



Universidad de Concepción
Facultad de Ingeniería
Ingeniería Civil Informática

**Modelo de clasificación etaria basado en conectividad funcional y estructural usando base de
datos HCP de envejecimiento**

Memoria de Título

Autor:	Bruno Andrés Arce Rodríguez
Profesora Patrocinante:	Cecilia Paola Hernández Rivas
Profesoras Co-Patrocinantes:	Pamela Beatriz Guevara Alvez Mónica Otero Ferreiro

Chile, Región del Biobío, Concepción
2 de septiembre de 2025

Agradecimientos

Quiero expresar mi más sincero agradecimiento a todas las personas que hicieron posible la realización de este trabajo. A mi familia, quienes estuvieron conmigo y me animaron durante todo el proceso. A mis perros Ronco y Blanca que me acompañaron durante muchas horas de trabajo. A mis compañeros y amigos, que aliviaron el peso del trabajo con su compañía y que me escucharon hablar sobre cualquiera fuese el problema de turno. A mi profesora guía, Cecilia Hernández por su constante apoyo, consejo y curiosidad sin las que este trabajo, el póster y publicación asociadas no hubiesen existido. A las profesoras co-patrocinantes Pamela Guevara y Mónica Otero, por sus observaciones precisas y el tiempo invertido en mejorar este trabajo. A Catalina Vergara y al grupo de imágenes médicas por brindarme los materiales para este trabajo y acogerme. A los proyectos FONDECYT 1221665, FONDECYT 11240971, FONDECYT 11241484. Y a todos quienes de alguna u otra forma estuvieron presentes.

A todos ustedes, gracias.

Resumen

Comprender cómo cambia el cerebro normal con la edad se ha vuelto cada vez más importante, ya que las personas con desviaciones significativas son susceptibles a posibles trastornos cerebrales. Aunque existen diversos factores que influyen en el desarrollo de la edad cerebral, existen varios métodos utilizan diferentes acercamientos para la predicción de la edad. Este trabajo propone un enfoque de aprendizaje automático que utiliza características de conectividad basadas en grafos, extraídas de redes de conectividad estructural y funcional de 606 sujetos sanos de la base de datos del Human Connectome Project - Aging (HCP-A). Nuestros resultados muestran que, al combinar un conjunto reducido de características relevantes de la conectividad estructural y funcional con modelos de Support Vector Machine (SVM) y Logistic Regression (LR), es posible alcanzar un F1-Score superior a 0,75, resultado que se encuentra en el rango reportado por otros trabajos relacionados. Además, utilizando el mismo conjunto reducido de características, un regresor con regularizador de LASSO puede alcanzar un MAE de 7,51.

Índice general

1. Introducción y problema a abordar	1
1.1. Contexto general	1
1.2. Presentación del problema	2
2. Objetivos del proyecto	5
2.1. Objetivo General	5
2.2. Objetivos Específicos	5
2.3. Alcances y limitaciones	5
2.3.1. Análisis de riesgos	6
2.3.2. Restricciones	6
2.3.3. Costos y beneficios	7
3. Marco teórico	8
3.1. Resonancia magnética	8
3.1.1. Resonancia magnética por difusión	8
3.1.2. Resonancia magnética funcional	9
3.2. Grafos	9
3.3. Machine Learning	12
3.3.1. Algoritmos de Machine Learning	12
3.3.2. Evaluación de modelos de ML	13
3.4. Estado del arte	14
4. Desarrollo del proyecto	17
4.1. Materiales	17
4.1.1. Base de datos Human Connectome Project - Aging	17
4.1.2. Atlases de referencia	18
4.1.3. Herramientas existentes para el procesamiento de imágenes	18
4.2. Métodos	18
4.2.1. Procesamiento de imágenes de difusión	18

4.2.2.	Procesamiento de rs-fMRI	20
4.2.3.	Exploración de los datos	21
4.2.4.	Grafos y métricas de conectividad	25
4.2.5.	Umbralización en grafos	26
4.2.6.	Entrenamiento	27
4.2.7.	Análisis posteriores y re-entrenamiento	28
4.2.8.	Aproximaciones al problema de regresión	29
5.	Evaluación experimental	30
5.1.	Configuración de los modelos	30
5.2.	Modelos de única representación	31
5.3.	Modelos de representación combinada	32
5.3.1.	Modelos basados en RF	33
5.3.2.	Modelos basados en LR	33
5.3.3.	Modelos basados en SVM	35
5.3.4.	Resumen de los resultados	37
5.4.	Reducción de características	40
5.5.	Re-entrenamiento de modelos	40
5.5.1.	Re-entrenamiento de LR	42
5.5.2.	Re-entrenamiento de SVM	42
5.5.3.	Resumen de los resultados luego de la reducción	42
5.6.	Regresores	43
6.	Conclusiones	47

Índice de figuras

4.1. Imagen estructural T1 utilizada en el preprocesamiento de las imágenes de difusión.	19
4.2. Distribución de edades para los 606 sujetos.	21
4.3. Distribución de rangos etarios HCP-A.	22
4.4. Distribución de rangos etarios equidistantes.	23
4.5. Distribución de rangos etarios balanceados.	23
4.6. Distribución de valores estructurales con escala logarítmica en el eje Y.	24
4.7. Cambio en la distribución de los valores de conectividad funcional.	25
4.8. Diagrama completo del pipeline utilizado. a) Recolección de datos y procesamiento de estos hasta obtención de métricas basadas en conectividad. b) Entrenamiento de los modelos de clasificación y regresión.	27
5.1. F1-score para los modelos de única representación a través de diferentes umbrales en las características.	32
5.2. F1-score para los modelos de representación combinada basados en Random Forest a través de diferentes umbrales estructurales.	33
5.2. F1-score para los modelos de representación combinada basados en Random Forest a través de diferentes umbrales estructurales.	34
5.3. F1-score para los modelos de representación combinada basados en Logistic Regression a través de diferentes umbrales estructurales.	35
5.3. F1-score para los modelos de representación combinada basados en Logistic Regression a través de diferentes umbrales estructurales.	36
5.4. F1-score para los modelos de representación combinada basados en Support Vector Machine a través de diferentes umbrales estructurales.	37
5.4. F1-score para los modelos de representación combinada basados en Support Vector Machine a través de diferentes umbrales estructurales.	38

5.5. Puntuaciones F1 para modelos SVM, RF y LR con diferentes umbrales aplicados a las características estructurales y funcionales. La columna izquierda muestra las puntuaciones F1 utilizando solo 34 características estructurales o 34 funcionales. La columna derecha muestra las puntuaciones F1 con las 68 características combinadas.	39
5.6. 20 características más relevantes en promedio para diferentes algoritmos, ordenadas por importancia.	41
5.7. Mejores resultados de modelos reentrenados basados en LR.	42
5.8. Mejores resultados de modelos reentrenados basados en SVM.	43
5.9. Resultados de modelos de regresión.	44
5.10. Mejores 15 características	45
5.11. Comportamiento de las 3 métricas más importantes utilizadas por los regresores, separadas por género.	46

Índice de tablas

3.1. Resumen de trabajos relacionados.	16
4.1. Relación entre las medidas de conectividad y las características extraídas.	26
5.1. Configuración de los resultados de la Tabla 5.2 en su orden respectivo.	38
5.2. Mejores resultados en modelos tipo 2 (combinación de características estructurales y funcionales) desagregados por métrica de evaluación y clases. Se destacan en negrita los mejores resultados de cada modelo en cada rango.	40
5.3. Resultados de los modelos SVM y LR utilizando diferentes tamaños en los conjuntos de características, para distintos rangos etarios.	43

Capítulo 1

Introducción y problema a abordar

1.1. Contexto general

El envejecimiento del ser humano, naturalmente, conlleva cambios tanto físicos como cognitivos, impactando de manera gradual pero sostenida el estilo de vida de las personas [1]. A diferencia de los cambios físicos que son notorios y observables, como las arrugas en la piel, pérdida de masa muscular o el encanecimiento del cabello, los cambios cognitivos provienen del envejecimiento cerebral, y dada la naturaleza del órgano no son observables. Dentro del espectro de cambios cognitivos se pueden destacar la disminución en la capacidad de concentración, memoria, procesamiento y una mayor resistencia al aprendizaje [2, 3].

Sin embargo, el paso del tiempo no es el único causante de los cambios en el cerebro. Diversos factores pueden alterar el envejecimiento de este órgano, produciendo una discrepancia relevante entre la edad cerebral y la edad cronológica del individuo. Esta brecha es referida en la literatura como brain age gap, y existe evidencia que vincula la existencia de grandes brechas con condiciones como la diabetes tipo II [4] o un nacimiento prematuro [5].

La resonancia magnética (MRI) es una técnica de imagenología médica no invasiva que utiliza campos magnéticos y ondas de radiofrecuencia para obtener imágenes detalladas del cerebro. Existen distintas variantes de MRI, cada una enfocada en distintos aspectos: las imágenes anatómicas permiten caracterizar propiedades como el espesor cortical o el área superficial; la resonancia de difusión (dMRI) entrega información sobre las fibras de la materia blanca midiendo la difusión de moléculas de agua; y la resonancia funcional (fMRI), particularmente en estado de reposo (rs-fMRI), mide la señal dependiente del nivel de oxígeno en sangre para inferir la actividad e interacciones entre regiones cerebrales.

En particular, la dMRI y la fMRI facilitan el estudio de la conectividad cerebral estructural

y funcional, respectivamente [6, 7]. Estas aproximaciones han permitido analizar de manera más precisa cómo se organizan y comunican las diferentes regiones cerebrales, y han contribuido a entender condiciones en las que el envejecimiento cerebral se ve alterado, como el trastorno obsesivo compulsivo (OCD) [8] o el Alzheimer [9].

Estos datos resultan de gran utilidad para el estudio de los cambios cerebrales y las diferentes condiciones neurológicas o psiquiátricas, como en el caso de la esquizofrenia [10], para la cual se ha detectado un incremento de las conexiones funcionales en la red cerebral en estado de reposo y cambios en la homogeneidad de la red, medida que permite evaluar la cohesión dentro de una red funcional específica [11]. De este modo, es posible realizar tratamientos e intervenciones adecuadas para prevenir o mejorar la salud del individuo.

1.2. Presentación del problema

Resulta entonces de particular interés estudiar los cambios que se producen en el cerebro, tanto en su estructura como en su funcionalidad, a lo largo de la vida, para obtener una mejor comprensión de las causas y efectos del *brain age gap*. Particularmente, la identificación prematura de una edad cerebral muy discordante con la edad cronológica puede levantar alertas respecto al riesgo de padecer ciertas afecciones no detectadas [12].

Diversos estudios han abordado el problema de la estimación de la edad cerebral [13], existiendo dos enfoques principales: aquellos centrados en el análisis directo de las imágenes de resonancia magnética (métodos basados en píxeles, vóxeles o superficies), y aquellos que se centran en el desarrollo de algoritmos de Machine Learning (ML) utilizando tanto técnicas tradicionales como Deep Learning.

Este último enfoque ha ganado popularidad en los últimos 10 años, existiendo estudios con diversas técnicas y enfoques los que a su vez presentan diferentes resultados y conclusiones. Sin embargo, una comparativa entre ellos no es del todo justa, pues las bases de datos utilizadas para el entrenamiento de los modelos varían demasiado entre estudios. Esta diversidad en el origen de los datos es una de las desventajas principales de estos métodos, pues estos modelos son altamente dependientes de la calidad y cantidad de datos. En su mayoría, los datos utilizados provienen de *datasets*, como puede ser el de *Alzheimer's Disease Neuroimaging Initiative* (ADNI) [14] o el

International Consortium of Brain Mapping (ICBM), entre otros [13]. Si bien estos datos contemplan pacientes sanos, no suelen representar ampliamente a la población, lo que afecta directamente la capacidad de generalización de los modelos. Se añade a esto la interpretabilidad limitada que ofrecen algunos modelos, particularmente aquellos que utilizan Deep Learning. Esto dificulta la observación y análisis posterior de los patrones encontrados. Por último, no se puede pasar por alto el gran costo computacional asociado al procesamiento de datos y entrenamiento de estos modelos, lo cual puede derivar en una disminución de la practicidad y aplicabilidad en contextos reales. [13].

