde se identifican problemas relacionados con los trastornos del sueño, común en enfermedades como el parkinson, demencia o esquizofrenia y problemas relacionados con la ingesta de alcohol, hábito también estrechamente relacionado con los trastornos de sueño.

En segundo lugar, la banda *theta* muestra una gran correlación con la dificultad de las operaciones mentales, tales como la adquisición de información, su procesamiento, el aprendizaje, etc. En otras palabras, toda tarea que implique carga de trabajo mental.

En tercer lugar, se observa la banda *Alpha*, cuyas frecuencias identifican funciones sensoriales, de movimiento y de memorización. Según (Pfurtscheller y Aranibar, 1977) la activación de dicha señal está correlacionada con la actividad de reposo mental y físico del cuerpo, en contraposición la señal se ve drásticamente reducida cuando la actividad mental está en pleno funcionamiento.

En cuarto lugar, se identifica la banda *Beta*, cuyas frecuencias tienen un gran grado de correlación con actividades de pensamiento activo, ocupado o ansioso junto a la concentración activa. Se observa una excitación mayor cuando se realizan tareas motoras que precisan de movimientos precisos junto a una gran concentración y cuando el sujeto observa el movimiento motor de otras personas, (Zhang et al., 2008) demostró que el cerebro humano tiende a imitar el movimiento de otros sujetos, realizando un proceso de movimiento “espejo”.

En quinto y último lugar, se observa la banda *gamma*, cuyas frecuencias están provocando un debate científico, ya que actualmente no se relaciona con un proceso particular. Los estudios realizados afirman que dichas frecuencias sirven para procesos de atención o que reflejan un subproceso de actividades motoras como el movimiento ocular, no relacionadas con procesos cognitivos.

Casos reales de uso delta
Trastornos del sueño
Alcoholismo y sueño
Casos reales de uso theta
Tareas N-Back
Navegación espacial
Monitorización cerebral en entornos operacionales
Casos reales de uso alpha
Meditación
Entrenamiento “Bio feedback”
Atención
Casos reales de uso beta
Control motor
Inducción del estado de alerta mediante impulsos
Casos reales de uso gamma
Correlación movimiento ocular frente a estímulos visuales

Tabla 2.2: Casos reales de uso bandas

2.1.4. Instrumentación

En la tarea de obtención de señal en el área cerebral se utilizan electrodos en la superficie craneal, justo encima del cuello cabelludo, cuyo proceso se identifica como “electroencefalograma” o “EEG”.

El proceso consiste en la medición potencial de los campos eléctricos generados por las neuronas en el proceso de sinapsis.

Cada electrodo consiste en un sensor de forma plana, circular, que habitualmente se une al cuero cabelludo mediante el uso de geles, pasta o crema conductora, soluciones salinas que ayudan a mejorar la calidad de la medición, reduciendo significativamente el ruido en la señal.



Figura 2.4: Grupo de electrodos, Natus

Existen electrodos que no necesitan de sustancias para la conexión con la piel, pero añaden más ruido a la señal.

Entendiendo condiciones óptimas, la conjunción de la piel, electrodos y la solución salina, realizan la función de un condensador, atenuando la transmisión de frecuencias bajas.

Los electrodos normalmente son producidos con las aleaciones de Plata (Ag) junto a una capa superficial de cloruro de plata (AgCl).

Debido al material utilizado junto a la solución salina, en el transcurso del tiempo los electrodos se deterioran, provocando mediciones más ruidosas, por ello antes de realizar cualquier experimento se debe de verificar el estado de los electrodos, para así minimizar el ruido.

2.1.5. Técnicas de registro

La técnica de registro más utilizada se denomina “sistema 10-20”, definido por la American Encephalographic Society (1994) y por Ostenveld & Praamstra (2001).

La técnica consiste en la localización de los electrodos con una agrupación matricial, donde la distancia entre cada uno de los electrodos corresponde con el 10 % o 20 % de la longitud total de la matriz de medición.

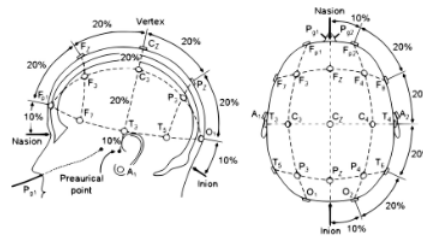


Figura 2.5: Sistema 10-20, Oscar Humberto, ResearchGate

En la matriz del sistema se identifican tres puntos importantes, que corresponden con el nasion, inion y los puntos preauriculares. Como se observa en la figura 2.5 el nasion se localiza en la zona central de la frente, el inion está localizado en la parte opuesta y por último los puntos preauriculares izquierdo y derecho se localizan en la parte superior posterior del pabellón de cada una de las orejas.

Para la identificación de los electrodos se utiliza la siguiente nomenclatura que identifica a cada una de las regiones:

Letra	Descripción
Fp	Frontopolar
F	Frontal
C	Central
P	Parietal
O	Occipital
T	Temporal

Tabla 2.3: Nomenclatura electrodos en sistema 10-20

El siguiente término que sigue a la denominación de la región corresponde con un número entero o letra que identifica la distancia respecto a la línea central. Los números impares corresponden a los electrodos del hemisferio izquierdo, los números pares a los

electrodos del hemisferio derecho. Por último, los electrodos que corresponden con la línea central se identifican con la letra “z”.

Como técnica alternativa de registro existe otra configuración de electrodos llamada “geodésica”, en la cuál todos los electrodos tienen una distancia similar entre vecinos.

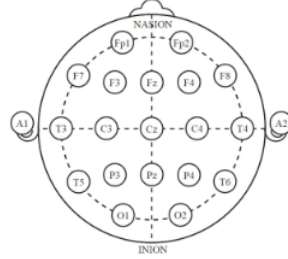


Figura 2.6: Sistema geodésico. EEG Pocketguide.

Dado que cada sujeto tiene unas dimensiones del cráneo diferentes el posicionamiento de los electrodos deberá de ser personalizado para el tamaño craneal.

Otro de los puntos importantes es el número de electrodos utilizados, que en el sistema internacional 10-20 se identifican hasta 21 electrodos.

Para poder medir el campo eléctrico, los electrodos tienen un punto de referencia a tierra, identificado por la letra “G”, que provoca por razones físicas ruido en la muestra medida.

Para solucionar el problema de ruido generado por el punto de conexión de tierra se define otro punto de referencia identificado por la letra “R”.

Primero se calcula la diferencia de potencial entre (Cz, G) y (R, G):

$$T_1 = Cz - G$$

$$T_2 = R - G$$

Que se traduce en:

$$T_3 = Cz - G - R + G$$

Y simplificado en:

$$T = Cz - R$$

Dónde se concluye que la potencia de un electrodo es respecto al punto de referencia, minimizando el problema de ruido obtenido desde G.

Otro de los problemas que se debe de tener en cuenta es la medición de la impedancia, como punto de control de la lectura realizada y cuyo valor tiene que ser lo más reducido posible, ya que la conexión de los electrodos con el cuello cabelludo implica que la superficie no es del todo limpia.

Existen células muertas en la superficie de la piel, piel grasa y sudor que pueden provocar una impedancia mayor, que se reduce siguiendo diferentes recomendaciones:

- Los sujetos del experimento a quienes se les colocan los electrodos deben de tener el pelo lavado y seco. El uso de elementos metálicos para la sujeción del pelo o la utilización de productos para fijar la posición del peinado aumenta el valor de la impedancia.
- Los electrodos tienen que estar debidamente limpios por alcohol isopropanol al 70 %.
- Por último, para la aplicación de los electrodos se recomienda una solución salina, como gel, como se ha comentado anteriormente en la presente memoria.

2.1.6. Bases de interpretación del EEG

La señal obtenida por los electrodos precisa de procesamiento para poder ser almacenada digitalmente. El proceso se identifica como amplificación y digitalización.

La lectura se realiza con la toma de mediciones instantáneas en un periodo de tiempo, donde se recomienda una tasa de muestreo de 128 Hz para que no haya sobrecarga de datos almacenados.

Una vez realizada la amplificación y digitalización de la señal se envía a un ordenador, dónde se almacena y utiliza para su interpretación.

En las mediciones se observan los artefactos que deberán de ser eliminados de la señal para poder interpretar la señal de manera idónea, sin interferencias.

2.1.7. Artefactos

Los artefactos son todas las mediciones que quedan reflejadas en la señal y que distan de la actividad cerebral a identificar en los experimentos realizados.

Los artefactos aparecen debido a que los electrones recogen campos eléctricos de diferentes fuentes del entorno, por ello deben de tratarse, realizando un proceso de limpieza manual, ya que no existe algoritmo que pueda realizar la tarea de manera automática.

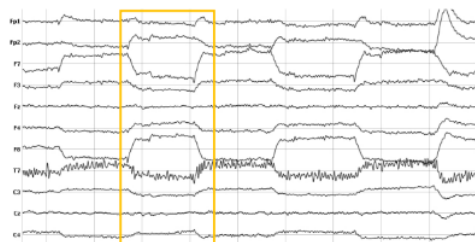


Figura 2.7: Ejemplo de artefacto fisiológico. EEG Pocketguide.

Se identifican dos tipos de artefactos, por una parte los generados por agentes fisiológicos y por otra los generados por agentes externos:

Agentes fisiológicos
Actividad muscular
Movimiento ocular
Parpadeo

Tabla 2.4: Agentes fisiológicos

En primer lugar, como agente fisiológico generador de artefactos se identifica la actividad muscular, que afecta en gran medida a la lectura de datos desde los electrodos. Los músculos más influyentes son los que se localizan cerca del área craneal, como todos los músculos de la cara, boca o cuello.

Otra de las señales musculares registradas por los electrodos corresponde con el latido del corazón, que deberá de ser eliminado mediante el uso de sensores basados en fotopletimografía, PPG o en otro dispositivo basado en electroencefalograma EEG para la obtención de la medición correspondiente al latido.

Los movimientos oculares también afectan a la medición generando campos eléctricos detectables por los electrodos, debido al gran número de neuronas existentes en la retina. Con el movimiento ocular dicho campo eléctrico se desplaza y genera el artefacto. Se utilizan elementos de seguimiento ocular o más electrodos EEG para intentar identificar dichos campos, para así poder ser eliminados de la señal original. Por último el parpadeo del ojo humano también provoca cambios significativos en la señal, dónde se recomienda que a la hora de obtener los datos en el experimento el sujeto no parpadee de manera sistemática, sino se utilizan soluciones basadas en la estadística como la regresión lineal, interpolación o la separación ciega de fuentes para minimizar dicho artefacto.

En el caso de los agentes externos, se identifican los movimientos de los electrodos en la cabeza, el ruido provocado por la señal y el balanceo y la oscilación del sujeto del que se está obteniendo la medición.

Agentes externos
Movimientos de electrodos en la cabeza
Ruido de la señal
Balanceo y oscilación del sujeto

Tabla 2.5: Agentes externos

En primer lugar, el movimiento de los electrodos puede afectar significativamente a la medición, por ello se tiene que asegurar que los electrodos están debidamente unidos y ajustados al cuello cabelludo para evitar así dicho artefacto.

En segundo lugar, se observa el ruido natural de la señal existente en todo campo eléctrico. Dado que el objetivo para mejorar la señal obtenida era reducir la impedancia medida como referencia, ésta puede provocar un aumento del ruido en la señal. Como las frecuencias del cerebro humano son más reducidas se pueden discriminar fácilmente del resto de señales adquiridas.

Por último, se identifican los artefactos provocados por el balanceo y la oscilación del sujeto al que se le está realizando el experimento. Si los movimientos realizados son bruscos puede afectar de manera significativa a la interpretación de la señal, dificultando la discriminación de la misma.

2.1.8. Técnicas especiales de análisis y registro

Transformada de Fourier

Para realizar el proceso de registro de información se cambia el concepto de lectura del campo eléctrico por el de la frecuencia como dato que va a almacenarse.

Para poder utilizar la señal de manera digital y ser almacenada, ésta tiene que pre-procesar y se discretiza.

La discretización consiste en representar la señal analógica con unos valores digitales que representan la amplitud de la señal.