Estudios recientes proponen modelos que usan alguno de los tres tipos de imágenes (MRI estructural de origen anatómico [15, 16], dMRI [17] o fMRI [18, 19]) para encontrar patrones o relaciones complejas de explorar en el conectoma humano. Para esto se basan en la extracción de características distintivas de cada una de las representaciones. Sin embargo, la combinación de características provenientes de distintos tipos de imágenes ha sido menos estudiada, aunque existen enfoques que ya han explorado la aplicación de multimétricas obtenidas a través de registros multimodales [20].

Aunque existen múltiples factores que explican la variabilidad entre cerebros humanos, distintos trabajos coinciden en que el cerebro se degrada con la edad [21, 13]. Comprender cómo el cerebro cambia en relación con la edad cronológica es un desafío fundamental en neurociencia, ya que resulta clave para diferenciar el envejecimiento natural de condiciones patológicas como el Alzheimer y otros trastornos neurológicos [22, 13].

En este contexto, varios trabajos han propuesto estimar la edad cerebral a partir de diferentes tipos de características y metodologías. Por ejemplo, Damoiseaux [23] analiza los cambios en la conectividad funcional y estructural (sustancia blanca) asociados a la edad. Este estudio integra hallazgos que muestran cómo la desconexión estructural y la integración funcional alterada contribuyen al envejecimiento cognitivo, destacando la relación entre la conectividad estructural de la sustancia blanca y la organización funcional de redes cerebrales.

Recientes investigaciones han desarrollado modelos de machine learning para predecir la edad cerebral a partir de imágenes rs-fMRI, así como utilizando modalidades complementarias como resonancia estructural, funcional y de difusión [24], e incluso señales EEG o análisis estadísticos de series temporales en regiones cerebrales específicas [25]. Entre las estrategias empleadas, destaca la aplicación de umbralización en matrices de conectividad estructural para seleccionar las conexiones más relevantes [26]. Por ejemplo, Han et al. [27] propusieron un modelo basado en conectividad

funcional que extrae 286 métricas derivadas de estructuras comunitarias en la red cerebral, utilizando correlación de Pearson para eliminar características irrelevantes, aunque sin especificar cuáles fueron finalmente retenidas.

En particular, este trabajo se focaliza en extender el enfoque centrado en la conectividad de la red, considerando únicamente características basadas en medidas de grafos y centrándose en la inclusión de medidas que capturen información local y global de la conectividad cerebral. Estas características se utilizarán para entrenar modelos de machine learning clásicos, adoptando un enfoque ya probado en la literatura, para así lograr clasificar a los individuos en tres distintos grupos etarios. Para la distribución de grupos etarios se considerará el propuesto por HCP-A y algunas variantes.

Capítulo 2

Objetivos del proyecto

2.1. Objetivo General

Proponer un modelo de clasificación mediante Machine Learning para detectar diferentes grupos etarios basado en patrones de conectividad funcional y estructural de los sujetos disponibles en la base de datos *Human Connectome Project* (HCP).

2.2. Objetivos Específicos

1. Aplicar procesamiento de imágenes para obtener representaciones cerebrales de los sujetos extrayendo matrices de conectividad funcional y estructural para cada sujeto. Definir grupos etarios obteniendo conjuntos divididos de manera uniforme por edad.
2. Entrenar modelos de clasificación, dividiendo los datos en sets de entrenamiento, validación y *testing*, utilizando validación cruzada.
3. Determinar y extraer un conjunto de características relevantes de los conectomas en base a la representatividad e importancia para los modelos.
4. Evaluar la efectividad de los modelos tradicionales entrenados, utilizando las métricas tradicionales de Machine Learning (*precision*, *accuracy*, *F1-Score*).

La metodología utilizada para lograr estos objetivos se detalla en el Capítulo 4

2.3. Alcances y limitaciones

El contexto en el que se ve enmarcado este trabajo inherentemente impone limitaciones en el alcance de este. Las siguientes secciones abordan los riesgos, restricciones y potenciales beneficios

que afectan a este trabajo.

En este trabajo se utiliza parcialmente los datos de los sujetos de la base de datos HCP-A, en particular 606 de los 725 disponibles y por lo tanto los resultados alcanzados no reflejan la totalidad de la base. El total de sujetos no pudo ser considerado por la gran cantidad de tiempo requerida para el preprocesamiento de las imágenes y la obtención de las matrices de conectividad para cada sujeto. Además, los resultados son difícilmente generalizables, tanto por la cantidad de sujetos como por el origen desconocido de estos, pues no se incluye información sociodemográfica en los datos de HCP-A.

Una limitación de este trabajo es que solo considera medidas basadas en grafos para la determinación de características y trabajos futuros pueden incorporar otras características y ver como esto impacta los resultados de los modelos. También se puede abordar la inclusión de todos los sujetos disponibles en la base de datos.

2.3.1. Análisis de riesgos

Esta memoria plantea el entrenamiento de un clasificador por rangos, técnica que se ha utilizado, pero con diferentes representaciones. Si bien es cierto que el tamaño del set de datos es suficiente para que los modelos de Machine Learning sean capaces de generalizar, no deja de ser factible que la inclusión de ambos conectomas, a la hora de representar a los individuos, entorpezca el proceso de clasificación. En ese sentido, este trabajo tiene un carácter exploratorio, sin garantías de mejora respecto al estado del arte.

En lo que a riesgos de gestión se refiere, no se encontraron riesgos relevantes más allá de dificultades en la coordinación de la transmisión de los datos.

2.3.2. Restricciones

Respecto a las restricciones pertinentes a esta propuesta se tiene, en primer lugar, el costo computacional del procesamiento de imágenes de resonancia magnética. Esto pone una limitante directa sobre la cantidad de datos disponible para el entrenamiento del modelo. Actualmente se encuentra ya procesada una cantidad de 606 individuos provenientes de la base de datos HCP-A, que si bien son suficientes para entrenar un modelo, sigue siendo una cantidad limitada en comparativa al volumen de datos que se maneja usualmente en la ciencia de datos. No obstante, es una

restricción característica del área, proveniente de la complejidad que implica la obtención de datos multimodales de MRI.

En segundo lugar, dada la discrepancia que puede existir entre la edad cerebral y la edad cronológica, la estimación del rango etario se complejiza. Aquellos individuos que posean cerebros más discordantes con su edad biológica tenderán a ser mal clasificados. Esto no necesariamente representa una desventaja; sin embargo, en la frontera de los rangos etarios puede volverse problemático, aumentando el error de los modelos.

2.3.3. Costos y beneficios

En lo que a costos refiere, el procesamiento de las imágenes de resonancia magnética tiene un alto costo computacional y temporal. A esto se suma el costo monetario de equipos que tengan las características necesarias para el procesamiento. La base de datos utilizada no está abierta al público de forma libre y el acceso a esta fue obtenido mediante una colaboración de proyectos e investigadores.

Los beneficios de este trabajo contemplan principalmente una contribución al conocimiento en neurociencia. Con los resultados obtenidos en este trabajo de MdeT un póster científico fue presentado en la Jornada Técnica 2025 del AC3E y un artículo fue escrito y enviado a una conferencia internacional.

Capítulo 3

Marco teórico

3.1. Resonancia magnética

La resonancia magnética (MR) es una técnica no invasiva de imagenología médica que utiliza campos magnéticos y ondas de radiofrecuencia para generar imágenes detalladas de los tejidos internos del cerebro. Es capaz de capturar detalles internos de un objeto mediante la medición de señales de resonancia magnética nuclear desde el exterior. Produce una representación en forma de matriz de datos multidimensional que muestra la distribución espacial de alguna propiedad física medida y tiene la capacidad única de generar imágenes bidimensionales en cualquier orientación, así como imágenes tridimensionales.

Existen diversas modalidades de resonancia magnética, variando así la forma en las que se obtienen y las utilidades que prestan. Particularmente, para este trabajo profundizaremos las nociones de resonancia magnética por difusión y funcional.

3.1.1. Resonancia magnética por difusión

Las imágenes de resonancia magnética por difusión (dMRI), se basan en la medición de la difusión molecular del agua en los tejidos, proporcionando información detallada sobre la organización de las fibras y la integridad estructural del cerebro. El agua se distribuye entre las células y los compartimentos fuera de las células, proporcionando datos tanto cualitativos como cuantitativos sobre cómo las moléculas se mueven en el tejido. Esto agrega una nueva dimensión a los estudios de MR, ya que va más allá de la información anatómica proporcionada por las secuencias convencionales y añade detalles sobre la microestructura del tejido.

En el cuerpo humano, el agua se distribuye entre las células y los compartimentos extracelulares. En la materia blanca del cerebro, la difusión del agua es anisotrópica y la dirección en la que se observa la máxima difusividad se ha encontrado que coincide con la orientación de los

tractos de fibras en la materia blanca. Esta propiedad es la que utilizan los modelos de difusión para llevar a cabo la reconstrucción de los tractos y así ofrecer información sobre su integridad.

3.1.2. Resonancia magnética funcional

Otra modalidad es la resonancia magnética funcional (fMRI), que generalmente utiliza el efecto de dependencia del nivel de oxígeno en sangre (BOLD). Este método se basa en los cambios en los niveles de oxígeno en la sangre. Cuando una persona realiza una actividad, el flujo sanguíneo aumenta en las áreas del cerebro involucradas. Inicialmente, el oxígeno en la sangre disminuye levemente, pero luego aumenta para suplir la demanda adicional, y finalmente vuelve a su nivel basal. La sangre oxigenada emite una señal magnética más fuerte, lo que permite que la fMRI detecte estas variaciones. A partir de estos cambios en el flujo sanguíneo y los niveles de oxígeno, la fMRI puede identificar qué regiones del cerebro están activas durante una tarea.

Una extensión de las fMRI es la resonancia magnética de estado de reposo (rs-fMRI), que se centra en las activaciones cerebrales espontáneas entre diferentes áreas. Para esto se solicita a los pacientes cerrar los ojos y relajarse sin pensar en nada en particular, manteniendo los ojos cerrados mientras se realiza el escaneo. Así, seleccionando una región base, se compara con otras en busca de altas correlaciones entre las activaciones de las dos regiones.

3.2. Grafos

Un grafo G consta de un conjunto no vacío y finito $V(G)$ de elementos llamados vértices (o nodos), y un conjunto finito $E(G)$ de pares no ordenados de elementos de $V(G)$ denominados aristas. Por otro lado, se define un grafo con peso como $G = (V, E, w)$, donde $w \in \mathbb{R}_{>0}$ representa el peso de cada arista. Esta noción y toda la teoría asociada a los grafos, es aplicable tanto a las redes funcionales y estructurales del cerebro, que se pueden definir en forma de grafo, donde el conjunto de nodos refleja las regiones cerebrales y las aristas corresponden a las conexiones entre regiones cerebrales. En muchas aplicaciones se usa el término de red para referirse a un grafo, de manera que en este trabajo en ocasiones se habla de red para referirse a red de conectividad cerebral o simplemente red.

El envejecimiento cerebral puede manifestarse de diferentes formas, entre ellas, la disminución de la conectividad cerebral. Gracias a la teoría de grafos podemos obtener métricas de conectividad que permitan caracterizar el estado de la red para cada sujeto.

Las métricas que se consideran para este trabajo se relacionan o buscan medir algún aspecto de la conectividad de la red. La mayoría de estas medidas son globales, dado que proporcionan información global de la estructura del grafo. Sin embargo, algunas de ellas pueden considerarse locales porque capturan información a nivel de subestructuras en el grafo. Las medidas utilizadas se describen a continuación.

- **Centralidad de cercanía / Closeness Centrality:** La centralidad de cercanía es una medida de la distancia media más corta de cada vértice a otro vértice. Un valor más pequeño indica una posición más central en la red [28, 29]. Se calcula para cada nodo y luego se considera tanto el promedio como la desviación estándar.
- **Centralidad de vector propio / Eigenvector centrality:** La centralidad de vector propio mide la importancia de un nodo teniendo en cuenta la importancia de sus vecinos. De manera más intuitiva, para un nodo no sólo tiene en cuenta «cuántos nodos conoce», sino también «a qué nodos conoce» [28]. Se calcula para cada nodo y luego se considera tanto el promedio como la desviación estándar.
- **Coeficiente de agrupación / Clustering coefficient:** Mide lo conectados que están entre sí los vecinos de un vértice. Más concretamente, se calcula como: el número de aristas que conectan a los vecinos de un vértice sobre el número total de aristas posibles entre los vecinos del vértice [28]. Se calcula para cada nodo y luego se considera tanto el promedio como la desviación estándar.
- **Conductancia / Conductance:** La conductancia mide qué tan separada está un subconjunto del grafo (comunidad) del resto del grafo [30]. En base al algoritmo de Newman para detectar comunidades, se considera la comunidad más grande y se calcula su conductancia.
- **Densidad / Density:** Corresponde a la proporción de aristas presentes en el grafo en relación al total posible de aristas. Una densidad de 1 indica un grafo completamente conectado, mientras que un valor de 0 nos indica un grafo desconectado [30].
- **Eficiencia global / Global efficiency:** La eficiencia se define como el inverso de la cantidad de vértices en el camino más corto entre dos nodos. La eficiencia global es entonces la eficiencia promedio del grafo [31]. Caracteriza la eficiencia en el traspaso de información a través de la red.

- **Fuerza / Strength:** A nivel global corresponde a la sumatoria de los pesos de todos los aristas del grafo. Se calcula tanto el promedio como la desviación estándar.
- **Grado / Degree:** Para cada nodo corresponde a la cantidad de aristas que salen de él. Consideramos en grado promedio para cada sujeto y su desviación estándar.
- **Intermediación / Betweenness:** Es la medida en que un determinado vértice se encuentra en los caminos más cortos entre otros vértices. En otras palabras, ayuda a identificar a los individuos que desempeñan un papel de «puente» en una red [28]. Se computa para cada nodo y luego se considera el promedio y la desviación estándar.
- **K Core / Core number:** El K -Core corresponde a un subgrafo maximal que contiene nodos de grado K o más. El Core number de un nodo es el mayor valor K de un K -Core que contenga ese nodo [30]. Se calcula para cada nodo y luego se considera tanto el promedio como la desviación estándar, para capturar el grado de integración de los nodos en la estructura de la red. Esta medida se considera como medida local.
- **Modularidad / Modularity:** Con base en la maximización greedy Clauset-Newman-Moore para modularidad se encuentra una partición del grafo con la mayor modularidad [30]. Sobre esta partición se computa la modularidad, que mide qué tan bien la red se divide en comunidades. Valores más altos indican una fuerte estructura de comunidad (conjuntos altamente conectados de nodos, con conexiones esparzas entre dichos grupos).
- **Número de caminos de largo K :** Contempla la cantidad de caminos (conexiones continuas) de largo 3 y 4.
- **Número de cliques maximales de tamaño K :** Para un grafo, un clique es un subconjunto de nodos completamente conectados entre sí, con k la cantidad de nodos en el subconjunto. Un clique maximal de tamaño k es un clique que no soporta otro nodo en el subconjunto sin dejar de ser un clique. Consideramos el uso de la herramienta Quick Cliques [32], que permite el cómputo de manera eficiente de los cliques maximales de un grafo, particularmente si el grafo es esparzo. Se consideran cliques desde $k = 3$ hasta $k = 11$. Esta medida se considera como medida de información local.
- **Número de componentes conexas:** Corresponde a la cantidad de subgrafos independientes dentro del grafo.
- **Page Rank:** PageRank calcula una clasificación de los nodos del grafo basándose en la estructura de los enlaces entrantes en cada nodo [30]. Se diseñó originalmente como algoritmo

para clasificar páginas web. Se calcula para cada nodo y luego se considera tanto el promedio como la desviación estándar.

- **Tamaño del máximo clique:** Considera el tamaño del clique maximal más grande encontrado sobre el grafo. Esta medida es global.
- **Transitividad / Transitivity:** La transitividad es la probabilidad global de que el grafo tenga nodos adyacentes interconectados, lo que revela la existencia de comunidades estrechamente conectadas. Se calcula mediante la división entre el número observado de cliques de tamaño 3 y el número máximo posible de cliques de tamaño 3 en el grafo [33].

3.3. Machine Learning

El aprendizaje automático es un campo de estudio de la inteligencia artificial que se ocupa del desarrollo y el estudio de algoritmos estadísticos capaces de aprender de los datos y generalizarlos a datos desconocidos.

Combinado con la extracción de métricas de conectividad para las representaciones de las redes funcionales y estructurales, el Machine Learning permite la identificación de patrones subyacentes que permitan profundizar la comprensión que se tiene de estas.

Los problemas de Machine Learning se dividen de acuerdo a la naturaleza de la predicción a obtener. Para datos continuos corresponde el entrenamiento de regresores, mientras que para datos discretos corresponde el entrenamiento de clasificadores. Si bien, la edad es inherentemente un valor continuo, existen aproximaciones que discretizan el valor, utilizando rangos etarios para trabajar.

3.3.1. Algoritmos de Machine Learning

Dentro del área de Machine Learning existen 3 clasificaciones principales para los algoritmos: los de aprendizaje supervisado, aprendizaje no supervisado y aprendizaje por refuerzo. El uso de los algoritmos dentro de estas categorías varía dependiendo de la naturaleza de los datos.

En este trabajo, se implementaron algoritmos de aprendizaje supervisado, seleccionados principalmente por la naturaleza del problema (predecir etiquetas conocidas). Estos algoritmos se entrenan a partir de datos previamente etiquetados, lo que permite estimar una función que relaciona

las entradas y las salidas realizando así predicciones en sujetos nuevos. En este trabajo, tales entradas corresponden a características que fueron extraídas de la conectividad cerebral estructural y funcional.

La selección de los modelos a continuación, fue influenciada tanto por su uso frecuente en la literatura del área, como es el caso de SVM, como por el hecho de que son buenos para trabajar con conjuntos de características numéricas (en este caso las medidas extraídas de los grafos). En este tipo de escenarios, estos modelos presentan una buena capacidad de generalización, robustez y cierto grado de simplicidad a la hora de interpretar los resultados. Se añade a esto el bajo costo computacional y temporal en el que se incurre al entrenar con conjuntos de datos relativamente pequeños y la rápida disponibilidad en términos de implementación. El objetivo entonces es mantener los modelos simples en formulación conservando la mayor cantidad de comparabilidad con otros trabajos.

1. **Support Vector Machine:** El algoritmo de SVM funciona encontrando el hiperplano óptimo que separa los datos en diferentes clases. En los casos en los que no se puedan separar los datos, SVM transforma los valores a un espacio dimensionalmente superior, en el que sí pueda existir una separación [34, 35]. Al tener su base de clasificación en los hiperplanos, pertenece a la familia de algoritmos basados en planos/hiperplanos.
2. **Random Forest:** Random Forest es un algoritmo basado en *decision trees* que utiliza dichas estructuras para conseguir una gran cantidad de resultados, los que promediados construyen la predicción. Estos árboles de decisión, plantean "preguntas" que al responderse con la información disponible encaminan el eventual resultado en una dirección específica, descartando las demás posibilidades [36, 35].
3. **Logistic Regression:** Por último, la regresión logística, también dentro de la familia de algoritmos basados en hiperplanos, utiliza la función sigmoide para realizar la clasificación de los datos. Sin embargo, corresponde utilizar regresión logística multiclase, que aplica la función softmax para asignar una probabilidad a cada respuesta posible [35].

3.3.2. Evaluación de modelos de ML

Para evaluar el desempeño de los modelos de Machine Learning, ya sea en tareas de regresión o clasificación, existe un conjunto de métricas estandarizadas. Estas buscan describir la capacidad del modelo de realizar predicciones acertadas y/o qué tan buenas son estas predicciones.

Las métricas más comunes para evaluar a los modelos de clasificación corresponden a:

1. Exactitud/*Accuracy*: La exactitud es la proporción de todas las clasificaciones que fueron correctas, ya sean positivas o negativas.

$$Accuracy = \frac{\text{clasificaciones correctas}}{\text{total de clasificaciones}} = \frac{TP + TN}{TP + TN + FP + FN}$$

2. Precisión/*Precision*: La precisión es la proporción de todas las clasificaciones positivas del modelo que realmente son positivas.

$$Precision = \frac{\text{clasificaciones correctas de la clase}}{\text{total de clasificaciones para la clase}} = \frac{TP}{TP + FP}$$

3. Recuperación/*Recall*: La recuperación o tasa de verdaderos positivos (TPR) es la proporción de todos los positivos reales que se clasificaron correctamente como positivos.

$$Recall = \frac{\text{clasificaciones correctas de la clase}}{\text{total elementos de la clase}} = \frac{TP}{TP + FN}$$

4. F1-Score: El F1-Score es una medida compuesta que equilibra la importancia de la precisión y la recuperación. Además preferible a la precisión para los conjuntos de datos con desequilibrio de clases.

$$F1 = 2 * \frac{precision * recall}{precision + recall} = \frac{2TP}{2TP + FP + FN}$$

Si bien la exactitud mide la proporción de predicciones correctas en general, esta no logra distinguir entre los diferentes tipos de errores. Por este motivo, aunque se presenta su definición, no se considera como métrica de evaluación para nuestros modelos.

3.4. Estado del arte

Diversas investigaciones actuales se enfocan en la predicción de la edad cerebral mediante técnicas de machine learning aplicadas a rs-fMRI. Un ejemplo relevante es el previamente mencionado

estudio de Han et al. [27], en el que se desarrolló un método basado en características extraídas de redes funcionales cerebrales. Su enfoque incorpora 286 métricas de conectividad por individuo y aplica un filtrado mediante correlación de Pearson para descartar variables poco relevantes. Sin embargo, este trabajo no especifica las características que fueron conservadas. En esa propuesta se realiza una validación sobre los *datasets Consortium for Reliability and Reproducibility (CoRR)* y *NKI/Rockland*, que cubren un amplio rango de edades, obteniendo buenos resultados en la precisión de predicción.

Existen también otros métodos que emplean medidas relacionadas con la estructura de las redes cerebrales como *input* para modelos de regresión. Un trabajo con una metodología similar [37] utiliza indicadores estructurales de nodos y conexiones, entrenando modelos con más de 500 sujetos del *dataset* de HCP-A, alcanzando un error absoluto medio (MAE) de 8.89 años. Además, la integración de diferentes modalidades de imágenes cerebrales ha comenzado a explorarse para capturar patrones más complejos en el conectoma, aunque la combinación sistemática de características multimodales aún representa un área en desarrollo con mucho potencial de mejora.

Dentro de los trabajos recientes y que se relacionan con la predicción de la edad cerebral, el número de sujetos, los métodos y sobre todo los resultados son muy variables. La Tabla 3.1 resume los aspectos más relevantes de tres trabajos asociados al área, describiendo los métodos, datos, cantidad de sujetos utilizados y los resultados de cada uno.

Como evidencia la revisión de la literatura reciente, la mayoría de los trabajos se centran en el uso exclusivo de rs-fMRI, existiendo algunas otras que han explorado de forma parcial la combinación de diferentes modalidades. Estas observaciones definen el enfoque propuesto, planteando la combinación de características estructurales y funcionales aún es una línea de trabajo en la que se pueden realizar avances.

Cuadro 3.1: Resumen de trabajos relacionados.