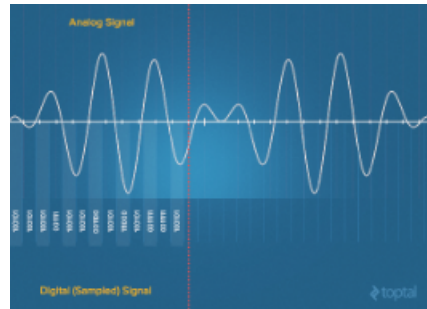


Figura 2.8: Proceso de discretización. Toptal

Gracias a la investigación de Jean-Baptiste Joseph Fourier, la demostración de la serie de Fourier permite representar una señal periódica sólo con el valor de las frecuencias, amplitud y fase de cada senoide, convirtiendo la señal en una suma de senos y cosenos que permiten realizar una aproximación de la señal real:

$$\frac{a_0}{2} + \sum_{n=1}^a \cos \frac{2n\pi}{T} + b_n \sin \frac{2n\pi}{T} t$$

Gracias a la serie de Fourier, se obtiene la representación discreta o la frecuencia de dominio como “huella digital”, característica para poder identificar la señal obtenida por los electrodos del EEG. En resumen, se obtienen todas las frecuencias y las magnitudes correspondientes de una señal en un intervalo temporal definido.

Para el proceso se puede utilizar el algoritmo “Fast Fourier Transform” o (FFT) con un coste computacional (n^2). Sin embargo, con la variación del algoritmo llamada “Cooley-Tukey” (DFT) se consigue un coste computacional de ($n \log n$).

El objetivo para el diseño de la “huella digital” es encontrar las frecuencias más importantes que identifican a actividades cerebrales concretas para realizar el análisis de la señal.

Método del Welch

Aplicar el método de Welch es una mejora que permite suavizar el espectro de potencia generado por FFT completa, para obtener así una señal mucho menos oscilante.

En el método anterior, se obtiene la transformada de Fourier aplicando la operación en toda la ventana de máxima temporal de una señal. Sin embargo, en el método de Welch se aplica una ventana deslizante con tiempo reducido, con un valor determinado donde se van aplicando FFT en cada una de las ventanas individualmente, generando una señal resultante más suavizada.

El desplazamiento de la ventana no corresponde con el tamaño de la ventana, el mismo se mueve a una distancia menor para ir superponiendo los diferentes segmentos temporales.

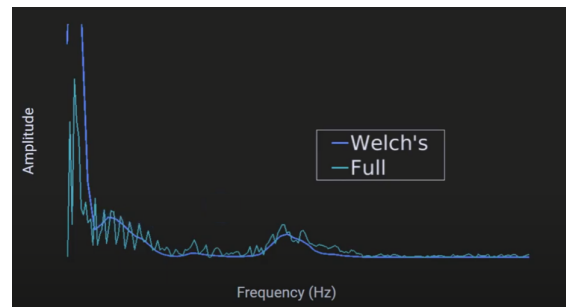


Figura 2.9: Comparativa FFT con Welch, Mike X Cohen

El método de Welch permite suavizar parte del ruido que puede ser recogido por la FFT completa, pero por contra se reduce la precisión espectral por el movimiento de la ventana, algo que sí aporta el método completo. El método descrito en segundo lugar se ha utilizado para el preprocesamiento de la señal para el presente trabajo.

2.2 Aprendizaje automático

2.2.1. Fundamentos

El aprendizaje automático *machine learning* es una de las ramas de la inteligencia artificial que dota a la computadoras el poder del autoaprendizaje para la resolución de problemas. El término aglutina diferentes disciplinas que se desarrollan con anterioridad y que participan en la acuñación del término en los años 80, siendo las más importantes el reconocimiento de patrones (*pattern recognition*) y la inteligencia artificial.

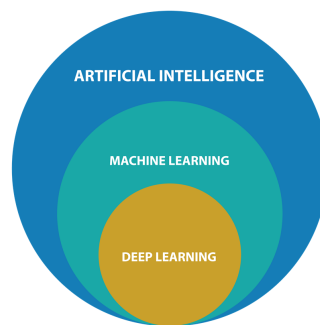


Figura 2.10: Artificial Intelligence. Medium

Se enumeran las tres técnicas de aprendizaje automático más importantes:

- Aprendizaje supervisado
- Aprendizaje no supervisado
- Aprendizaje semi-supervisado

En primer lugar, se observan los sistemas con aprendizaje supervisado, donde no se tiene conocimiento previo sobre el problema a resolver y se aprende mediante datos de entrenamiento de entrada y salida.

Se clasifican en cuatro tipos: inductivos, de memorización, deductivos, por analogía, todos ellos aplicados en problemas de clasificación o segmentación de imágenes.

Seguidamente se identifican los sistemas de aprendizaje no supervisado, donde los datos de entrenamiento utilizados solo tienen datos de entrada y cuyo aprendizaje da como resultado un agrupamiento de muestras de datos en diferentes grupos o *clusters*. Estos se clasifican en dos tipos: los jerárquicos y no jerárquicos.

Los clusters jerárquicos no precisan el número de *clusters* como punto de partida para ser encontrados: los algoritmos detectan todos los grupos posibles en los datos. En cambio, en los no jerárquicos, los algoritmos disponen del número de grupos a detectar como premisa.

En último lugar, se contemplan los sistemas con aprendizaje semi-supervisado, un planteamiento que comparte características de los dos tipos enumerados en los dos párrafos anteriores. Estos algoritmos disponen de datos etiquetados y no etiquetados. Buscan ser más efectivos que con un planteamiento supervisado o no supervisado.

La mejora de los algoritmos y el aumento del poder computacional a lo largo de los años ha permitido que el uso y aceptación del aprendizaje automático sea generalizado dentro del campo de la investigación y en el mundo empresarial, identificando los siguientes cuatro campos de aplicación más importantes:

- Reconocimiento de imágenes y vídeo
- Reconocimiento de audio
- Reconocimiento de texto
- Análisis de datos

El reconocimiento de imágenes se aborda dentro del campo de la visión por computador y dota a cualquier sistema del poder de percepción de su entorno. Estos se aplican en tareas de reconocimiento facial en dispositivos móviles, en la conducción autónoma de vehículos o para la detección de enfermedades en imágenes médicas, siendo las arquitecturas convolucionales las más utilizadas para la resolución de este tipo de problemas.

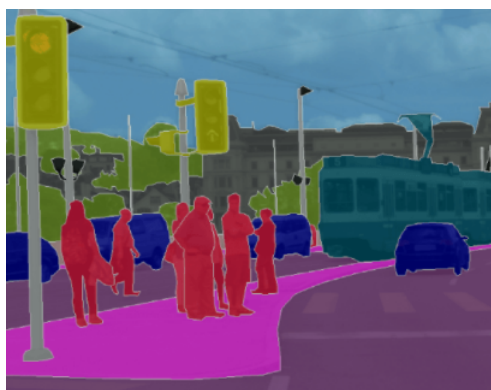


Figura 2.11: Ejemplo de segmentación semántica. Analyticsindiamag

El reconocimiento de texto también se engloba dentro del campo de la visión por ordenador, pero también en el procesamiento del texto natural (*Natural Language Processing*). Éste se utiliza para digitalizar libros de formato físico a digital mediante el reconocimiento de texto manuscrito con técnicas de reconocimiento óptico de caracteres OCR

con arquitectura convolucional o mediante el uso de redes recurrentes *RNN*, más recientemente, mediante el uso de redes *Transformers*.

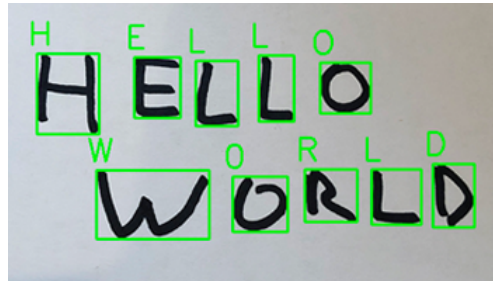


Figura 2.12: Ejemplo de reconocimiento de caracteres (OCR). Pyimagesearch

El reconocimiento de audio realiza un proceso de transformación de audio a texto mediante arquitecturas convolucionales que detectan patrones con mucha efectividad, utilizado en dispositivos móviles para transcripción a texto o en la traducción instantánea, también en los sistemas de detección de voz para el control de dispositivos domóticos o para conocer el estado anímico de las personas en una conversación.

Una vez transformados los datos sonoros a texto, se utilizan las arquitecturas de redes recurrentes utilizadas en el reconocimiento de texto descritas en el párrafo anterior.

El campo de aplicación descrito ha hecho posible la utilización de dispositivos de personas con diversidad funcional de diferente índole, pudiendo adaptar la tecnología a las necesidades del usuario.

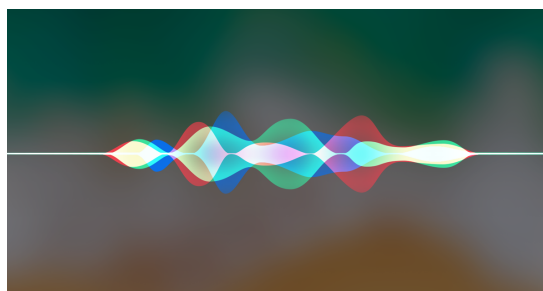


Figura 2.13: Ejemplo de reconocimiento de voz, Apple (SIRI). Apple

Por último, el análisis de datos permite analizar gran cantidad de datos y detectar patrones en ello, utilizando algoritmos de clustering y ciencia de datos. Éstos son utilizados, por ejemplo, por grandes empresas para sus decisiones directivas o por equipos deportivos para fichar nuevos jugadores o estudiar a rivales por parte de los equipos técnicos.

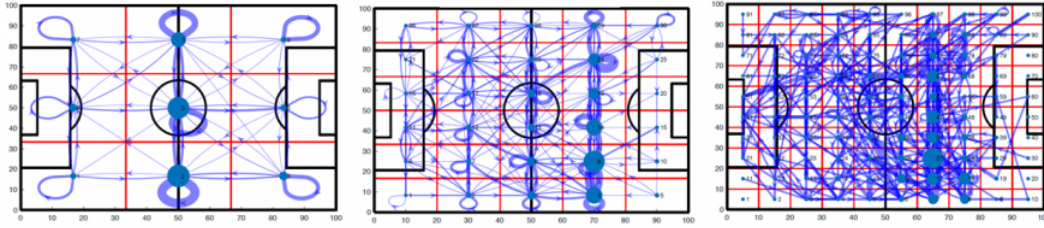


Figura 2.14: Ejemplo de reconocimiento de patrones de movimiento en un partido de fútbol. Laboratory of Biological Networks, Center for Biomedical Technology (UPM)

El uso de datos está aumentando debido a que las grandes organizaciones están dedicando mayores esfuerzos en la recopilación de los mismos, cuidando la calidad y el etiquetado.

Todo ello es importante, pero sin olvidar la necesidad de poder computacional para el entrenamiento de arquitecturas más complejas que han ayudado al auge del *Cloud Computing* y la publicación de librerías como *Eddi* o *Tensor Flow*, reduciendo la complejidad matemática en la implementación de grandes soluciones.

Una vez definido los aspectos teóricos más importantes que afectan al trabajo, se enumeran las técnicas utilizadas en la realización del presente trabajo: los árboles de decisión para los modelos de *Random Forest* y *XGBoost* y, por último, el modelo *Multi Layer Perceptron*.

2.2.2. Árboles de decisión

Los árboles de decisión forman una estructura jerárquica basada en nodos y hojas, donde se evalúa una condición. Cuando se evalúa una condición dependiendo del resultado obtenido, se recorren las diferentes ramas del árbol donde, después de haber recorrido los nodos hasta una hoja final, se obtiene un resultado.

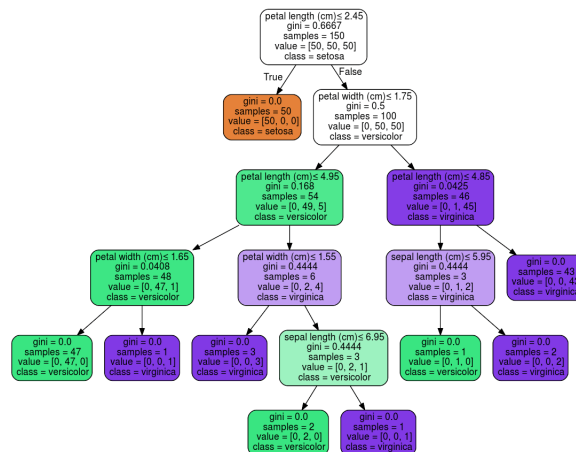


Figura 2.15: Árbol de decisión, Cristobal Aguirre

La condición aplicada en cada nodo implica que una característica cumple o no una condición, si se cumple se deriva por la rama izquierda, si no se cumple por la rama derecha.