Autor	Año	Método(s)	Datos	Sujetos	Resultados
Chang et al. [18]	2023	Regresor utilizando selección de características con <i>Least Absolute Shrinkage and Selection Operator</i>	Uso de rs-fMRI para calcular una matriz de conectividad de 264 regiones que luego aplanan y reducen	176	MAE de 1.58 (12.32 considerando outliers).
Sorooshyari [25]	2024	Clasificador utilizando pares de rangos etarios. Uno se considera joven y el otro viejo.	Uso de rs-fMRI para obtener señales de 42 regiones y sus medidas de tendencia central.	887	Mean Accuracy Rate de 0.637 (0.740 entre edades 21–30 y 71+).
Baghernezhad y Reza [19]	2024	Clasificador utilizando selección de características basado en valores estadísticos.	Uso de rs-fMRI para obtención de matriz de conectividad, grafo y métricas de conectividad.	40	Accuracy Rate de 0.822.
Kim et al. [38]	2022	Uso de <i>Connectome-based predictive modeling</i> (CPM) sobre redes de conexión funcional de 248 regiones.	Uso de rs-fMRI para obtención de nodos y aristas más importantes mediante <i>Leave One Out Cross Validation</i> (LOOCV).	198	RMSE de 10.55.
Gonneaud et al. [39]	2021	Red neuronal utilizando selección de características sobre modelos de SVM y Regression Tree entrenados con métricas de grafos.	Uso de rs-fMRI para obtención de matriz de conectividad y 26 métricas de grafo.	1340	RMSE de 13.24 y MAE de 11.58.
Feng et al. [37]	2022	Evaluación de algoritmos de regresión sobre diferentes métricas.	Medidas de nodo y aristas relacionadas con la materia gris y blanca, espesor cortical y conectividad funcional.	1193	MAE de 8.89.

Capítulo 4

Desarrollo del proyecto

4.1. Materiales

4.1.1. Base de datos Human Connectome Project - Aging

La base de datos utilizada para este trabajo corresponde a la de Human Connectome Project - Aging (HCPA) [40], la cual cuenta con 725 sujetos con edades que comprenden entre los 36 y 100 años. Para el análisis realizado, tanto de conectividad funcional y estructural, se utilizaron 606 sujetos del total disponible. La razón reside en el tiempo necesario para el procesamiento de los datos.

Para transparentar el proceso de obtención de imágenes, desde HCP-A se explicita la siguiente metodología utilizada para la construcción de la base de datos:

Las imágenes estructurales se obtuvieron mediante una secuencia MPRAGE multieco ($TE = 1.8\text{--}7.2$ ms, $TR = 2,500$ ms, $FOV = 256 \times 256$ mm², 208 cortes, tamaño de voxel = 0.8 mm isotrópico, ángulo = 8°) [41]. Las imágenes de difusión no procesadas incluyeron dos *shells* ($b = 1500/3000$ s/mm²) con 98–99 direcciones cada una, vóxeles de 1.5 mm, multibanda = 4 y $TR = 3.23$ s [41]. Finalmente, las resonancias de estado de reposo se adquirieron con una secuencia GRE multibanda ($MB = 8$, $TR/TE = 800/37$ ms, ángulo = 52°) con vóxeles isotrópicos de 2 mm y cobertura total del cerebro (72 cortes). Se registraron 4 series alternando codificaciones AP/PA, cada una con 488 volúmenes, mientras los sujetos fijaban la vista en una cruz para reducir artefactos por movimiento [41].

Junto a esto se cuenta además con información complementaria de cada sujeto, como su edad en meses, sexo al nacer y otras métricas asociadas a la realización de la toma de imágenes.

4.1.2. Atlases de referencia

Para el procesamiento de imágenes de resonancia magnética se utilizan atlases corticales y subcorticales. Los atlases utilizados por este trabajo corresponden al atlas cortical de Schaefer [42] en su versión de 100 parcelas y el atlas subcortical de Tian [43] de 16 parcelas. Estos atlases sirven como mapas ya estandarizados del cerebro, permitiendo asignar cada píxel o vóxel de las imágenes cerebrales a una región anatómica. En este caso, el de Schaefer al ser cortical, divide la corteza cerebral en 100 regiones distintas, mientras que el atlas de Tian se centra en estructuras subcorticales, definiendo 16 regiones.

4.1.3. Herramientas existentes para el procesamiento de imágenes

El procesamiento de imágenes de resonancia magnética cuenta con una metodología completamente establecida y rigurosa. Gracias al apoyo del grupo de imágenes médicas de la Universidad de Concepción, diversos *scripts* fueron facilitados con el objetivo de simplificar el proceso y mantener los estándares de procesamiento.

Los códigos han sido escritos en diferentes lenguajes, principalmente Python, mientras que otros corresponden a implementaciones en Matlab. Junto a estos también se consideran los códigos incluidos en algunos softwares, como es el caso de *HCP Pipelines*. Todo el procesamiento de imágenes se ha llevado a cabo en equipos de laboratorio de la Universidad de Concepción.

4.2. Métodos

4.2.1. Procesamiento de imágenes de difusión

El procesamiento de las imágenes (que para este trabajo resulta ser más bien el preprocesamiento), consiste en la transformación de las imágenes de dMRI a una matriz de conectividad funcional que indica la fuerza de la conexión entre pares de regiones corticales y subcorticales. Esta matriz será utilizada para la construcción de un grafo para así obtener medidas que servirán como input para los modelos.

Para el procesamiento de las imágenes de difusión se ha utilizado el pipeline de Human Connectome Project [44], apoyado de los softwares FSL [45] y FreeSurfer [46]. Este preprocesamiento busca eliminar artefactos y mejorar la calidad de las imágenes para los siguientes pasos.

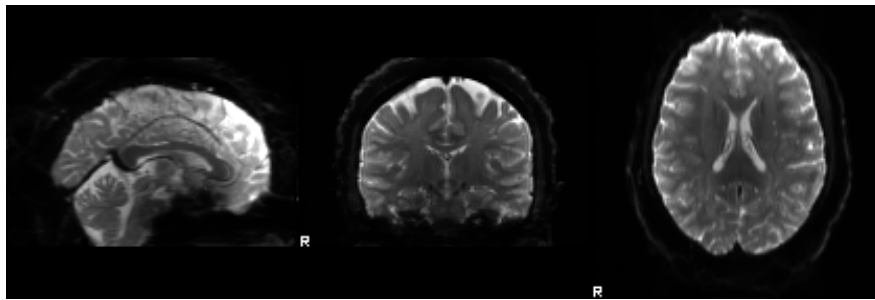


Figura 4.1: Imagen estructural T1 utilizada en el preprocesamiento de las imágenes de difusión.

Se comienza realizando una normalización de las imágenes sin difusión. Esto se hace para las cuatro series de difusión (98 direcciones anterior-posterior, 98 direcciones posterior-anterior, 99 direcciones anterior-posterior y 99 direcciones posterior-anterior). Luego, junto a la herramienta *topup* provista por FSL, se calcula la magnitud de la distorsión. A continuación, con la serie temporal de ambas direcciones de codificación se modelan las distorsiones (provenientes de la corriente de Foucault y el movimiento de la cabeza), utilizando la herramienta *eddy*, incluida en FSL. Por último, se corrige la distorsión y las imágenes sin difusión son alineadas con la imagen T1w del sujeto. Los datos de difusión finalmente se ajustan al espacio estructural nativo con una resolución de 1.25 mm.

Luego de esto, se realizó la parcelación cortical utilizando el atlas de Schaefer 2018 de 100 parcelas y 7 redes. La parcelación fue proyectada en la imagen estructural T1 (ilustrada en la Figura 4.1) de cada sujeto utilizando herramientas de FreeSurfer. También se realizó la parcelación subcortical utilizando el atlas de Tian Subcortex 2020 de 16 parcelas subcorticales. El software Brainstorm [47] para MATLAB permitió la proyección de la parcelación sobre el espacio estructural T1 de cada sujeto. Antes de finalizar la etapa de parcelación, se editaron las etiquetas de las parcelaciones, modificando también la parcelación subcortical para que corresponda en dimensión y tamaño de voxel con la parcelación cortical. Como último paso se unieron ambas parcelaciones en un único archivo NIfTI (Neuroimaging Informatics Technology Initiative) [48].

Reuniendo los resultados de los pasos anteriores y finalizando el procesamiento de imágenes de difusión, se realizó el cálculo de la tractografía probabilística, utilizando MRtrix3 [49]. Para ello se estima la distribución de la orientación de las fibras (FOD) utilizando el modelo de deconvolución esférica restringida de múltiples tejidos (MSMT-CSD), que requiere de una función de respuesta representativa para el conjunto de sujetos. Inicialmente se estiman estas funciones para cada uno

de los sujetos disponibles, capturando características particulares del escaneo de cada sujeto. Sin embargo, para reducir la variabilidad intersujeto, se genera una respuesta promedio para el conjunto completo. Es esta la función que finalmente se utiliza para la estimación de las FOD, garantizando mayor consistencia entre los individuos procesados. Realizado esto, se genera el tractograma probabilístico de 30 millones de fibras utilizando tractografía anatómicamente restringida (ACT). Una vez generado el tractograma se aplica el algoritmo SIFT2 (Spherica-deconvolution Informed Filtering of Tractograms 2) [50], que asigna un peso a cada una. Finalmente, se construyeron matrices de conectividad estructural para cada sujeto aplicando los tractogramas filtrados sobre la parcelación previamente combinada.

Se cuenta en total con matrices de conectividad estructural para un total de 606 sujetos de la base de datos. Los 121 sujetos restantes continúan procesándose.

4.2.2. Procesamiento de rs-fMRI

El caso de las imágenes de resonancia magnética es particular. Al comienzo del trabajo, se recibieron la totalidad de las matrices de conectividad funcional asociadas a las rs-fMRI. El trabajo de procesamiento de estos datos fue realizado por Catalina Vergara, del grupo de Análisis de imágenes médicas de la Universidad de Concepción. El procesamiento para la obtención de estas es como sigue.

Teniendo como punto de partida las imágenes ya preprocesadas disponibles en la base de datos, se realizó una parcelación utilizando el software Connectome Workbench [51]. De un modo análogo a las imágenes de difusión, se utilizaron los atlas corticales y subcorticales de Schaefer y Tian, respectivamente.

Con esto se calcularon las señales de la actividad cerebral (correspondientes a series temporales) para cada región parcelada, promediando los cursos temporales de todos los vóxeles de la región. A partir de estas señales se construyó entonces una matriz de correlación, que determina la asociación por pares entre las series temporales de las parcelas cerebrales. Esto representa la red de conectividad funcional, y fue generada para cada sujeto.

Si bien esta matriz se puede utilizar para realizar análisis, se le aplicó la transformación Z de Fisher para mejorar la normalidad de los coeficientes de correlación.

4.2.3. Exploración de los datos

Utilizando como insumos las representaciones funcionales y estructurales de los 606 sujetos en forma matricial, se realizó primeramente una exploración de los datos disponibles. Para ello, utilizando la información complementaria, observamos la distribución de la edad en los sujetos de estudio.

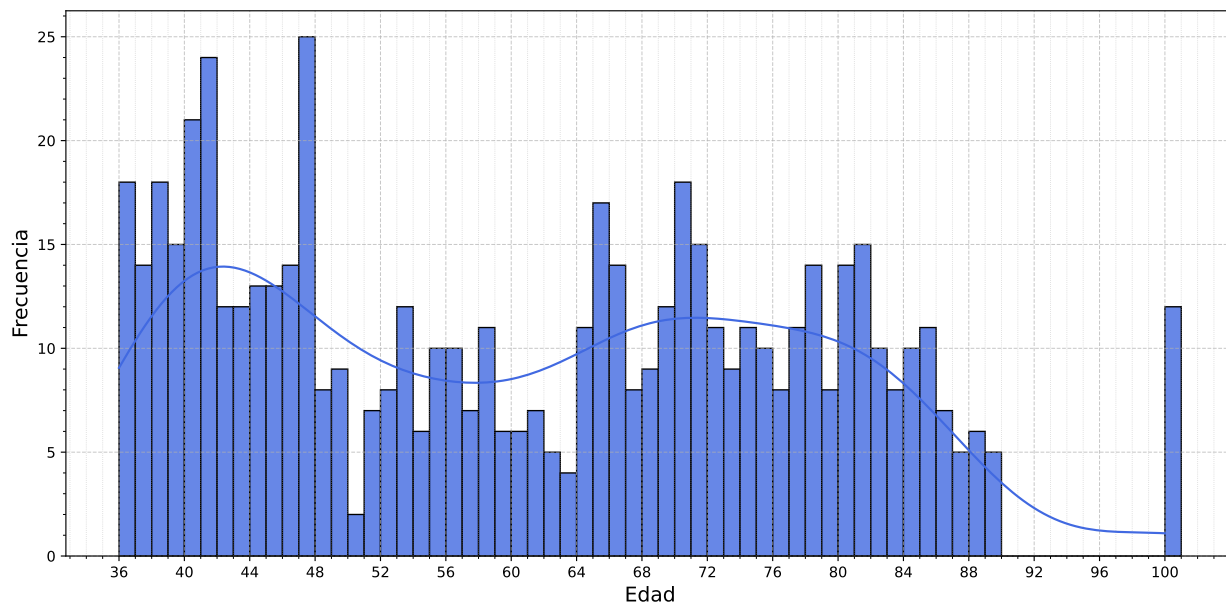


Figura 4.2: Distribución de edades para los 606 sujetos.

La Figura 4.2, que muestra la distribución total de edades en los sujetos, evidencia una gran cantidad de estos con edades menores a 50 años. También se observa un buen nivel de representación en el segmento de los adultos mayores (mayores de 60 de acuerdo a la Ley 19.828). Sin embargo, existe una subrepresentación del rango entre 50 y 60 años de edad.

Con el objetivo de definir grupos etarios para realizar la clasificación, observamos la propia división de edades que realiza HCP para los individuos en su base de datos. En su trabajo, HCP-A [40] distingue entre 3 grandes grupos etarios: **Maduros** (*Mature*, entre 36 y 64 años), **Mayores** (*Old*, entre 65 y 79 años) y **Más mayores** (*Oldest old*, mayores de 80 años) [40]. Notamos un desbalance de clases inherente a la captura de datos, con una concentración de más del 50 % de los sujetos en la clase Maduros, debido al gran rango de edad que abarca la clase. La Figura 4.3 describe gráficamente la distribución etaria del rango HCP-A.

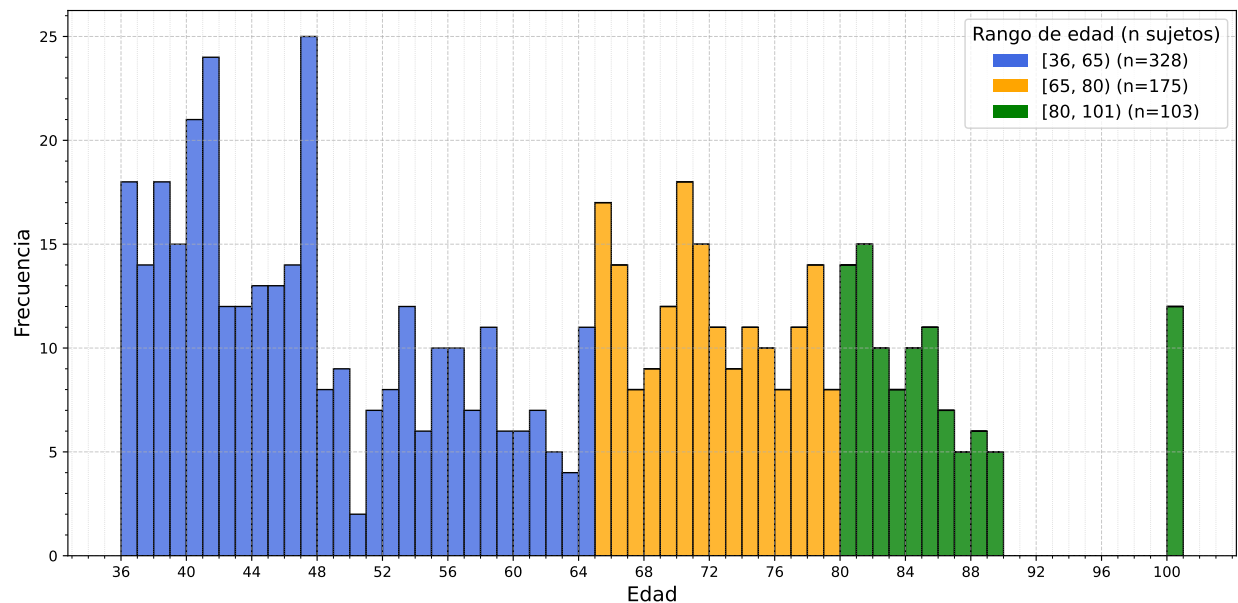


Figura 4.3: Distribución de rangos etarios HCP-A.

Una aproximación intuitiva consiste en la definición de rangos etarios que abarquen cada uno la misma cantidad de años. Se proponen entonces los siguientes rangos etarios: **Grupo joven** (desde los 36 años hasta los 56 años), **Grupo medio** (entre 57 y 77 años) y **Grupo mayor** (desde los 78 hasta los 100 años). La división en estos grupos, como se observa en la Figura 4.4, balancea un poco más las clases y desde una perspectiva biológica distingue de manera un poco más clara ciertos estadios de la vida humana.

Finalmente, se propone una clase que busque tener el mayor balance posible entre clases, ilustrado en la Figura 4.5. Estos rangos se componen por: **Adulto joven** (entre 36 y 47 años), **Adulto pleno** (entre 48 y 69 años) y **Adulto mayor** (desde los 70 hasta los 100 años). En línea con el objetivo, esta división etaria mantiene las clases bien balanceadas a costa de unos rangos poco interpretables.

Matrices estructurales

Ahondando en los datos que conforman las representaciones estructurales de los sujetos, encontramos una distribución de valores en extremo sesgada. Confirmando lo que se observa en la

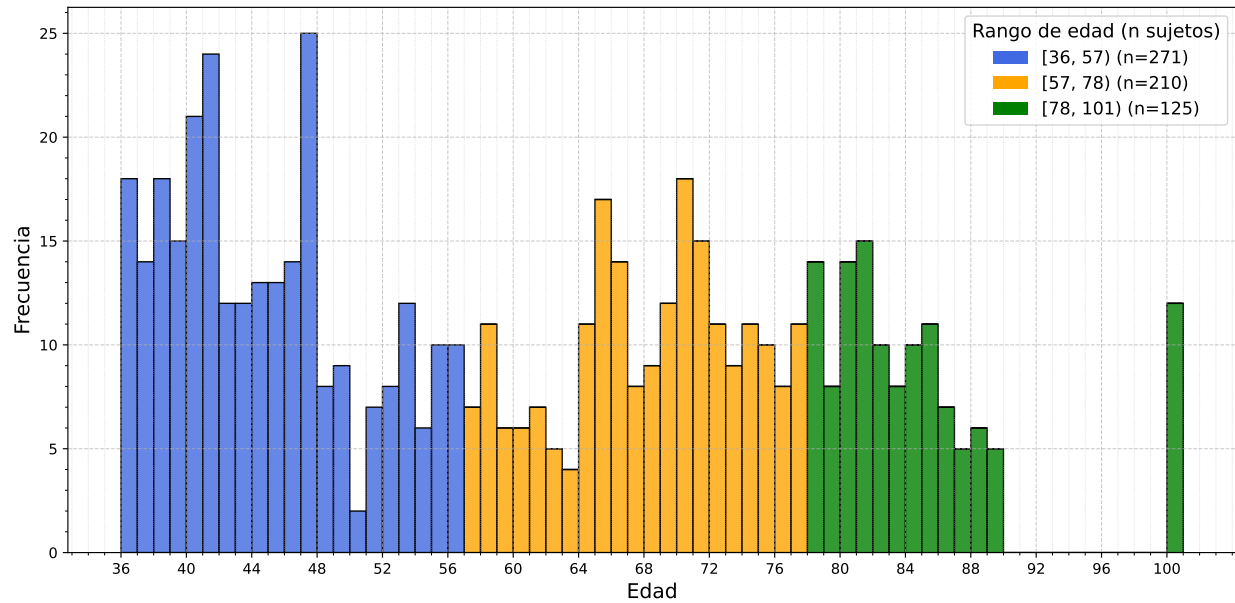


Figura 4.4: Distribución de rangos etarios equidistantes.

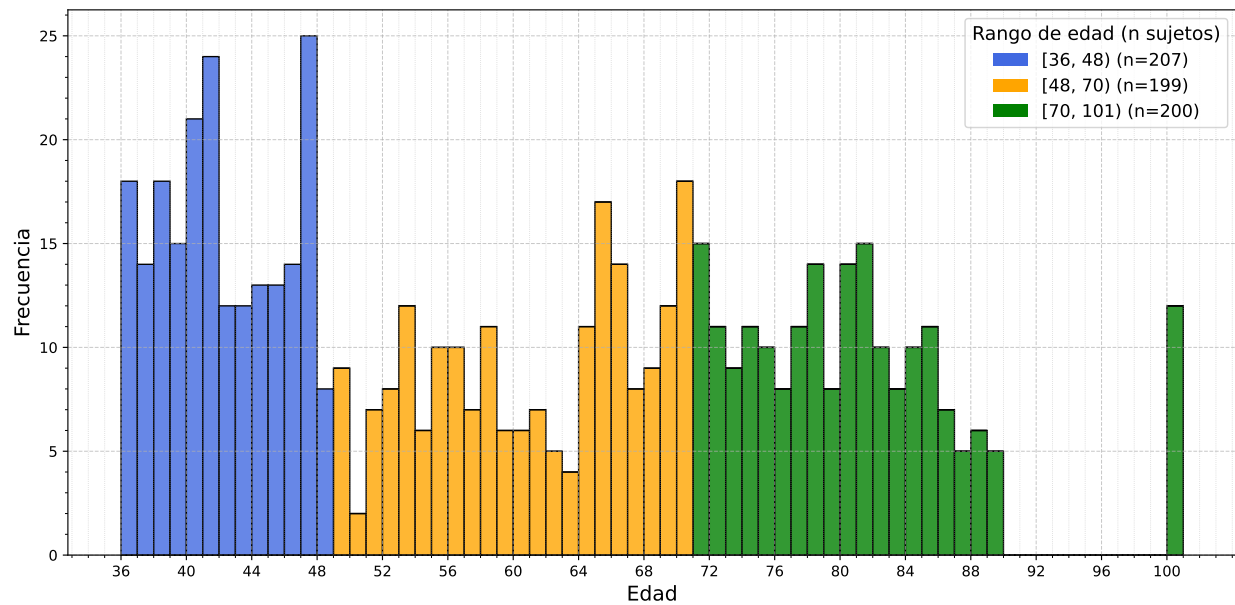


Figura 4.5: Distribución de rangos etarios balanceados.

Figura 4.6 el 95 % de los valores de conectividad estructural son menores o iguales a **2.356**. El sesgo es tal, que el 1 % de los datos tiene un valor superior a **7.829**. Esto puede sugerir que estos valores extremos (aquellos fuera del percentil 99) corresponden a valores atípicos. Sin embargo, es muy probable que estos valores sean errores arrastrados desde las mediciones realizadas por los que fueron considerados para el análisis.