Para el aprendizaje del algoritmo en cada nodo se aplica una operación de Doctomización, donde se aplica un criterio de partición de conjunto basado en el decremento de la impureza (entropía) para una partición dada:

$$I(t) = - \sum_{c=1}^C \hat{P}(c|t) \log_2 \hat{P}(c|t)$$

$$\hat{P}(c|t) = \frac{N_c(t)}{N(t)}$$

Para utilizar la medida de impureza en el entrenamiento se maximiza el decremento de entropía:

$$(d^*, r^*) = \arg \max_{d,r} \Delta(d, r)$$

$$\Delta I(d, r) = I(t) - \frac{N(L(t))}{N(t)} I(L(t)) - \frac{N(R(t))}{N(t)} I(R(t))$$

Una vez entrenado el modelo las condiciones de cada nodo están definidas. Así, se puede utilizar el modelo para la tarea de clasificación o regresión.

Random Forest

Una vez claras las bases de un árbol de decisión se describe el clasificador Random Forest. El modelo está formado por un gran número de árboles de decisión formando un mismo ente. Cada árbol predice una clase y, para la decisión final, se escogen cuáles son la clase o las clases más votadas por dichos árboles.

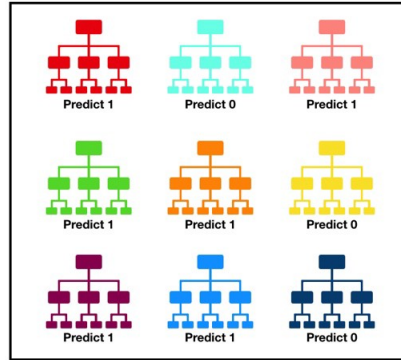


Figura 2.16: Random Forest, towards data science

El funcionamiento conjunto de un gran número de árboles de decisión mejora el resultado obtenido por un solo árbol de decisión. [1]

Además, mediante las estimaciones realizadas por los random Forest se puede obtener la correlación de las características más significativas para la clasificación.

XGBoost

Otro de los algoritmos más importantes basados en árboles de decisión es XGboost[3]. El algoritmo toma como estructura la utilización de "k"árboles de decisión, como en el

caso de los *Random Forest* pero la diferencia más significativa es que cada nuevo árbol intenta corregir los errores del árbol anterior.

En otras palabras, cuando un árbol realiza una predicción, observa las muestras que clasifica de manera errónea y le indica a los siguientes árboles que lo tengan en cuenta para la minimización del error. En el caso de *Random Forestm* en cambio directamente cada árbol tiene un funcionamiento autónomo:

$$l_{regresion} = (y_i - \hat{y}_i)^2$$

$$l_{clas} = y_i \ln(1 + e^{-\hat{y}_i}) + (1 - y_i) \ln(1 + e^{\hat{y}_i})$$

Para determinar el resultado final, se elige de la misma manera, dónde la predicción coincide con la decisión más unánime entre todos los árboles:

$$\hat{y}_i^{(t)} = - \sum_{k=1}^t f_k(X_i) = \hat{y}_i^{(t-1)} + f_t(x_i)$$

2.2.3. Multi Layer Perceptron

El *Multi Layer Perceptron* o (MLP)[10] es uno de los últimos modelos utilizados en el proyecto. Identificado como un algoritmo de aprendizaje automático basado en redes neuronales, su arquitectura está basada en un número determinado de capas de entrada, ocultas y de salida, que contienen como unidad elemental el perceptrón[9], elemento que se conecta con otros perceptrones de capas anteriores y posteriores.

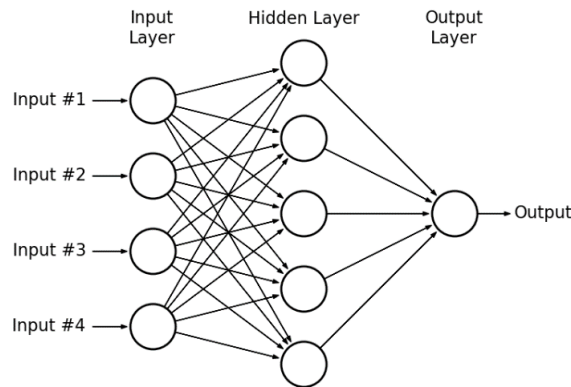


Figura 2.17: Multi Layer Perceptron (MLP), Mohamed Zahran

La red neuronal se construye en forma matricial para facilitar el cómputo de los cálculos en el entrenamiento y minimizar el uso de memoria dentro del computador. El proceso de entrenamiento se realiza en dos fases: la propagación hacia adelante y la retropropagación. De esta manera la red es capaz de entrenar hasta converger a una solución obteniendo el menor error posible a la salida.

En la propagación hacia adelante o (Forward Propagation), cada neurona utiliza una función de activación y la matriz de pesos almacenada para calcular el valor de salida, todo antes de ser propagado a la capa siguiente. Cuando las últimas neuronas propagan el resultado a la salida, se calcula el gradiente de cada capa con el algoritmo de retropropagación (Back Propagation) y, en dicho proceso, se calcula el valor de la matriz de pesos

de las neuronas que minimizan el error a la salida de la red, actualizando los valores de la matriz.

Actualmente, las redes neuronales se utilizan en numerosas aplicaciones, especialmente en visión por computador, en el reconocimiento de lenguaje natural humano o en la detección de patrones en los datos.

Uno de los principales problemas de las redes neuronales es que para su entrenamiento se necesitan más datos que en los modelos clásicos como *Random Forest* o *Vecinos más cercanos (KNN)*, para que el modelo pueda generalizar debidamente y no se produzcan problemas de sobreajuste.

2.3 Métricas

Las métricas son utilizadas para comprobar el rendimiento de los modelos utilizados. En el presente trabajo se identifica el *Coeficiente de Sorensen-Dice (F1-Score)*.

2.3.1. Coeficiente de Sorensen-Dice (F1-Score)

Métrica definida en los trabajos de Thorvald Sørensen [7] y Lee Raymond Dice [8] en la década de los 40, cuya métrica se utiliza para comprobar la similitud de dos muestras teniendo en cuenta dos variables, la precisión y el *recall*.

Las dos variables de la función se obtienen gracias a la casuística observada en la Figura 2.18:

	Predicted 0	Predicted 1
Actual 0	TN	FP
Actual 1	FN	TP

Figura 2.18: *F1-Score*, The Data Scientist

La precisión se utiliza para observar, en un conjunto de predicciones realizadas por un modelo, cuantas son realmente positivas. Al observar la fórmula de la Figura 2.19, se comprueba que si se incrementan los casos de falsos positivos, el valor se decrementa, demostrando que el modelo no es preciso.

$$\text{Precisión} = \frac{TP}{TP + FP}$$

Figura 2.19: Cálculo de la precisión

Su uso es bastante relevante, por ejemplo, para observar el buen funcionamiento de los filtros de *e-mails spam* de los clientes de correo, para filtrar el correo que realmente sea útil para el usuario.

En el caso del *recall* se utiliza para medir de todo el conjunto de positivos reales, cuántos se han predicho como falsos positivos, observando el cálculo en la la Figura 2.20.

$$Recall = \frac{TP}{TP + FN}$$

Figura 2.20: Cálculo del *recall*

El *recall* es una buena métrica para ver si en un hospital, donde se utilizan modelos predictivos, se dan diagnósticos negativos de una enfermedad cuando el paciente la tiene en realidad.

El coeficiente de Sorensen-Dice utiliza los valores de precisión y recall como se indica en la Figura 2.21 para obtener un balance entre ellos:

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

Figura 2.21: Métrica *F1-Score*

La métrica de puntuación *F1* se utiliza para medir el rendimiento modelo de aprendizaje automático implementado en el proyecto dentro del proceso de experimentación de los modelos. Sus valores están comprendidos en el rango [0, 1] siendo el valor de uno el mejor resultado obtenido por la métrica.

2.4 Situación actual de la tecnología

En el siguiente punto, se analizan los artículos científicos más relevantes en la clasificación de emociones con modelos basados en aprendizaje automático a partir de datos de electroencefalograma (EEG), almacenado en DEAP Dataset.

2.4.1. Investigating the Use of Pretrained Convolutional Neural Network on Cross-Subject and Cross-Dataset EEG Emotion Recognition

La solución propuesta por Yucel Cimtay, Erhan Ekmekcioglu [4] se basa en un modelo de red neuronal formada por una red *InceptionResnetV2* preentrenada, encargada de obtener las características automáticamente, a diferencia de los métodos clásicos donde las características se obtienen manualmente, identificando los métodos de *Welch* y la *transformada de Fourier (FFT)*, cómo métodos más utilizados en trabajos similares.

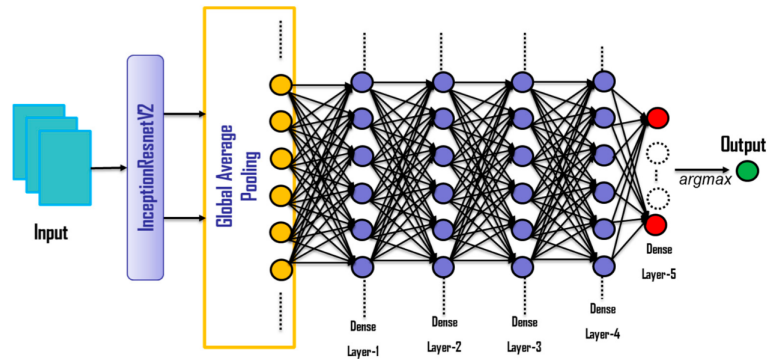


Figura 2.22: Arquitectura de la solución, InceptionResNetV2, Fully Connected

Una vez obtenidas las características, se utiliza una red *Fully Connected* con cuatro capas densas para el entrenamiento y una última capa de salida con dos neuronas.

Una vez identificada la arquitectura, se analiza cada uno de los aspectos más importantes recogidos en el artículo científico que hacen referencia a los procesos a tener en cuenta a la hora de utilizar la misma:

Ventanas deslizantes con superposición

Se define una ventana temporal de tiempo N utilizada para obtener segmentos de datos de una muestra, cuyo desplazamiento se realiza con una superposición de tamaño $(N/6)$.

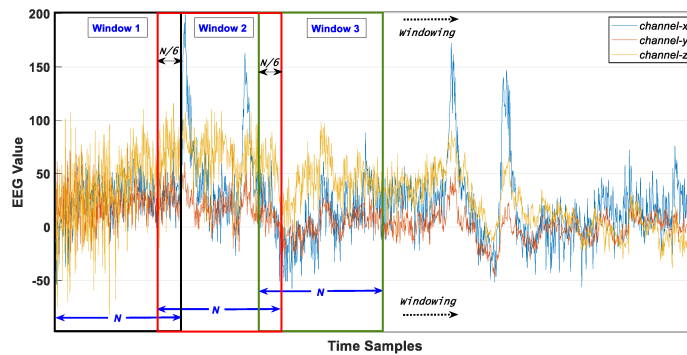


Figura 2.23: Ejemplo en el uso de ventanas deslizantes con superposición

En la muestra se obtienen los datos de los tres canales (x , y , z). Después del uso de ventanas deslizantes, los datos precisan de una remodelación de tamaño y una normalización. Una vez aplicado el preproceso descrito se derivan como entrada a la red *InceptionResNetV2* para su entrenamiento, experimentación y evaluación.

Transformación en la dimensión de los datos y Data Augmentation

Como primera fase del preproceso de los datos, se realiza un *Reshaping* de los datos obtenidos para que el tamaño del vector de los datos entradas coincidan con las dimensiones de la red *InceptionResNetV2*.

Al utilizar Keras, las dimensiones mínimas exigidas son de la orden:

$$(N_1, N_2, 3)$$

donde:

$$N_1 \geq 75, N_2 \geq 75$$

Una vez definido el tamaño se realiza un proceso de *Data Augmentation*, dónde se obtienen nuevas muestras similares con ruido, cuyo tamaño total se calcula con la siguiente ecuación:

$$NNC = \text{ceil}\left(\frac{N_1}{S} - 1\right)$$

Para aplicar ruido en el proceso de *Data Augmentation* se añaden muestras con valores aleatorios siguiendo una distribución gaussiana de media μ , y varianza σ^2 .

Normalización

Después del *Reshaping* anterior, se aplica una normalización para eliminar la media de cada ventana en cada muestra, tanto en los datos normales como en los datos añadidos en el *Data Augmentation*.