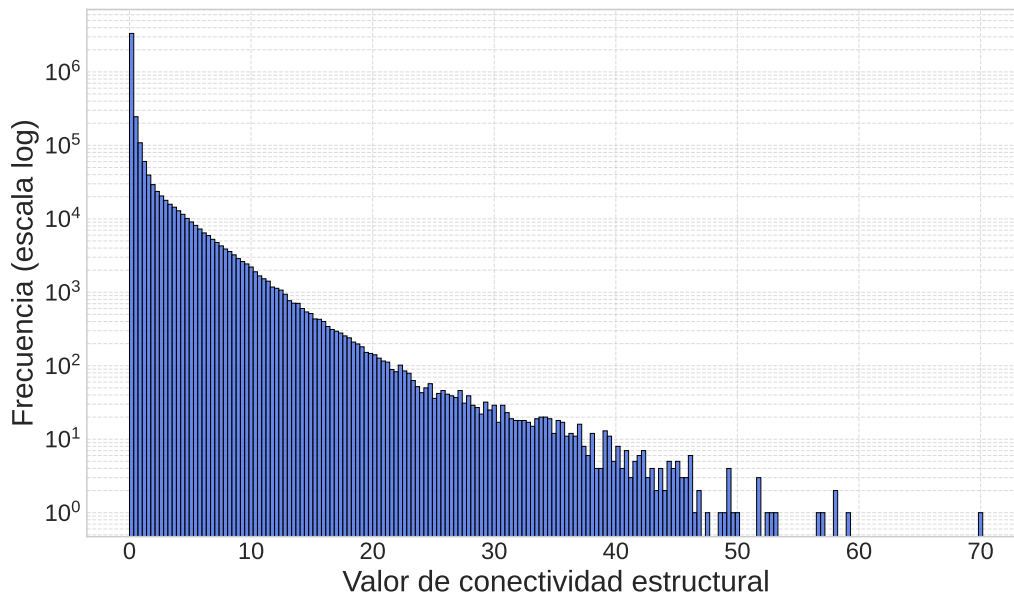


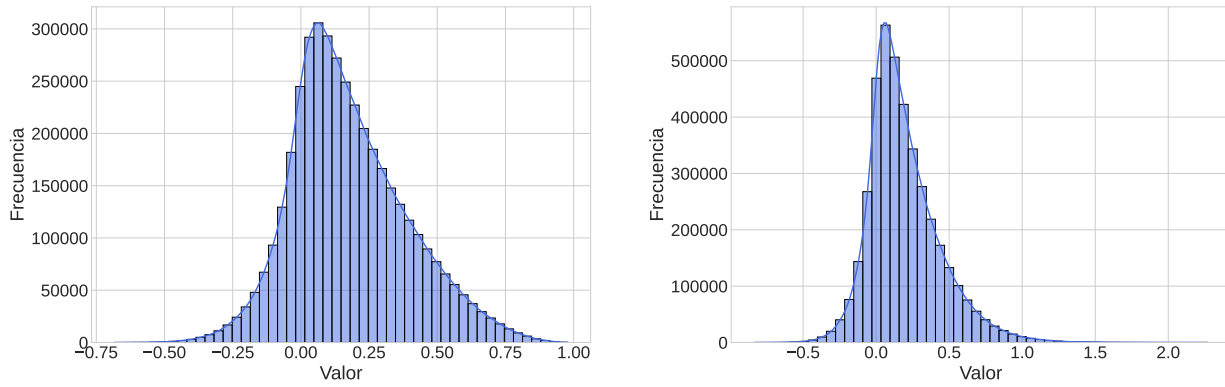
Figura 4.6: Distribución de valores estructurales con escala logarítmica en el eje Y.

Alternativas evaluadas para la corrección de esta distribución de valores fueron las transformaciones logarítmicas y de potencia, particularmente la transformada de Yeo-Jhonson que es una adaptación de la transformación Box-Cox para variables con valores nulos y/ negativos. Estas alternativas fueron descartadas debido a resultados pobres, en los que no lograron acercar la distribución de valores a una más útil sin agregar distorsión de por medio.

Matrices funcionales

Tal como se mencionó anteriormente, los datos funcionales ya se encontraban procesados, por lo que no presentaron mayores dificultades en su tratamiento. La Figura 4.7 muestra el efecto de la transformación Z de Fisher al aplicarla a los datos originales. Notamos que la transformada acerca el centro de la distribución a 0 y que, al pasar de coeficientes de Pearson a valores Z, extiende el

dominio de los valores más allá del intervalo $[-1, 1]$. Los valores funcionales, por lo tanto, fueron modificados solo por la transformada Z para proporcionarles mayor normalidad a los datos.



(a) Distribución de los valores de las matrices de conectividad funcional original.

(b) Distribución de los valores de las matrices de conectividad funcional luego de la transformada de Fisher.

Figura 4.7: Cambio en la distribución de los valores de conectividad funcional.

4.2.4. Grafos y métricas de conectividad

Para cada una de las representaciones se construyó un grafo, en base a la matriz de conectividad. Cada grafo cuenta con un total de 116 vértices, correspondientes a las regiones obtenidas por la parcelación. Las aristas del grafo corresponden a los pesos entre pares de regiones.

Para los valores estructurales se construyó el grafo estructural (G_S) sin modificaciones, pues todos los pesos son positivos. Sin embargo, para los valores funcionales se consideraron dos aproximaciones: se utilizó el valor absoluto de los pesos para construir el grafo funcional absoluto (G_{FAbs}) y se consideró la separación de los valores positivos y negativos en dos grafos separados, el grafo funcional positivo (G_{F+}) y el grafo funcional negativo (G_{F-}).

Se seleccionaron entonces un total de 16 medidas de conectividad (descritas en la sección 3.2) con las que en total se obtuvieron 34 características para cada representación, como se observa en la Tabla 4.1. Algunas de estas fueron computadas sobre el grafo original, otras sobre el grafo con pesos invertidos (para hacer efectiva la noción de «fuerza» en las aristas), y otras sobre el grafo binarizado.

Cuadro 4.1: Relación entre las medidas de conectividad y las características extraídas.

Medida	Características
Grado / Degree	Degree promedio (1) y desviación estándar (2).
Fuerza / Strength	Strength promedio (3) y desviación estándar (4).
Centralidad de cercanía / Closeness Centrality	Closeness promedio (5) y desviación estándar (6).
Intermediación / Betweenness	Betweenness promedio (7) y desviación estándar (8).
Coefficiente de agrupación / Clustering coefficient	Clustering coef. promedio (9) y desviación estándar (10).
Densidad / Density	Density (11).
Eficiencia global / Global efficiency	Global Efficiency (12).
Transitividad / Transitivity	Transitivity (13).
Número de componentes conexas	Nº componentes conexas (14).
Tamaño del máximo clique	Tamaño clique máximo (15).
Número de caminos de largo k	Caminos de largo 3 (16) y 4 (17).
Centralidad de vector propio / Eigenvector centrality	Eigenvector centrality promedio (18) y desviación estándar (19).
Modularidad / Modularity	Modularity (20).
Conductancia / Conductance	Conductance (21).
K -Core / Core number	K-core promedio (22) y desviación estándar (23).
Page Rank	Page Rank promedio (24) y desviación estándar (25).
Número de cliques maximales de tamaño k	Cliques de tamaño 3 (27), 4 (28), ..., 11 (34).

4.2.5. Umbralización en grafos

Construidos los grafos para cada representación, debido a la gran cantidad de valores de conectividad muy pequeños en las aristas de G_S , como se observa en la Figura 4.6, se proponen diferentes umbralizaciones. Estas se realizan antes de la obtención de las métricas con el objetivo de conservar aquellas conexiones fuertes (aristas de mayor peso), por considerarse de mayor relevancia en la conectividad..

Para los grafos estructurales se consideran distintas umbralizaciones, eliminando aquellas conexiones con valores menores a 0.01, 0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.40, 0.50 y 0.60. Mientras que para los grafos G_{F+} y G_{F-} se consideraron umbrales idénticos a las de G_S , incluyendo además una versión sin umbralizar.

La elección del umbral comienza con valores que aunque pequeños eliminan un gran porcentaje de las conexiones. Este umbral aumenta lentamente (en aumentos de 0.05), para luego crecer más 90 % de las conexiones. Finalmente, para los grafos G_{FAbs} solo se consideraron umbralizaciones

desde 0.0 hasta 0.6 con incrementos de 0.1 en el umbral. Estos datos presentan una umbralización más simple pues su distribución no está sesgada.

La Figura 4.8 resume el flujo completo utilizado como metodología. Se incluyen en la sección superior los 3 rangos etarios considerados.

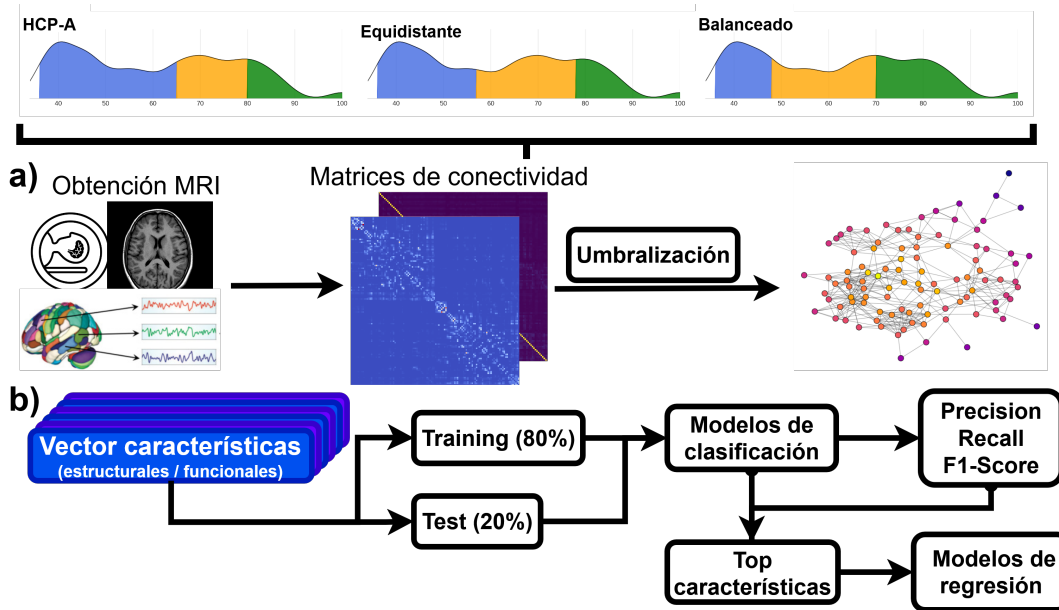


Figura 4.8: Diagrama completo del pipeline utilizado. a) Recolección de datos y procesamiento de estos hasta obtención de métricas basadas en conectividad. b) Entrenamiento de los modelos de clasificación y regresión.

4.2.6. Entrenamiento

Para el entrenamiento de modelos capaces de clasificar etariamente, se consideraron los algoritmos descritos en la sección 3.3.1.

Con los algoritmos elegidos, se consideraron diferentes combinaciones de las características para evaluar el impacto de las mismas sobre las clasificaciones. Así, se entrenaron modelos utilizando cada representación por separado (solo con métricas estructurales, funcionales positivas y negativas, y funcionales absolutas) con el fin de comparar el rendimiento en modelos entrenados con más de un tipo de característica. En este punto es importante mencionar que si bien la separación de los valores funcionales negativos fue realizada, sus resultados son poco significativos, pues no es un procedimiento realizado comúnmente en la literatura [52]. Por este motivo, en el Capítulo

5 no se consideran.

Luego, se combinaron las métricas de las diferentes representaciones funcionales con las métricas estructurales, que actuaron como base de las clasificaciones debido a su buen desempeño. De esta forma, se consideran diferentes combinaciones, abordadas con mayor detalle en la sección 5.1. La combinación de estas características además considera la dimensión del umbral. Es decir, se combinaron también variando los umbrales utilizados.

Para preparar los datos de cara al entrenamiento, se siguieron los procedimientos estándar utilizados en Machine Learning, realizando una estandarización sobre las características. Realizar esto sobre el conjunto de datos es un requisito común para muchos estimadores de ML pues pueden presentar resultados inconsistentes si las características individuales no se aproximan a datos distribuidos normalmente. Además, se utilizó la división clásica de ML, dejando un 80 % de los datos para entrenamiento y un 20 % para *testing*. Por último, antes del entrenamiento como tal, se buscaron los mejores hiperparámetros para cada uno de los algoritmos, utilizando la técnica de *K*-Fold Cross Validation, con $K = 5$.

De esta forma se obtuvieron resultados para diferentes experimentos, los que fueron evaluados utilizando la precisión, el recall y el F1-Score de cada modelo.

4.2.7. Análisis posteriores y re-entrenamiento

Con los modelos entrenados y evaluados, se obtuvo la *permutation feature importance* promedio para cada uno de los algoritmos. *Permutation Feature Importance* (PFI) es una técnica de inspección de modelos que mide la contribución de cada característica al rendimiento estadístico de un modelo. Así, para cada uno de los algoritmos (SVM, Logistic Regression y Random Forest) se filtraron las 20 características más relevantes.

Con estas características, se realizó un re-entrenamiento de los modelos, obteniendo nuevos resultados y métricas de evaluación, que se utilizaron en la comparación con los modelos ya entrenados sobre el conjunto completo de características.

4.2.8. Aproximaciones al problema de regresión

Como el problema de la predicción de edad cerebral es en su naturaleza un problema de regresión más que de clasificación, se utilizaron las características ya computadas y las nociones adquiridas tras la evaluación de los modelos de clasificación para entrenar algunos modelos de regresión.

Para esto se utilizó la edad real de cada sujeto, y los diferentes conjuntos de métricas de conectividad. Esto permitió evaluar el sobreajuste del modelo y, al mismo tiempo, tener un punto de comparativa con modelos de regresión existentes en la literatura.

Para evaluar el modelo de regresión se utilizó el error absoluto medio (MAE), el error cuadrático medio (MSE) y el R^2 . La métrica más interpretable resulta ser el MAE, pues refleja la cantidad de años promedio en la que falla la predicción.

Capítulo 5

Evaluación experimental

5.1. Configuración de los modelos

En concordancia con lo mencionado en el capítulo anterior, se entrenaron modelos basados en los diferentes algoritmos, variando las características utilizadas como *input* y los umbrales utilizados para obtener dichas características.

Las configuraciones pueden ser agrupadas entre aquellas que utilizan solo un tipo de representación y aquellas que combinan ambas representaciones.

1. **Características de una única representación (Modelos tipo 1):** En todas estas configuraciones, los vectores de características que entran al entrenamiento de los modelos se conforman únicamente por las 34 características obtenidas en la subsección 4.2.4. Además, para cada configuración se consideran las distintas umbralizaciones.
 - Medidas estructurales: Se consideran solo los valores de conectividad extraídos de G_S .
 - Medidas funcionales positivas: Se consideran solo los valores de conectividad extraídos de G_{F+} .
 - Medidas funcionales absolutas: Se consideran solo los valores de conectividad extraídos de G_{FAbs} .
2. **Características combinadas (Modelos tipo 2):** Estas configuraciones contemplan la complementación de la información estructural con la funcional. En todos estos casos se considera un vector de 68 características. La aplicación de los umbrales se basa en los *thresholds* estructurales, es decir, desde 0.01, 0.05, 0.10, ..., 0.60.
 - Medidas estructurales y funcionales positivas: Considera la unión de las características obtenidas de G_s con las provenientes de G_{F+} .

- Medidas estructurales y funcionales positivas no umbralizadas: Considera la unión de las características obtenidas de G_s con las provenientes de G_{F+} , con la particularidad de que estas no fueron umbralizadas.
- Medidas estructurales y funcionales absolutas: Considera la unión de las características obtenidas de G_s con las provenientes de G_{FAbs} . En este caso se entrenó un modelo para cada umbral aplicado a G_{FAbs} .

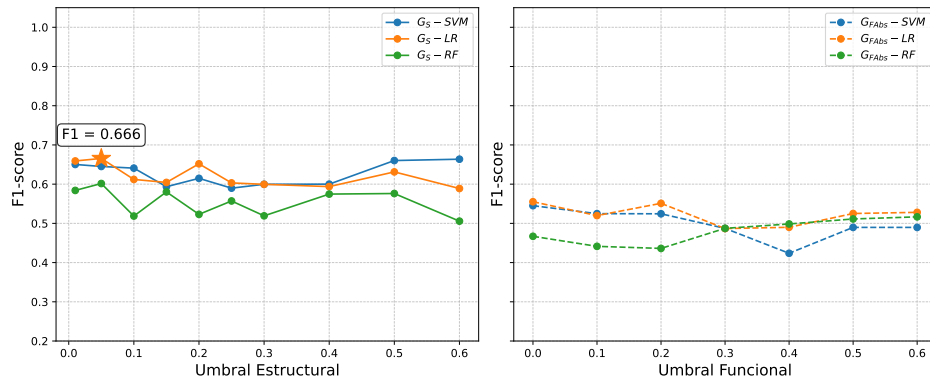
El tamaño del set de entrenamiento para todos los casos fue de 484 sujetos, conservando 122 sujetos para pruebas. Durante el entrenamiento además se consideró la optimización de parámetros correspondiente a cada modelo. Como se mencionó anteriormente, esta se realizó utilizando la técnica de *K-Fold Cross Validation*, con $K = 5$, lo que nos deja 97 sujetos en el conjunto de validación. Estos sujetos fueron reemplazados sistemáticamente por otros 97 distintos en cada iteración de la optimización de parámetros.

Los resultados presentados a continuación se centran principalmente en los modelos que utilizan características funcionales y características funcionales absolutas. La utilización de las características provenientes de G_{F+} , a pesar de estar respaldada por la literatura, no resulta tan relevante en comparativa con las de G_{FAbs} . Para facilitar la visualización de los gráficos, no se presentan en este capítulo.

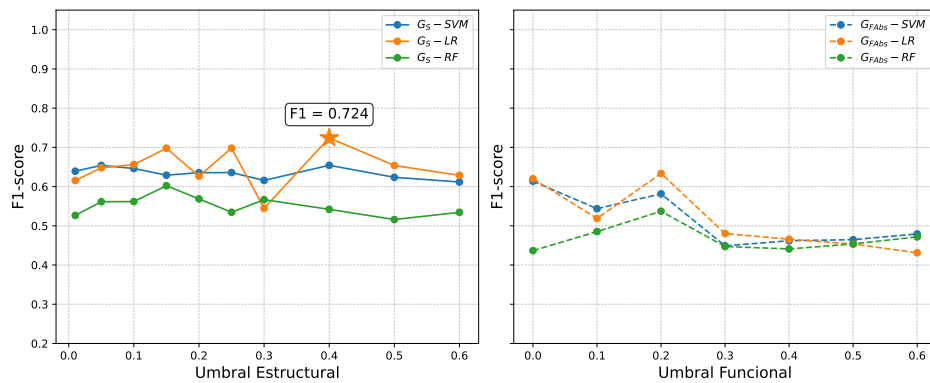
5.2. Modelos de única representación

Los resultados de los modelos que utilizan solo un tipo de característica (estructurales separadas de funcionales) se encuentran en la Figura 5.1. En ella se evidencia un mejor desempeño de las características estructurales a lo largo de todos los umbrales y para todos los rangos etarios considerados. Las características de origen funcional presentan resultados consistentemente peores que su contraparte estructural. Aquellos modelos basados en SVM y LR obtienen además resultados más robustos que aquellos obtenidos por los modelos basados en RF.

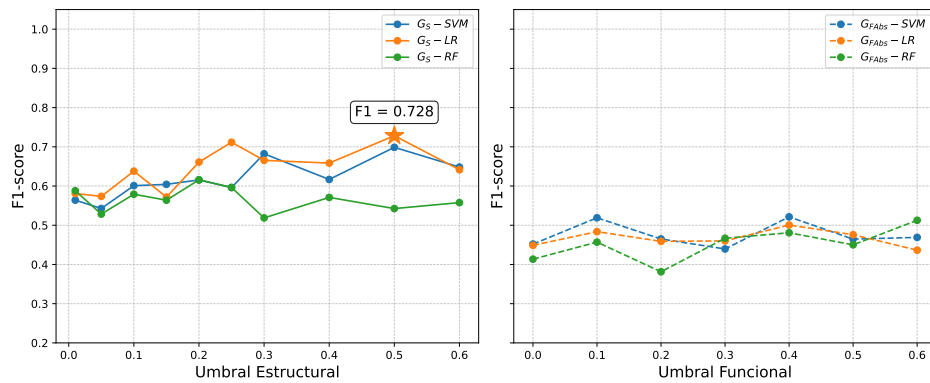
Estos resultados sirven como base para la comparativa posterior con los modelos que combinan las características estructurales y funcionales



(a) Rango balanceado



(b) Rango equidistante



(c) Rango HCP-A

Figura 5.1: F1-score para los modelos de única representación a través de diferentes umbrales en las características.

5.3. Modelos de representación combinada

Los resultados de los modelos que combinan diferentes características se presentan a continuación desagregados por algoritmo base.

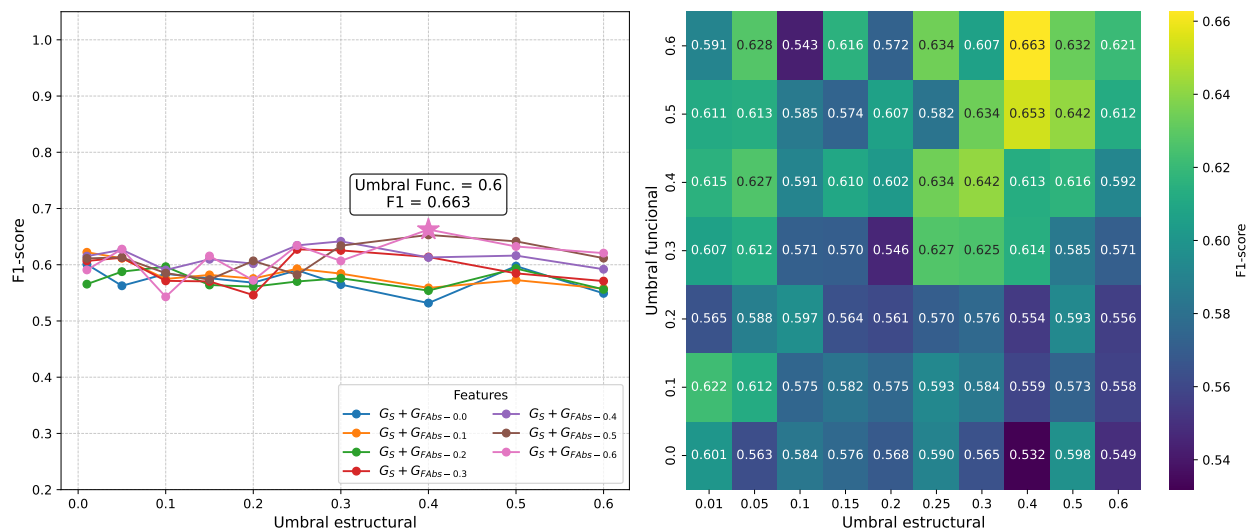
5.3.1. Modelos basados en RF

Los modelos de características combinadas basados en Random Forest obtienen los peores resultados entre todos los modelos. Estos resultados, como se observan en la Figura 5.2, tienen su origen en los aquellos modelos de un tipo de característica. Tal como se mencionó antes, la Figura 5.1 muestra que para todos los rangos etarios Random Forest realiza una clasificación deficiente. Otro motivo posible recae en la naturaleza del algoritmo, indicando que los algoritmos basados en *decision trees* pueden no ser ideales para el problema.

El mejor resultado para estos modelos se observa en la Subfigura 5.2c con un F1-Score de 0.672 sobre el rango propuesto por HCP-A, utilizando un umbral estructural de 0.50 y un umbral funcional de 0.30.