Selección de canales

Una vez aplicada la normalización se realiza una selección de características. Los autores del artículo concluyen que, al utilizar características que no son importantes para detectar una emoción en particular, reducen la precisión obtenida por el modelo.

Las características seleccionadas en el artículo corresponden con los electrodos (AF1, F3, F4, F7, T7, AF2, F5, F8, T8). Cabe destacar que en la solución que proponen no se discrimina con el tipo de onda.

Resultados en artículo

La arquitectura *InceptionResnetV2*, *Fully Connected* obtiene una precisión media entre sujetos del 72.81 % con el uso de del dataset *Database for Emotion Analysis using Physiological Signals (DEAP)*.

2.4.2. Classification of Human Emotions from Electroencephalogram (EEG) Signal using Deep Neural Network

La solución propuesta por el equipo *College of Computer and Information Sciences* de la *King Saud University Riyadh, Saudi Arabia* [6] proponen una aproximación basada en redes neuronales utilizando *DEAP Dataset*.

Las características generadas se recogen en los dos siguientes puntos:

Power Spectral Density (PSD)

El artículo propone transformar transformar los datos del dominio *tiempo* al dominio de la *frecuencia*, utilizando la transformada rápida de Fourier (FFT).

La realización del proceso PSD, es similar al realizado en el presente trabajo, donde se obtiene la distribución de la potencia de la señal en diferentes intervalos de frecuencias, para obtener de manera individualizada las frecuencias theta (4-8 Hz), alfa (8-13 Hz), beta (13-30 Hz) y gamma (30-40 Hz). Para la realización del proceso utilizan la librería *Python Signal*.

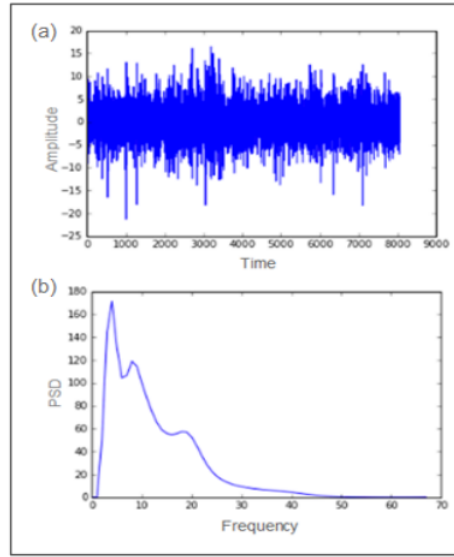


Figura 2.24: Power Spectral Density (PSD)

Frontal EEG Asymmetry

La propuesta diferencial respecto al resto de artículos y trabajos es la utilización de la Asimetría EEG Frontal, dado que en estudios anteriores consideran que existe una relación directa entre la asimetría de activación frontal *Alpha* y la valencia del estado emocional del sujeto analizado.

El hemisferio izquierdo se activa significativamente en estados de valencia positiva, mientras que en el hemisferio derecho en estados de valencia negativa, por ello utilizan la característica para la clasificación de *Valencia y Arousal*.

Para la obtención de las características se utilizan los siguientes electrodos individualizados por tipo de onda *Alpha* y *Beta*:

$$Valence = \frac{\alpha(F4)}{\beta(F4)} - \frac{\alpha(F3)}{\beta(F3)}$$

$$Arousal = \frac{\alpha(AF3 + AF4 + F3 + F4)}{\alpha(AF3 + AF4 + F3 + F4)}$$

Modelo de clasificación

El artículo propone la utilización de una red neuronal densa de 4 capas ocultas, con unas dimensiones de (2194, 1310, 786, 472) neuronas respectivamente.

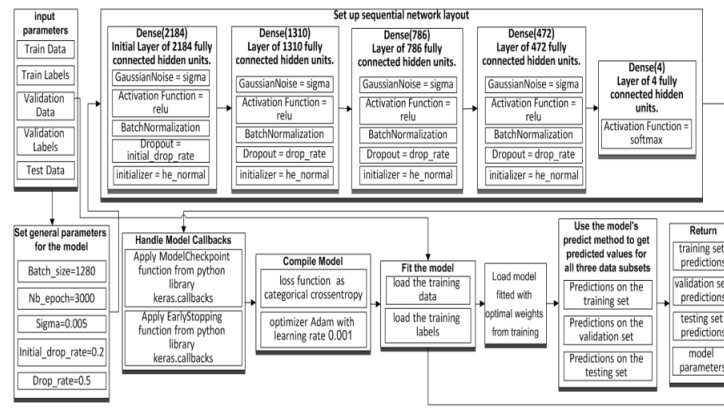


Figura 2.25: Modelo de clasificación Denso Fully Connected para clasificación ECG

Resultados en artículo

La obtención de características utilizada en el artículo en combinación con la red neuronal densa *Fully Connected* obtiene una precisión media entre sujetos del 82 % con el uso de del dataset *Database for Emotion Analysis using Physiological Signals (DEAP)* tanto para *Valencia* como para *Arousal*.

2.4.3. Emotion Recognition Based on DEAP Database using EEG Time-Frequency Features and Machine Learning Methods

La solución propuesta por T.D.Kusumaningrum, A.Faqih, B.Kusumoputro [5] sigue un procedimiento similar al presente trabajo, pero la única diferencia significativa es que se utiliza la técnica de *Fast Transformer Fourier* para obtener las características de entrada al modelo de aprendizaje automático.

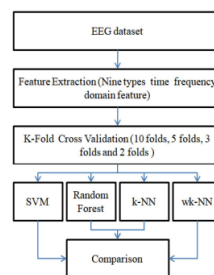


Figura 2.26: Procedimiento Artículo EEG, Modelo SVM, Random Fores, K-NN, wk-NN

Obtención de características

Para la obtención de las características se realiza el proceso PSD, es similar al realizado en el presente trabajo, donde se obtiene la distribución de la potencia de la señal en diferentes intervalos de frecuencias, para obtener de manera individualizada las frecuencias *alfa* (8-15 Hz), *beta* (16-31 Hz), *gamma* (>32 Hz) y *theta* (4-7 Hz).

Se excluye el uso de las frecuencias *delta* dado que la clasificación de *Valencia* y *Arousal* se realiza en un estado consciente, las señales *delta* se activan cuando el sujeto está en estado de sueño.

Con la señal transformada al dominio de la frecuencia se obtienen las características: peak to peak mean, mean squared value, variance, maximum power spectral frequency, power spectral density and power sum y del dominio de tiempo (parameters, complexity, mobility and activity).

Modelo de clasificación

Uno de los modelos utilizados en el artículo es el *Random Forest*, utilizado en el presente proyecto con otro tipo de aproximación original. Se selecciona como parámetro un número máximo de 100 árboles de clasificación, que corresponde con el 10 % del total de características de los datos.

Experimentación del método

Utilizan el método de validación cruzada *n-fold* y prueban el número de grupo de datos con el que se obtiene el mejor resultado. Seleccionan un valor de cinco para la variable (n).

Resultados en artículo

La obtención de características utilizada en el artículo en combinación de las características descritas y el modelo Random Forest obtiene una precisión media entre sujetos del 62.58 % con el uso de del dataset *Database for Emotion Analysis using Physiological Signals (DEAP)* tanto para *Valencia* como para *Arousal*.

CAPÍTULO 3

Materiales y métodos

3.1 DEAP Dataset

DEAP: A Database for Emotion Analysis using Physiological Signals [11] es el dataset seleccionado para la realización del presente proyecto. La elección del dataset es resultado de un proceso de búsqueda exhaustiva de datos, siendo DEAP el conjunto de datos más estandarizado para la tarea en particular de cara a la comparación de resultados con otros proyectos.

El dataset está formado por un gran número de datos disponibles, con sus respectivos formatos para realizar experimentos de todo tipo:

Datos	Formato
Online Ratings	xls, csv, ods spreadsheet
Video List	xls, csv, ods spreadsheet
Participant Ratings	xls, csv, ods spreadsheet
Participant Questionnaire	xls, csv, ods spreadsheet
Face Video	Zip file
Data Original	Zip file
Data Preprocessed	Zip file for Python and Matlab

Tabla 3.1: DEAP Dataset

Para la realización del proyecto se utiliza *Data Preprocessed* que contiene la información formateada para la carga desde el código de python.

Los datos preprocesados contienen dos tipos de archivos:

Nombre array	Array Shape	Array Contents
data	(40x40x8064)	video/trial x channel x data
labels	(40x4)	video/trial x label (valence, arousal, dominance, liking)

Tabla 3.2: DEAP Dataset, Data Preprocessed

En primer lugar se identifican los datos, 32 ficheros que contienen los 40 experimentos realizados a cada sujeto individual, con la obtención total de 40 características y un número de 8064 muestras en total. Los datos están limpios de artefactos pero no se ha aplicado

un preproceso previo con el método de Welch. Después del preproceso las muestras se reducirán considerablemente.

Se identifican 32 características relacionadas con los electrodos de EEG y unas última ocho características donde se identifica *hEOG*, *vEOG*, *zEMG*, *tEMG*, *GSR*, *Respiration belt*, *Plethysmograph* y *Temperature* que se obvian para la realización del presente trabajo.

En último lugar se identifican las etiquetas para cada uno de los experimentos realizados a cada sujeto, dónde se obtiene un valor real de las etiquetas *Valence*, *Arousal*, *Dominance*, *Liking*. El presente proyecto se centra en dos, en la *Valencia* y el *Arousal*.

Las etiquetas hacen referencia al nivel fisiológico de activación del cerebro en el cerebro humano.

3.1.1. Valencia

El término *Valencia* identifica a emociones que siente el ser humano. Cuando el valor de medición es alto se identifica la alegría, pero cuando el valor es bajo se identifica la ira o el miedo.



Figura 3.1: Interpretación valor de Valencia

3.1.2. Arousal

El término *Arousal* identifica a emociones que siente el ser humano. Cuando el valor de medición es alto se identifica la excitación intensa, pero cuando el valor es bajo se identifica el sueño profundo.

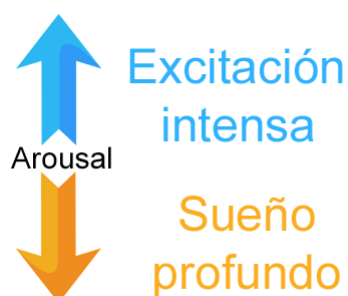


Figura 3.2: Interpretación valor de Arousal

3.1.3. Áreas cerebrales

Cada electrodo se agrupa en una zona cerebral *Left*, *Right*, *Frontal*, *Parietal*, *Occipital*, *Central*:

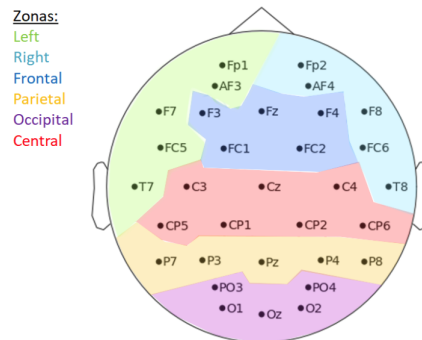


Figura 3.3: Localización de los electrodos en las áreas del cerebro humano

Con la clasificación en áreas de los electrodos se podrán identificar que zonas son las más activas para cada una de las clasificaciones.

3.2 Diseño

En el diseño de la solución se implementan tres modelos basados en aprendizaje automático, *Random Forest*, *Multi Layer Perceptron* y *XGBoost*, dónde se identifican las características más relevantes para la clasificación y se realiza un ajuste de hiperparámetros, para obtener así los hiperparámetros más significativos para la tarea en el reconocimiento de emociones mediante señales de EEG.

3.2.1. Especificación de requerimientos

El proyecto se organiza en diferentes cuadernos para organizar la información de manera adecuada, facilitando el desarrollo del proyecto, su experimentación y su posterior evaluación.

Se utilizan los datos EEG del dataset correspondiente, correspondientes a 32 características. El resto de características disponibles no son relevantes para el presente trabajo.

Los electrodos disponibles para la clasificación son: *Fp1*, *AF3*, *F3*, *F7*, *FC5*, *FC1*, *C3*, *T7*, *CP5*, *CP1*, *P3*, *P7*, *PO3*, *O1*, *Oz*, *Pz*, *Fp2*, *AF4*, *Fz*, *F4*, *F8*, *FC6*, *FC2*, *Cz*, *C4*, *T8*, *CP6*, *CP2*, *P4*, *P8*, *PO4*, *O2*.

Los datos de EEG se discriminan por ancho de banda *theta*, *alpha*, *beta*, *gamma*, obteniendo un total de 128 características.

Originalmente se dispone de cuatro etiquetas con un valor real de *Valencia*, *Arousal*, *Linking*, *Dominance*, dichas etiquetas se utilizan para obtener una etiqueta binaria que identifica la presencia o ausencia de *Valencia* o *Arousal*. Las etiquetas *Linking* y *Dominance* no se consideran en el presente trabajo.

Los datos se dividen en conjuntos de entrenamiento, desarrollo y test.

En primer lugar, los datos de entrenamiento se utilizan para obtener las características más importantes y para la búsqueda de hiperparámetros. En segundo lugar, los de

desarrollo para evaluar los hiperparámetros que mejor funcionan para la tarea. En última instancia, los de test para evaluar el modelo final.

Se definen los modelos de aprendizaje automático *Random Forest*, *Multi Layer Perceptron* y *XGBoost*, por recomendación de los tutores del trabajo. La elección de métodos clásicos viene dada porque en la primeras fases del proyecto, el dataset disponible para realizar los experimentos, no tenía un número alto de muestras. El dataset original es propiedad de la Universitat de Valencia y del equipo de investigación del Instituto de automática y informática industrial de la UPV, por ello son los datos que se utilizan en los primeros experimentos.

Los modelos de clasificación basados en redes neuronales necesitan un número elevado de datos para su entrenamiento en comparación con los métodos clásicos, por ello la importancia de utilizar modelos como *Random Forest* y *XGBoost*.

Para poder comparar el rendimiento de los métodos clásicos con una red neuronal, como recomendación de mis tutores se define una red neuronal de pocas capas, dado que se tienen pocos datos disponibles, con la finalidad de evitar problemas de sobreajuste.

La experimentación precisa de una ordenación inicial de características más significativas para la realización de la experimentación. En los modelos *Random Forest* y *Multi Layer Perceptron* la ordenación de variables más significativas se realiza con un modelo *Random Forest* con una parametrización por defecto. En el caso del modelo *XGBoost* se utiliza un modelo con la misma arquitectura para la selección previa de variables más significativas.

Se realiza una selección de hiperparámetros y el número de características con las que se obtiene el rendimiento más elevado, para así realizar la evaluación final de los modelos.

3.2.2. Metodología

El objetivo principal del proyecto es definir una metodología para la clasificación de emociones con modelos basados en aprendizaje automático y investigar si dicho planteamiento consigue mejorar los resultados obtenidos en otros trabajos similares.

Se implementan diferentes modelos basados en aprendizaje automático, para su entrenamiento y evaluación, para ello se diferencian dos procesos, el primero relacionado con la escritura de la presente memoria y el segundo con la implementación del código, la experimentación y la realización de una evaluación final.

El proyecto tiene una duración total de 360 horas, por ello para lograr el objetivo final se decide utilizar una metodología ágil, en concreto *Scrum*. En el presente proyecto los tutores del trabajo ocupan el rol del *Product Owner (PO)* y *Scrum Master*, en el caso del estudiante ocupa un rol de *Equipo de desarrollo*.

Primero se define un *Backlog* con todas los requisitos definidos en el proyecto. De manera iterativa se van aplicando cambios que modifican dichos requisitos para la obtención del resultado final.

Las tareas están definidas dentro de los *Sprints*. Periódicamente el estudiante informa a los tutores del proyecto de sus avances al fin de cada *Sprint*, se enseñan los resultados obtenidos, definiendo los siguientes pasos a seguir y cambios que se deben de aplicar al trabajo ya desarrollado, para así cumplir los objetivos finales.

Se define el backlog con ocho *Sprints*, dónde en cada uno de ellos se han realizado las tareas correspondientes:

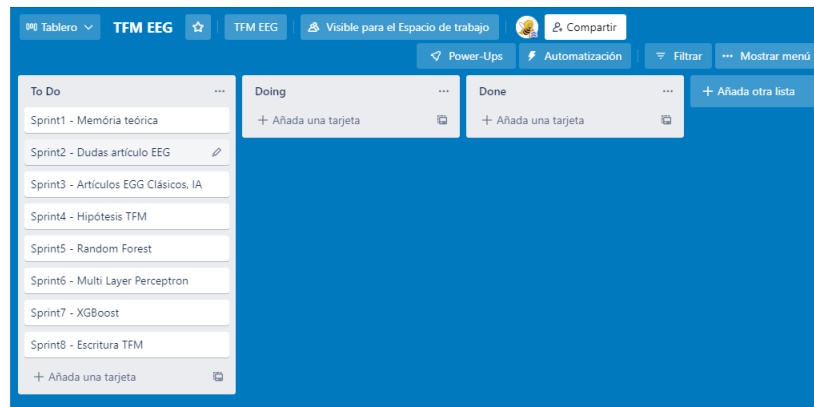


Figura 3.4: Backlog del proyecto

Para cada *Sprint* se ha definido la siguiente correspondencia en el cómputo total de horas disponibles para el trabajo:

Sprint	Duración (h)
1 - Memoria teórica	50
2 - Dudas artículo EEG	50
3 - Artículos EGG Clásicos IA	50
4 - Hipótesis TFM	20
5 - Random Forest	50
6 - Multi Layer Perceptron	50
7 - XGBoost	50
8 - Escritura TFM	40

Tabla 3.3: Duraciones de los *Sprints* para el desarrollo del proyecto

3.3 Herramientas utilizadas

3.3.1. Software

Google Colab

Herramienta en línea de google para la creación de cuadernos de código colaborativos en los que se puede escribir y ejecutar código en python en la nube, sin la necesidad de instalar las librerías más conocidas y utilizadas.

Es una herramienta potente para realizar pruebas de código y comprobaciones de funcionamiento previas de manera rápida y sencilla.

La herramienta ha sido muy útil para diferentes asignaturas de la carrera y para la realización del presente proyecto.

3.3.2. Lenguajes de programación

Python

Lenguaje de programación de alto nivel creado por Guido Van Rossum en el año 1991. El nombre hace referencia al gusto del creador por los *Monty Python*.

Es un lenguaje bastante extendido en el mundo del aprendizaje automático por su corta curva de aprendizaje, por la facilidad en la descarga y uso de librerías.

Al ser un lenguaje interpretado, la depuración de código puede ser sencilla, pero no con la existencia de errores semánticos.

3.3.3. Librerías

NumPy

NumPy es una librería de python que permite crear vectores o matrices para el almacenamiento y procesamiento de datos dentro del campo científico.

El uso de arrays permite estructurar el valor de los píxeles de las imágenes y sus respectivos canales para el tratamiento de imagen mediante el uso de filtros, para la conversión en imágenes *PNG* o en objetos de tipo Tensor.

Pandas

Librería ampliamente utilizada en el campo del aprendizaje automático y la ciencia de datos. Sus herramientas permiten formatear la información en tablas, llamadas (Data-Frames).

El formato *Data Frame* permite realizar operaciones matemáticas a los datos y obtener datos estadísticos de los mismos con facilidad.

Matplotlib

Librería de python para la creación de gráficas utilizada en el proyecto para la visualización de los resultados de evaluación del proyecto.

Dispone de documentación en línea muy extensa y una gran comunidad detrás para resolver dudas.

Scikit-learn

Librería de python utilizada por los modelos de *Random Forest* y *Multi Layer Perceptron*. Dispone de multitud de modelos que engloban problemas de clasificación, regresión, clustering.

Dispone de un catalogo muy amplio de herramientas para aplicar en el preprocesado de datos, sin obviar los datasets disponibles para aprender a utilizar modelos en aprendizaje automático en fases iniciales.

La documentación disponible es muy amplia y tiene una gran comunidad de personas detrás para responder dudas y solucionar problemas de implementación.

3.4 Técnicas de aprendizaje automático utilizadas

En el siguiente punto se enumeran los modelos utilizados para la realización del presente trabajo, justificando su uso en comparativa con el resto de modelos disponibles dentro del ámbito del aprendizaje automático.

En las primeras fases del proyecto se realizó un estudio exhaustivo del estado de la cuestión de la tecnología para la detección de emociones mediante el uso de (EEG) que, sumado al criterio de los tutores del presente trabajo, se llegó a un consenso para centrar el tiempo y el esfuerzo en tres modelos diferentes.

Según la bibliografía existente, el modelo del cual se obtienen los mejores resultados es la aproximación de árboles de clasificación. Concretamente *Random Forest*, pero se prueba una arquitectura más novedosa, un modelo basado en *XGBoost* para experimentar si mejora significativamente los resultados obtenidos.

En último lugar, se utiliza una red neuronal básica: *Multi Layer Perceptron (MLP)*. Esta se utilizará para realizar una comparativa con el resto de métodos clásicos enumerados, sin entrar a valorar arquitecturas más complejas, ya que el presente trabajo se centra en mejorar los resultados obtenidos con modelos basados en árboles de clasificación, planteando una metodología diferente para investigar si mejora los resultados de trabajos similares relacionados.

Por último, uno de los problemas más importantes que se presenta en la tarea de clasificación de emociones, en el uso de *DEAP Dataset* está reflejada en los diferentes artículos relacionados. Cada sujeto para una emoción en particular presenta una señal que puede ser diferente entre diferentes sujetos, por ello en todo intento de generalización de los modelos de clasificación, los resultados de precisión obtenidos por todos los métodos presentan valores de precisión en torno al 62 %.

3.4.1. Random Forest

Los modelos basados en árboles de decisión, concretamente los *Random Forest*, dentro de ámbito de los algoritmos clásicos basados en aprendizaje automático, son los que mayor rendimiento muestran en la tarea de clasificación de emociones como la *Valencia* y el *Arousal* [5].

En el presente trabajo se propone una visión alternativa dónde previamente se aplica el *método de Welch* a los datos, discriminando por tipo de onda y nombre de electrodo, buscando así que las características menos influyentes afecten lo más mínimo a la tarea de clasificación, dando valor a las características más significativas.

Otro de los aportes se basa en definir proceso de selección de hiperparámetros, realizando previamente una ordenación de las características más significativas, dónde se prueba el rendimiento del modelo para cada tamaño de (n) características más significativas.

A efectos comparativos con otros artículos relacionados se opta por intentar obtener un rendimiento mayor con el uso de un modelo *Random Forest*, porque como se indica en el artículo de Theresia Diah [5], el método clásico que obtiene los mejores resultados en el estado de la cuestión es el modelo *Random Forest*.

3.4.2. Multi Layer Perceptron

Seguidamente, se utiliza una red neuronal base *Multi Layer Perceptron (MLP)* para observar si la aproximación propuesta por el presente trabajo ayuda a obtener resultados de

precisión significativos con el uso de una red neuronal totalmente conectada, con pocas capas, en comparación con los otros modelos.

Para poder realizar una comparación entre los modelos propuestos, la ordenación y selección de características más significativas se realiza con un modelo *Random Forest* base.

Como se observa en el artículo de Yucel Cimtay, Erhan Ekmekcioglu [4], proponen utilizar una red neuronal pre entrenada para obtener las características de manera automática.

3.4.3. XGBoost

Por último, después de haber definido un modelo con una aproximación clásica y otra basada en redes neuronales, se prueba un modelo del grupo de árboles de decisión por recomendación de los tutores del presente trabajo. En este caso, un modelo *XGBoost*, que en las tareas observadas en la fase de estudio del estado del arte se concluye que en ciertas tareas de clasificación o regresión consiguen mejorar los resultados obtenidos con *Random Forest* [3].

Utilizar un modelo *XGBoost* nos permite experimentar si dicho modelo, junto a la metodología propuesta en el presente trabajo, mejora los resultados obtenidos por los modelos de *Random Forest* y *Multi Layer Perceptron*.

Dado que es un modelo basado en árboles de decisión el algoritmo permite obtener las variables más significativas en la tarea de clasificación. Asimismo, se sigue el mismo planteamiento de ordenación de variables definido en los modelos anteriores, pero en dicho caso con los métodos disponibles en *XGBoost*.

3.5 Implementación

En el siguiente capítulo se describe la implementación de la solución previamente diseñada, explicando detalladamente cada una de las fases marcadas en los objetivos del presente trabajo.

La implementación engloba, en primera instancia, la organización de la estructura del entorno en *Google Colab*, en el cual se almacena el dataset "*DEAP: A Database for Emotion Analysis using Physiological Signals*".

Una vez terminada la fase organizativa, se implementa el código necesario para el proyecto, estructurado en diferentes cuadernos en la plataforma de Google.

En la implementación primero se recoge la división de los datos en tres conjuntos diferentes, de entrenamiento, validación y test:

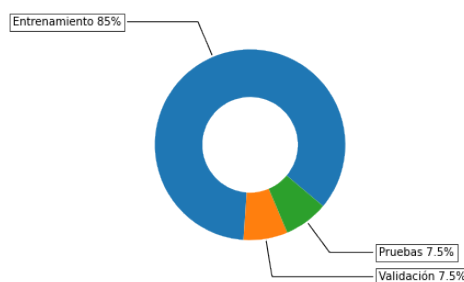


Figura 3.5: División de los datos en conjuntos

La mayoría de los datos engloban el conjunto de entrenamiento, porque como se ha indicado con anterioridad el número de muestras final es de 1280, aplicando la operación de *split* se obtienen el siguiente número de muestras en cada conjunto:

Conjunto	Nº de muestras
Entrenamiento	1088
Validación	96
Pruebas	96

Tabla 3.4: Nº de muestras por conjunto

No se sigue un criterio específico para seleccionar los porcentajes correspondientes, son parámetros por defecto para realizar los experimentos.

El conjunto de entrenamiento se utiliza, en primer lugar, para obtener en orden las características más importantes y para el entrenamiento del modelo para la clasificación, el conjunto de validación para la selección de hiperparámetros y por último, el conjunto de pruebas para medir el rendimiento del modelo correspondiente.

Cuando se acaba con la implementación se ejecutan los cuadernos de manera secuencial, para así realizar las pruebas pertinentes. En primera instancia se obtienen los resultados de precisión para elegir los hiperparámetros que otorgan un mayor rendimiento para la tarea.

Una vez seleccionado el mejor modelo, se realiza la evaluación final, dónde se obtienen los resultados con las métricas *Precisión*, *F1-Score*.

3.5.1. Búsqueda del dataset

Durante las primeras fases del trabajo, teniendo claras las etiquetas a clasificar, concretamente la *Valencia* y *Arousal*, se realiza una búsqueda exhaustiva de un dataset con un gran número de muestras para poder construir los conjuntos de entrenamiento, validación y pruebas.

En primer lugar, se estudia la posibilidad de utilizar un dataset disponible dentro de mi grupo de investigación del instituto de automática y informática industrial de la Universitat Politècnica de València, cedido por la facultad de Psicología de Universitat de València.

Después de la adecuación de los datos, se concluye que no se disponen de los datos suficientes para realizar los experimentos. Por ello, se buscan datos que se hayan utilizado en otros artículos relacionados.

Por último, después de la lectura de bibliografía relacionada, se decide optar por la utilización de "DEAP: A Database for Emotion Analysis using Physiological Signals" [11].

Inicialmente, se intenta contactar con los autores del *Dataset* sin obtener respuesta. Esto demoró bastante las fases iniciales del proyecto, teniendo que contar con la ayuda de mis tutores para realizar la solicitud, obteniendo respuesta semanas después. Después de la mediación de mis tutores se pudo disponer del "DEAP, Dataset". Se empieza a estudiar los datos disponibles en el dataset para su posterior utilización.

3.5.2. Carga del dataset y preprocesamiento de los datos, división de ancho de banda, Método de Welch

El dataset está formado por 32 ficheros con extensión (.dat) dónde cada archivo contiene los datos obtenidos por cada uno de los electrodos, sensores, en un experimento determinado. A cada sujeto se le induce un sentimiento mediante un video, el mismo psicólogo define una etiqueta de *Valencia*, *Arousal*, *Dominance*, *linking* correspondiente. Dichas etiquetas aún no se les ha aplicado un proceso de *One Hot Encoding*, para obtener la etiqueta binarizada. La etiqueta binaria es necesaria para identificar la presencia o ausencia de *Valencia*, *Arousal*.

```
array([[7.71, 7.6 , 6.9 , 7.83],
       [8.1 , 7.31, 7.28, 8.47],
       [8.58, 7.54, 9. , 7.08],
       ...,
       [3.05, 7.01, 5.1 , 1.1 ],
       [3.99, 7.17, 4.85, 1. ],
       [7.15, 4.03, 9. , 1.88]])
```

Figura 3.6: Etiquetas originales DEAP dataset en Numpy Array

Para la lectura de los ficheros se utiliza la librería *CPickle*, recomendada por los autores de "DEAP". Se almacenan los datos de los 32 experimentos y las etiquetas correspondientes en memoria del cuaderno en *Google Colab*.

Una vez leído los archivos se pasa a la fase de preprocesado de la información, fase en la que primero se transforman las etiquetas para centrar la atención en la *Valencia* y el *Arousal*.

Con la media total de los valores para una etiqueta se obtiene el valor binarizado de la presencia o ausencia de *Valencia* o *Arousal*. Si el valor de la etiqueta está por encima de dicho valor umbral, previamente calculado, se considera que la etiqueta se codifica con un uno, si la etiqueta está por debajo del umbral calculado se codifica con un cero en formato *One Hot Encoding*:

```
array([[0., 1.],
       [1., 0.],
       [1., 0.],
       ...,
       [1., 0.],
       [0., 1.],
       [1., 0.]], dtype=float32)
```

Figura 3.7: Formato One Hot Encoding de las etiquetas

Una vez transformadas las etiquetas se comprueba si los datos están desbalanceados, obteniendo los siguientes resultados:

Etiqueta	Datos
Positive Valence	680
Negative Valence	600
High Arousal	640
Low Arousal	640

Tabla 3.5: Resultados balance de las etiquetas en los datos

Se concluye que las clases no están desbalanceadas y se prosigue con el preproceso de los datos. Es importante que los datos estén debidamente balanceados entre clases, para que los modelos puedan generalizar correctamente sin sobre aprender.

El siguiente paso, consiste en discriminar la señal obtenida de los electrodos en los cuatro tipo de ondas *theta*, *alpha*, *beta* y *gamma* mediante el método de *Welch*. El proceso consiste en transformar el valor medido de los electrodos, que se encuentran dentro del dominio temporal al dominio de la frecuencia. Una vez transformados los datos se separan por tipo de onda, con su etiqueta correspondiente.

Al aplicar dicho método se obtiene el periodograma de cada una de las muestras:

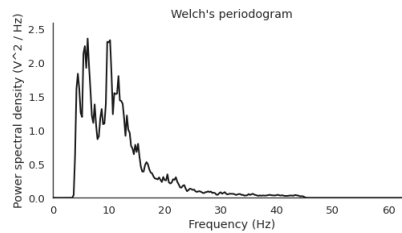


Figura 3.8: Periodograma de una muestra EEG al aplicar Welch

Dichas muestras se discriminan por tipo de onda, obteniendo el valor de la frecuencia en una ventana temporal concreta. Una muestra en el dominio temporal tiene una etiqueta asociada. Al transformar la muestra al dominio de frecuencias se obtiene una cuantificación de las frecuencias más comunes en cada una de las ventanas temporales. Por ello, en cada intervalo de onda se obtiene el valor correspondiente de un electrodo por tipo de onda.

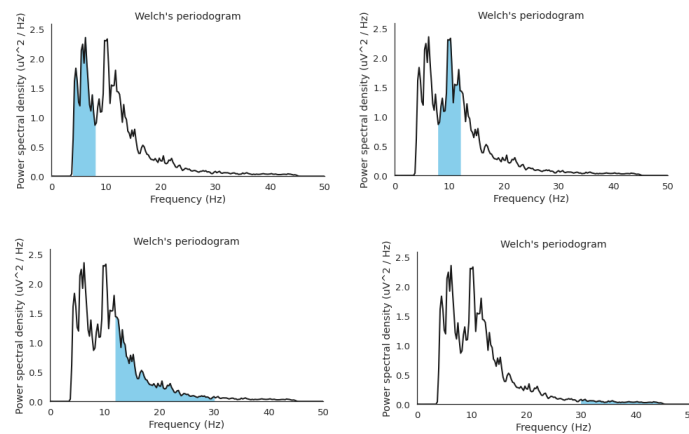


Figura 3.9: Discriminación tipo de onda al aplicar Welch

Inicialmente cada muestra tiene un número total de 32 características. Después de aplicar el proceso de discriminación de tipo de onda, como se han utilizado los cuatro tipo de onda *theta*, *alpha*, *beta* y *gamma*, se obtienen las nuevas 128 características, que identifican a un electrodo particular y el valor obtenido en la muestra para un tipo de onda.

3.5.3. Implementación y experimentación usando un modelo Random Forest

En la implementación del modelo Random Forest se utiliza la librería *Scikit Learn*. Para la realización del experimento, primero se define un modelo inicial para obtener una lista ordenada de las características más importantes. Para ello, se definen unos hiperparámetros aleatorios por defecto:

Parámetro	Valor
max_depth	500
random_state	10

Tabla 3.6: hiperparámetros Random Forest en la selección de características más significativas

Los hiperparámetros definidos en la tabla anterior son para realizar un entrenamiento a un modelo inicial, solo utilizado para la obtención de una lista ordenada, con las características más importantes. Solo es necesario realizar la fase de entrenamiento para la obtención de la lista.

La lista ordenada con las características más importantes se obtiene con la función "feature_importances". En las imágenes siguientes aparecen las variables más significativas para la clasificación de la *Valencia* y *Arousal* y por tipo de onda *Theta*, *Alpha*, *Beta*, y *Gamma*. Se muestran un pequeño número de las variables seleccionadas, siendo el total de 128 electrodos, onda:

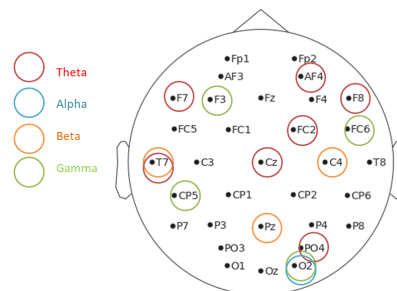


Figura 3.10: Características más importantes Random Forest, Valencia

Al observar los electrodos más significativos en la clasificación de la *Valencia*, el tipo de onda más relevante es la onda *Theta*. La mayoría de electrodos importantes se encuentran en el área *Izquierda, Frontal y Derecha*.

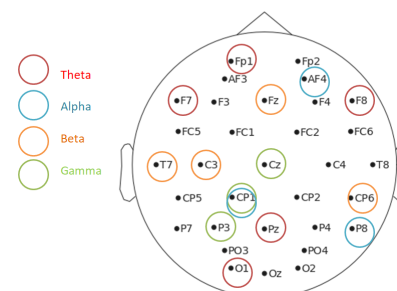


Figura 3.11: Características más importantes Random Forest, Arousal

En la clasificación de *Arousal*, el tipo de onda más relevante de nuevo es la onda *Tetha*, pero los electrodos se localizan en las áreas *Izquierda*, *Central*, *Parietal*.

A efectos prácticos, la lista de características más importantes nos indica los electrodos discriminados por tipo de onda que tienen que utilizarse para el entrenamiento, obviando el resto, aunque en la práctica dependiendo de la tarea se necesitan más o menos variables, dado que las menos significativas pueden aportar ruido al modelo.

Seguidamente se realiza una selección de hiperparámetros iterativa, para obtener los más relevantes. El criterio seguido para la selección de los hiperparámetros del modelo se ha consultado la documentación de la librería *Scikit Learn*, seleccionando los hiperparámetros que más pueden influir en el rendimiento del modelo *Random Forest*.

No se han utilizado más hiperparámetros, dado que no se ha implementado un algoritmo de búsqueda avanzado para la selección, con un algoritmo de búsqueda, se ha implementado un simple bucle iterativo anidado, una búsqueda de fuerza bruta para seleccionar cada uno. La lista de características más significativas también entra dentro del bucle anidado.

En el caso de los hiperparámetros se selecciona un valor existente en una lista, sin embargo para las características más importantes se define una variable iterativa que indica el índice del número de características que se seleccionan.

Para fijar el número total de hiperparámetros se ha utilizado la librería de python "tqdm", que permite medir el tiempo de ejecución estimado del bucle general en la búsqueda de hiperparámetros. Con dicha información se fija en cuatro el número de hiperparámetros para realizar los experimentos en un tiempo admisible, dado que con más, el cuaderno de *Google Colab* nunca finalizaba la ejecución.

Otro de los aspectos importantes después de la selección del número total de hiperparámetros es elegir el rango. Tanto la cantidad de hiperparámetros como el rango total y el contador de iteración influyen en el tiempo total de ejecución, por ello se definen finalmente los siguientes rangos para cada hiper parámetro:

Parámetro	Valores
n_estimators	list(range(100,1000, 100))
criteria	[gini, entropy]
max_depths	list(range(50,1000, 50))
max_features	[sqrt,log2]

Tabla 3.7: Selección de hiperparámetros en Random Forest

A continuación, se obtienen los mejores hiperparámetros para la tarea de clasificación de *Valencia* y *Arousal*.

En primera instancia se obtienen los mejores hiperparámetros para la etiqueta *Valencia*, con un 0.7083 de precisión con el conjunto de validación.

Parámetro	Valor
Datos más significativos	60
n_estimators	600
criteria	gini
max_depths	50
max_features	sqrt

Tabla 3.8: Mejores hiperparámetros para Valencia, Random Forest

En segunda instancia se obtienen los mejores hiperparámetros para la etiqueta *Arousal*, con un 0.6875 de precisión con el conjunto de validación.

Parámetro	Valor
Datos más significativos	110
n_estimators	300
criteria	gini
max_depths	50
max_features	sqrt

Tabla 3.9: Mejores hiperparámetros para Arousal, Random Forest

Una vez implementado, entrenado el modelo se han seleccionado los hiperparámetros relevantes para realizar la evaluación final.

3.5.4. Implementación y experimentación usando un modelo Multi Layer Perceptron

En la implementación del modelo Multi Layer Perceptron (MLP) se utiliza la librería *Scikit Learn*. Para la realización del experimento primero se utiliza un modelo *Random Forest*, como en el paso anterior, dónde se obtiene la lista ordenada de las características más importantes para la clasificación, ya que el modelo MLP no dispone de función para visualizar las características más relevantes.

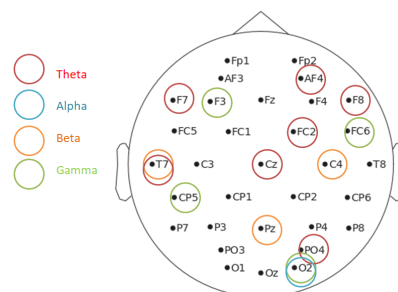


Figura 3.12: Características más importantes MLP, Valencia

Al observar los electrodos más significativos en la clasificación de la *Valencia*, el tipo de onda más relevante es la onda *Theta*. La mayoría de electrodos importantes se encuentran en el área *Izquierda, Frontal y Derecha*.

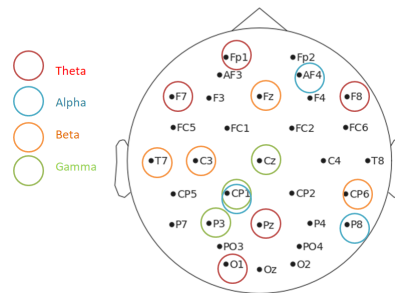


Figura 3.13: Características más importantes MLP, Arousal

En la clasificación de *Arousal*, el tipo de onda más relevante de nuevo es la onda *Theta*, pero los electrodos se localizan en las áreas *Izquierda, Central, Parietal*.

A continuación, se seleccionan los hiperparámetros relevantes para realizar una experimentación de selección de hiperparámetros más significativos, por ello se especifican los siguientes los siguientes rangos:

Parámetro	Valores
Normalization	[True, False]
Size Layer	[range(10, 1000, 100)]

Tabla 3.10: Selección de hiperparámetros en MLP

Una vez realizada la ejecución iterativa, se obtienen los mejores hiperparámetros para la tarea de clasificación de *Valencia y Arousal*.

En primera instancia, se obtienen los mejores hiperparámetros para la etiqueta *Valencia*, con un 0.6563 de precisión con el conjunto de validación.

Parámetro	Valor
Datos más significativos	120
Normalization	False
Size Layer	400

Tabla 3.11: Mejores hiperparámetros para Valencia, MLP

En segunda instancia, se obtienen los mejores hiperparámetros para la etiqueta *Arousal*, con un 0.6771 de precisión con el conjunto de validación.

Parámetro	Valor
Datos más significativos	60
Normalization	False
Size Layer	200

Tabla 3.12: Mejores hiperparámetros para Arousal, MLP

Una vez implementado, entrenado el modelo se han seleccionado los hiperparámetros relevantes para realizar la evaluación final.

3.5.5. Implementación y experimentación usando un modelo XGBoost

A diferencia de los dos modelos anteriores, para la implementación del modelo *XGBoost* se utiliza la librería *xgboost*, dado que *Scikit Learn* no tiene disponible el modelo en sus librerías.

Primero, se define un modelo inicial para obtener una lista ordenada de las características más importantes, como se ha realizado con el modelo *Random Forest*, utilizando un modelo *XGBoost*, dado que estos modelos también se basan en árboles de clasificación, pudiendo obtener las variables más significativas para la clasificación. Los hiperparámetros del modelo son por defecto en la librería.

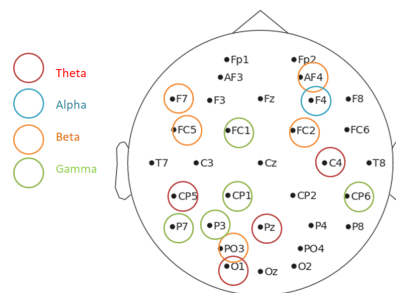


Figura 3.14: Características más importantes XGBoost, Valencia

Al observar los electrodos más significativos en la clasificación de la *Valencia*, el tipo de onda más relevante es la onda *Beta*. La mayoría de electrodos importantes se encuentran en el área *Parietal*, *Occipital*.

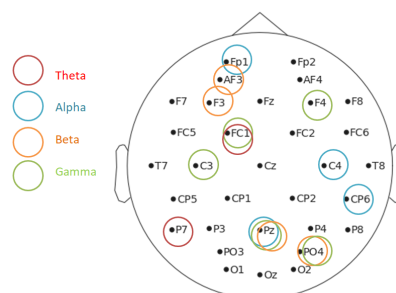


Figura 3.15: Características más importantes XGBoost, Arousal

En la clasificación de *Arousal*, el tipo de onda más relevante es la onda *Alpha*, pero los electrodos se localizan en las áreas *izquierda*, *Parietal*, *Occipital*.

A continuación, se seleccionan los hiperparámetros relevantes para realizar una experimentación de selección de los más significativos, por ello se eligen los siguientes con unos rangos definidos para poder realizar la experimentación en un tiempo razonable.

Parámetro	Valores
max_depth	range(1, 500, 100)
lambda	range(1,5)
alpha	range(0,5)
Learning_rate	[1e-2,1e-3,1e-4,1e-5]

Tabla 3.13: Selección de hiperparámetros en XGBoost

Seguidamente, se comprueba para la tarea en cuestión, que hiperparámetros son los que mejor resultado dan y cuántas mejores características se deben de utilizar en el experimento.

Una vez realizada la ejecución iterativa se obtienen los mejores hiperparámetros para la tarea de clasificación de *Valencia y Arousal*.

En primera instancia se obtienen los mejores hiperparámetros para la etiqueta *Valencia*, con un 0.7188 de precisión con el conjunto de validación.

Parámetro	Valor
Datos más significativos	70
max_depth	100
lambda	1
alpha	4
Learning_rate	1e-2

Tabla 3.14: Mejores hiperparámetros para Valencia, XGBoost

En segunda instancia se obtienen los mejores hiperparámetros para la etiqueta *Arousal*, con un 0.6771 de precisión con el conjunto de validación.

Parámetro	Valor
Datos más significativos	20
max_depth	100
lambda	1
alpha	2
Learning_rate	1e-2

Tabla 3.15: Mejores hiperparámetros para Arousal, XGBoost

Una vez implementado, entrenado el modelo se han seleccionado los hiperparámetros relevantes para realizar la evaluación final.

En la evaluación final, después de la obtención de los hiperparámetros que mejor rendimiento aportan al modelo, se entrena de nuevo con el conjunto de entrenamiento, pero a diferencia de la selección de hiperparámetros, se utiliza el conjunto de pruebas para evaluar.

CAPÍTULO 4

Resultados

Una vez definidos los hiperparámetros más significativos en la fase de implementación y experimentación, se procede a realizar la evaluación final con los hiperparámetros seleccionados y el conjunto de características más significativas, con los que hemos obtenido los mejores resultados.

Una vez entrenado los modelos, se evalúa con el conjunto de pruebas y se obtiene el *Classification Report*, que utiliza las métricas *Precisión*, *Recall*, *F1-Score*. Para una mayor claridad con los datos se formatea la información de la tabla en una gráfica *Heat map*.

Por último, se recoge en una gráfica el resumen de la precisión obtenida por los modelos del presente trabajo, *Random Forest*, *MLP* y *XGBoost* para la clasificación de las etiquetas *Valencia* y *Arousal*.

4.1 Evaluación modelo Random Forest

En primer lugar, se obtienen los resultados para la clasificación de *Valencia* con el mejor modelo *Random Forest* obtenido:

	precision	recall	f1-score	support
0	0.59	0.59	0.59	46
1	0.62	0.62	0.62	50
accuracy			0.60	96
macro avg	0.60	0.60	0.60	96
weighted avg	0.60	0.60	0.60	96

Figura 4.1: Resultados Random Forest, Valencia

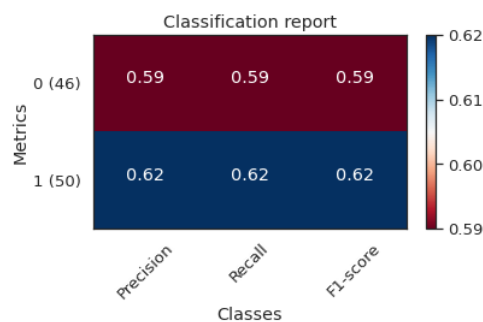


Figura 4.2: Resultados en gráfica, Random Forest, Valencia

Se evalúa el mejor modelo obtenido para la clasificación de *Arousal*:

	precision	recall	f1-score	support
0	0.65	0.71	0.68	49
1	0.67	0.60	0.63	47
accuracy			0.66	96
macro avg	0.66	0.66	0.65	96
weighted avg	0.66	0.66	0.65	96

Figura 4.3: Resultados Random Forest, Arousal

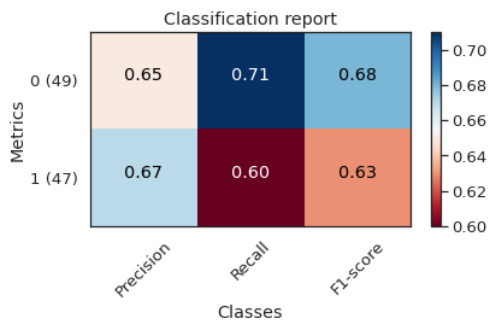


Figura 4.4: Resultados en gráfica, Random Forest, Arousal

4.2 Evaluación modelo Multi Layer Perceptron

En segundo lugar, se obtienen los resultados para la clasificación de *Valencia* con el mejor modelo *MLP* obtenido:

	precision	recall	f1-score	support
0	0.47	0.35	0.40	46
1	0.52	0.64	0.57	50
accuracy			0.50	96
macro avg	0.49	0.49	0.49	96
weighted avg	0.49	0.50	0.49	96

Figura 4.5: Resultados MLP, Valencia

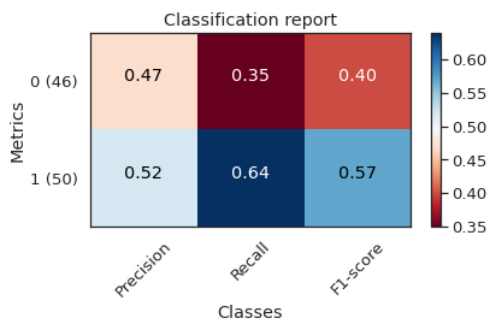


Figura 4.6: Resultados en gráfica, MLP, Valencia

Se evalúa el mejor modelo obtenido para la clasificación de *Arousal*:

	precision	recall	f1-score	support
0	0.51	0.70	0.59	46
1	0.58	0.38	0.46	50
accuracy			0.53	96
macro avg	0.54	0.54	0.52	96
weighted avg	0.54	0.53	0.52	96

Figura 4.7: Resultados MLP, Arousal

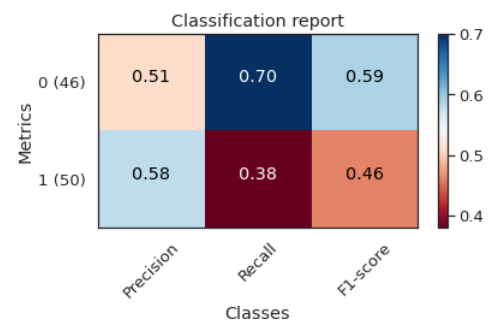


Figura 4.8: Resultados en gráfica, MLP, Arousal

4.3 Evaluación modelo XGBoost

En tercer lugar, se obtienen los resultados para la clasificación de *Valencia* con el mejor modelo *XGBoost* obtenido:

	precision	recall	f1-score	support
0	0.64	0.61	0.62	49
1	0.61	0.64	0.62	47
accuracy			0.62	96
macro avg	0.63	0.63	0.62	96
weighted avg	0.63	0.62	0.62	96

Figura 4.9: Resultados XGBoost, Valencia



Figura 4.10: Resultados en gráfica, XGBoost, Valencia

Se evalúa el mejor modelo obtenido para la clasificación de *Arousal*:

	precision	recall	f1-score	support
0	0.61	0.61	0.61	49
1	0.60	0.60	0.60	47
accuracy			0.60	96
macro avg	0.60	0.60	0.60	96
weighted avg	0.60	0.60	0.60	96

Figura 4.11: Resultados XGBoost, Arousal

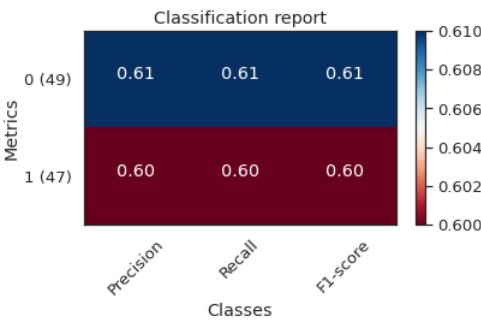


Figura 4.12: Resultados en gráfica, XGBoost, Arousal

En último lugar, se obtiene una gráfica global de los resultados obtenidos por los tres modelos:

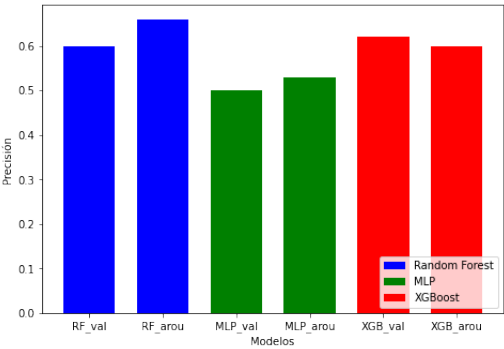


Figura 4.13: Gráfica resumen resultados evaluación modelos

CAPÍTULO 5

Discusión y conclusiones

5.1 Crítica y discusión de los resultados

El método propuesto en el presente trabajo mejora el rendimiento del modelo de tipo *Random Forest* en la tarea de clasificación de la etiqueta *Arousal* respecto al estado de la cuestión, con una precisión del 66 %. Sin embargo en la tarea de clasificación de la etiqueta *Valencia* no se consigue mejorar los resultados, dónde se obtiene un 60 % de precisión. El resultado con un método similar comparable está recogido en el artículo de A.Faqih [5] con una precisión del 62 % tanto para la clasificación de la etiqueta *Valencia* como para la clasificación de la etiqueta *Arousal*:

Modelo	Precisión
Random Forest, Nuestra aproximación, Arousal	66 %
T.D. Kusumaningrum [5] Valencia, Arousal	62 %
Random Forest, Nuestra aproximación, Valencia	60 %

Tabla 5.1: Comparativa resultados Random Forest en la tarea de clasificación de la etiqueta Valencia y Arousal

Si se sigue la misma metodología propuesta con el modelo de tipo *Multi Layer Perceptron* (MLP) no se mejora la precisión respecto al estado de la cuestión, en la tarea de clasificación de la etiqueta *Valencia* con una precisión del 50 % y de la etiqueta *Arousal* con una precisión del 53 %:

Modelo	Precisión
Al-Nafjan [6], Valencia, Arousal	82 %
Yucel Cimtay[4], Valencia, Arousal	72 %
MLP, Nuestra aproximación, Arousal	53 %
MLP, Nuestra aproximación, Valencia	50 %

Tabla 5.2: Comparativa resultados Multi Layer Perceptron (MLP) en la tarea de clasificación de la etiqueta Valencia y Arousal

Por último, si se sigue la misma metodología propuesta con el modelo de tipo *XGBoost* no se mejora la precisión respecto al estado de la cuestión, en la tarea de clasificación de la etiqueta *Valencia* con una precisión del 62 % y de la etiqueta *Arousal* con una precisión del 60 %:

Modelo	Precisión
T.D. Kusumaningrum, Valencia, Arousal [5]	62 %
XGBoost, Nuestra aproximación, Valencia	62 %
XGBoost, Nuestra aproximación, Arousal	60 %

Tabla 5.3: Comparativa resultados XGBoost en la tarea de clasificación de la etiqueta Valencia y Arousal

La discriminación de los electrodos por tipo de onda y la selección de las (n) características más importantes que entran como entrada del modelo demuestran ser una metodología eficaz en conjunción con los modelos *Random Forest* para la clasificación de la etiqueta *Arousal*, pero no para la *Valencia*.

Uno de los principales problemas que tienen los datos para la obtención de resultados más significativos, como recoge T.D. Kusumaningrum en su artículo [5] es que las señales cerebrales generadas por las neuronas en el cerebro no son muy discriminativas para los modelos con aproximación clásica, tienen unos valores que van unidos a un sujeto concreto, que identifican a una emoción concreta pero puede ser algo diferente entre sujetos, mostrando un poder reducido de generalización en los modelos.

El modelo *XGBoost* se utiliza para estudiar su rendimiento para la tarea de clasificación de las etiquetas de *Valencia* y *Arousal* en comparación con la aproximación clásica propuesta basada *Random Forest*, que demuestra que con el planteamiento seleccionado no mejora los resultados anteriormente obtenidos.

Los resultados obtenidos con el modelo *Multi Layer Perceptron (MLP)* en base al planteamiento inicial no han sido satisfactorios en comparación con los resultados del artículo de Yucel Cimtay [4] y de Abeer Al-Nafjan [6] con un 72.81 % y un 82 % de precisión respectivamente.

En la aproximación de Yucel, se utiliza una red neuronal de la que se obtienen las características de una señal automáticamente, obteniendo resultados que mejoran con holgura los métodos clásicos.

En la aproximación de Abeer Al-Nafjan, se generan nuevas características con los datos existentes, que otorgan un gran rendimiento y poder de generalización que el modelo propuesto en el presente trabajo y otras arquitecturas de red neuronal.

El uso de una red neuronal para la obtención de características demuestra ser un mecanismo eficaz que ayuda a eliminar el factor humano en la selección de características, pero en contraposición convierte el modelo en una caja negra, por ello utilizar modelos basados en árboles de clasificación es una aproximación más abierta y ética a la hora de explicar el funcionamiento del modelo y el criterio de clasificación seguido.

5.2 Principales aportaciones

El estado de la cuestión en la aproximación *Random forest* promueve la creación de nuevas características de entrada que otorgan al modelo una mayor capacidad de generalización en un planteamiento clásico, utilizan el método de cuantificación *Fast Fourier transform* para transformar los datos del dominio temporal al de frecuencias.

En primer lugar, en el presente trabajo sólo consideramos variables de electroencefalograma (EEG), aplicando el método de cuantificación de *Welch*, dónde se discrimina la señal por electrodo y tipo de onda, obteniendo 128 características totales como entrada al modelo.

En segundo lugar, la selección de la etiqueta binaria de cada muestra *Valencia, Arousal* se obtiene calculando la media total de los datos disponibles mediante la siguiente condición:

$$\begin{aligned} \text{si } \hat{y}_i &\geq \text{media}_{total} \rightarrow \text{High (Valencia, Arousal)} \\ \text{sino} &\rightarrow \text{Low (Valencia, Arousal)} \end{aligned}$$

En tercer lugar, se define una selección de hiperparámetros, *n° de estimadores, criterions, máx profundidad y máx características*.

En cuarto lugar, se utiliza el modelo de *Random Forest* para ordenar las variables que más influyen en la discriminación del clasificador. El número de (n) datos más significativos entra en el proceso de selección de hiperparámetros.

5.3 Trabajos futuros

El estado de la cuestión marca los pasos a seguir en futuras publicaciones, dónde se identifican dos fases bien diferenciadas, la metodología en la extracción de características y el modelo clasificador a utilizar.

El uso de una red neuronal para la obtención de características demuestra ser un mecanismo eficaz que ayuda a eliminar el factor humano en la selección de características, otorga mayor poder de generalización en la clasificación, pero implica tener un clasificador *caja negra*, en el cual no se puede dar una explicación clara de las decisiones tomadas en la clasificación.

Otro de los puntos fuertes es la limpieza de los datos, la eliminación automática de los *artefactos* puede ser beneficioso para limpiar los datos con mayor profundidad, obteniendo datos con el menor número de ruido posible que puedan afectar al entrenamiento y clasificación de los modelos.

Aunque los modelos basados en redes neuronales demuestran tener un gran rendimiento, utilizar modelos basados en métodos clásicos como *Random Forest* y *XGBoost* són una elección acertada, para así profundizar la búsqueda de características y seleccionar las que otorgan un mayor rendimiento de clasificación al modelo.

Seleccionar una arquitectura de red neuronal funcional para la detección de emociones también es un reto interesante, ya que al tener datos en el *dominio de tiempo* entran en escena modelos convolucionales o basados en transformers y en el *dominio de frecuencias* los modelos de clasificación como *Fully Connected, Redes Recurrentes, etc.*

Por último, la obtención de los datos de electroencefalograma (EEG) actual provoca la obtención de ruido en los datos, por ello investigar y utilizar métodos de obtención intracerebral favorecerá a prevenir el ruido provocado por el cuello cabelludo existente entre el cerebro humano y los electrodos de medición.

Bibliografía

- [1] Leo Breiman. Random Forests *Statistics Department, University of California Berkeley, CA 94720*, January 2001
- [2] Whole slide imaging in pathology: advantages, limitations, and emerging perspectives Farahani N, Parwani A, Pantanowitz L *Pathology and Laboratory Medicine International*, 11 June 2015 Volume 2015:7 Pages 23—33.
- [3] XGBoost: A Scalable Tree Boosting System Tianqi Chen, Carlos Guestrin, *University of Washington*, 10 Jun 2016.
- [4] Investigating the Use of Pretrained Convolutional Neural Network on Cross-Subject and Cross-Dataset EEG Emotion Recognition Yucel Cimtay, Erhan Ekmekcioglu, *Institute for Digital Technologies, Loughborough University London, London E20 3BS, UK*, 4 April 2020.
- [5] Emotion Recognition Based on DEAP Database using EEGTime-Frequency Features and Machine Learning Methods T.D. Kusumaningrum, A.Faqih, B.Kusumoputro *Dept of Electrical Engineering Universitas Indonesia*, International Conference on Science Technology (ICoST 2019).
- [6] Classification of Human Emotions from Electroencephalogram (EEG) Signal using Deep Neural Network Abeer Al-Nafjan, Areej Al-Wabil, Manar Hosny, Yousef Al-Ohali (*IJACSA*) *International Journal of Advanced Computer Science and Applications*, Vol. 8, No. 9, 2017
- [7] A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on danish commons Thorvald Sørensen *The Royal Danish Sciences Society Biological Writings*, Volume V, No. 4, 1948.
- [8] Lee Raymond Dice. *The biotic provinces of North America*. University of Michigan Press, viii + 78. 1943.
- [9] Marvin Minsky, Seymour A. Papert. *Perceptrons: An Introduction to Computational Geometry*. MIT, Massachusetts Institute of Technology. 1969.
- [10] Fionn Murtagh. *Multilayer perceptrons for classification and regression*. Neurocomputing, Volume 2, Issues 5–6, July 1991, Pages 183-197
- [11] S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, *DEAP: A Database for Emotion Analysis using Physiological Signals*. IEEE Transactions on Affective Computing, Special Issue on Naturalistic Affect Resources for System Building and Evaluation, in press