5.3.2. Modelos basados en LR

Los modelos de características combinadas basados en Logistic Regression obtienen resultados considerablemente mejores. Estos, ilustrados en la Figura 5.3, muestran una mejora respecto tanto a los resultados de los modelos de RF como a los modelos de única representación de la Figura 5.1. Respecto a estos últimos se destaca que la mejoría no es absoluta, especialmente para HCP-A y equidistante donde mejores modelos que utilizan solamente características estructurales.



(a) Rango balanceado

Figura 5.2: F1-score para los modelos de representación combinada basados en Random Forest a través de diferentes umbrales estructurales.

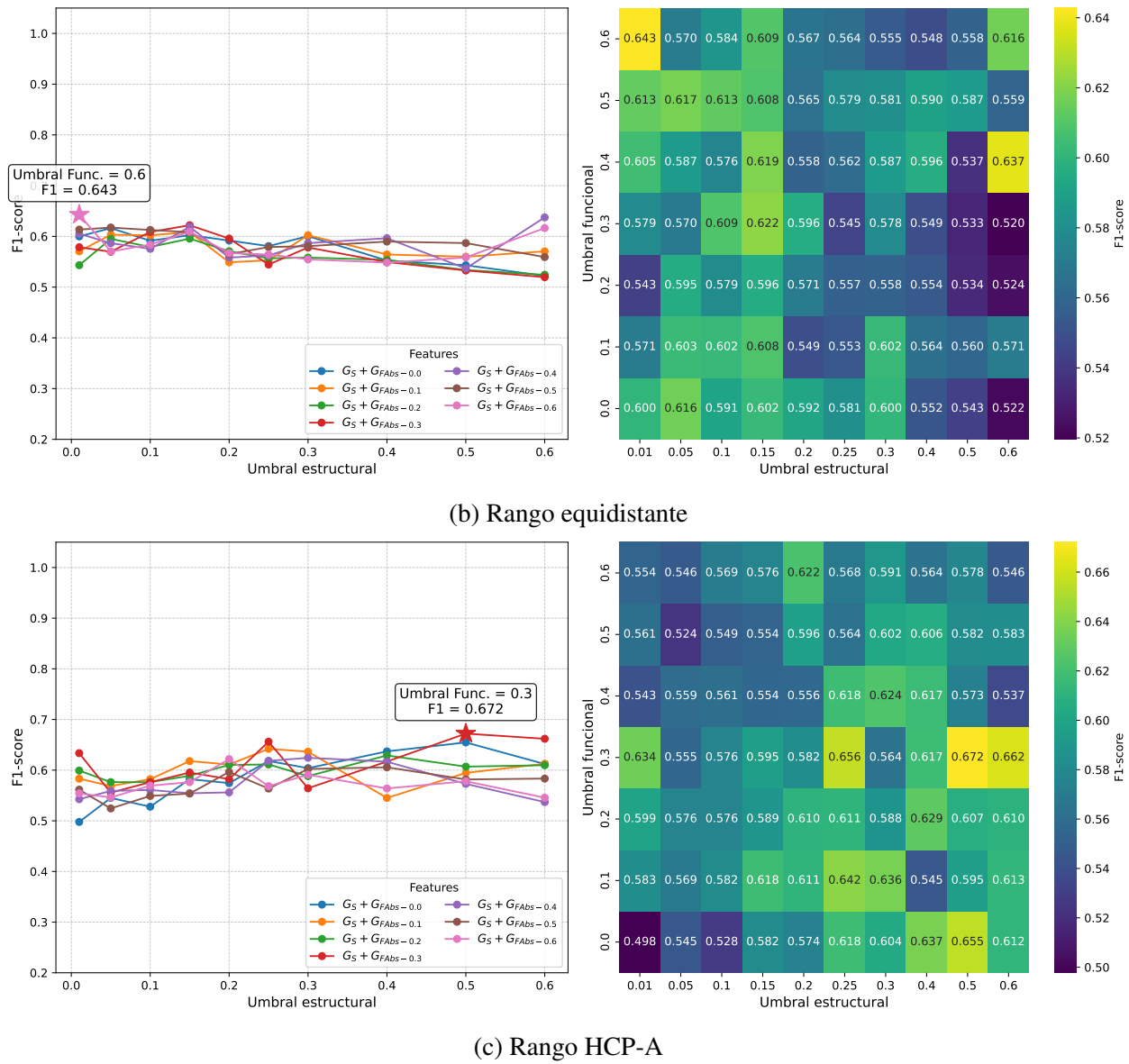


Figura 5.2: F1-score para los modelos de representación combinada basados en Random Forest a través de diferentes umbrales estructurales.

El mejor resultado para estos modelos se observa en la Subfigura 5.3c con un F1-Score de 0.729 sobre el rango propuesto por HCP-A, utilizando un umbral estructural de 0.20 y un umbral funcional de 0.30.

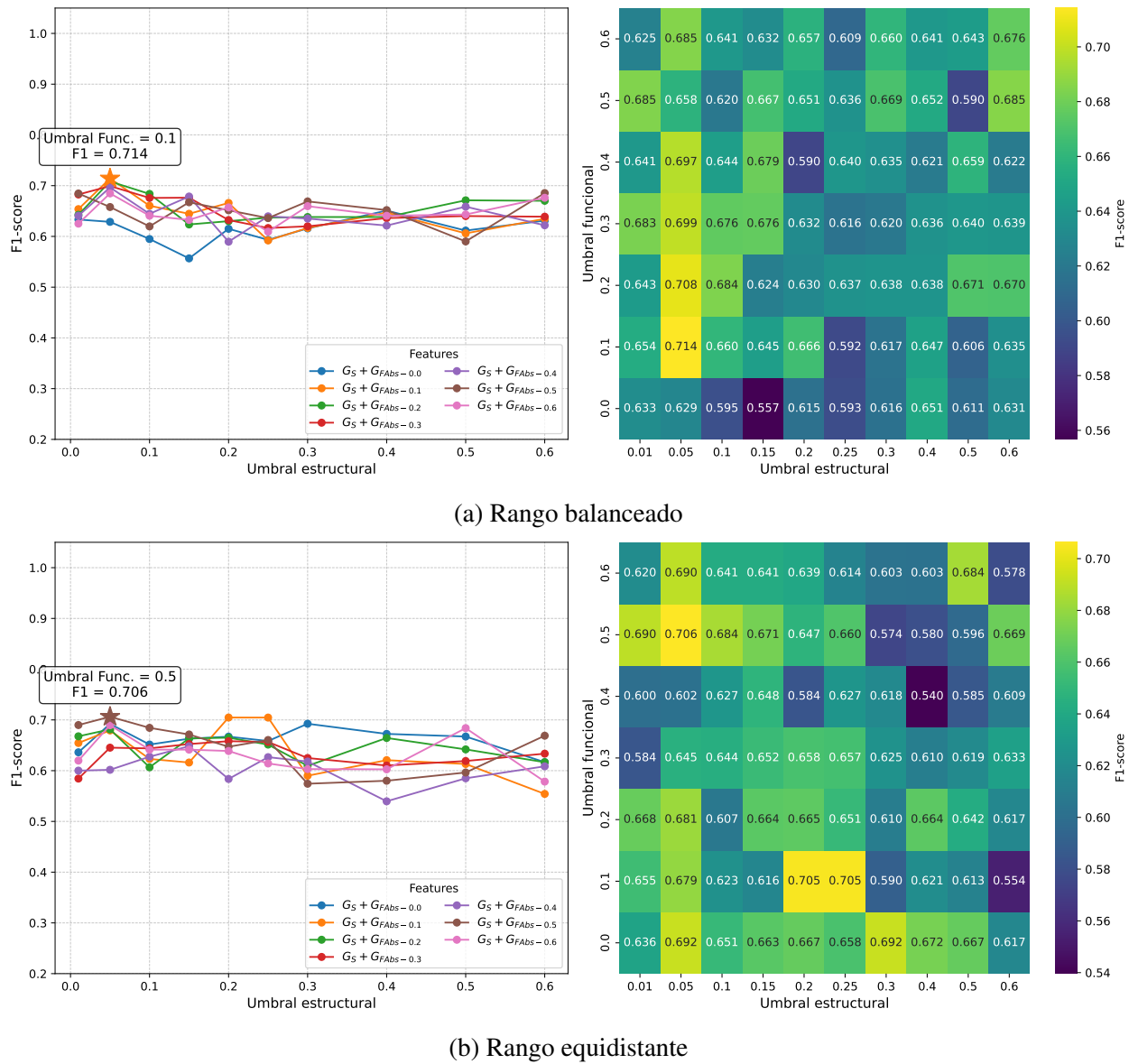
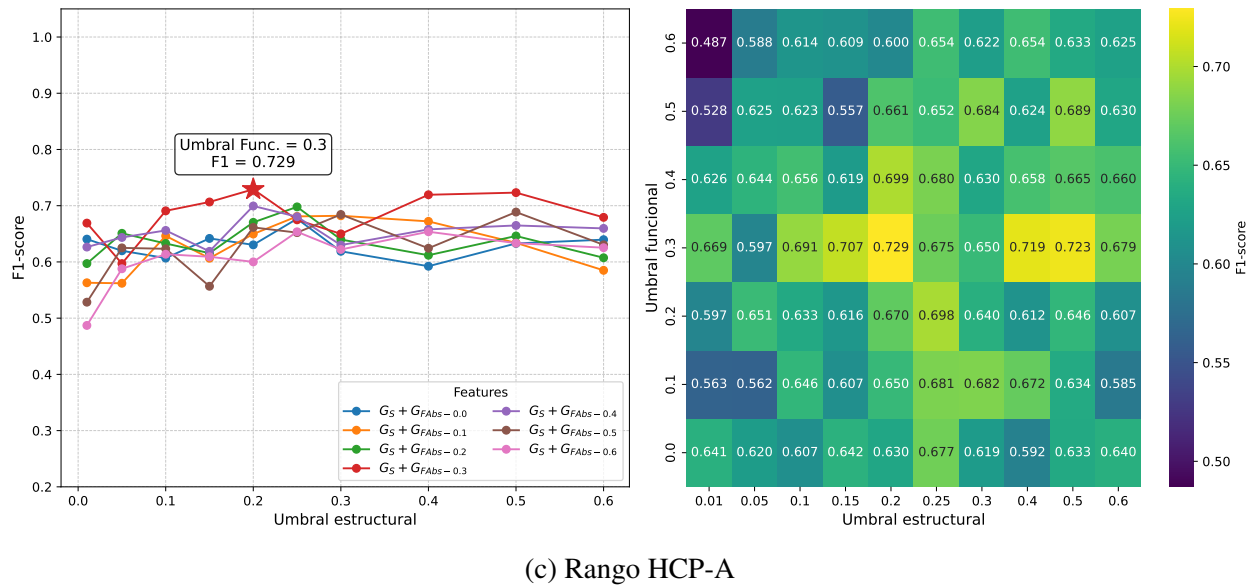


Figura 5.3: F1-score para los modelos de representación combinada basados en Logistic Regression a través de diferentes umbrales estructurales.

5.3.3. Modelos basados en SVM

Finalmente, para los modelos basados en Support Vector Machine se obtienen los mejores resultados de todos los modelos de características combinadas. La Figura 5.4 muestra estos mejores resultados en diferentes rangos. Tal como en los modelos de LR, los resultados de SVM son consistentemente mejores que los de su contraparte de una única característica. Logran también diferenciarse de los modelos de LR, pues no presentan el fenómeno de que ciertas configuraciones



(c) Rango HCP-A

Figura 5.3: F1-score para los modelos de representación combinada basados en Logistic Regression a través de diferentes umbrales estructurales.

de los modelos tipo 1 superaban a los tipo 2.

El mejor resultado para estos modelos se observa en la Subfigura 5.4b con un F1-Score de 0.737 sobre el rango equidistante, utilizando un umbral estructural de 0.05 y un umbral funcional de 0.00, es decir, sin umbralización funcional.

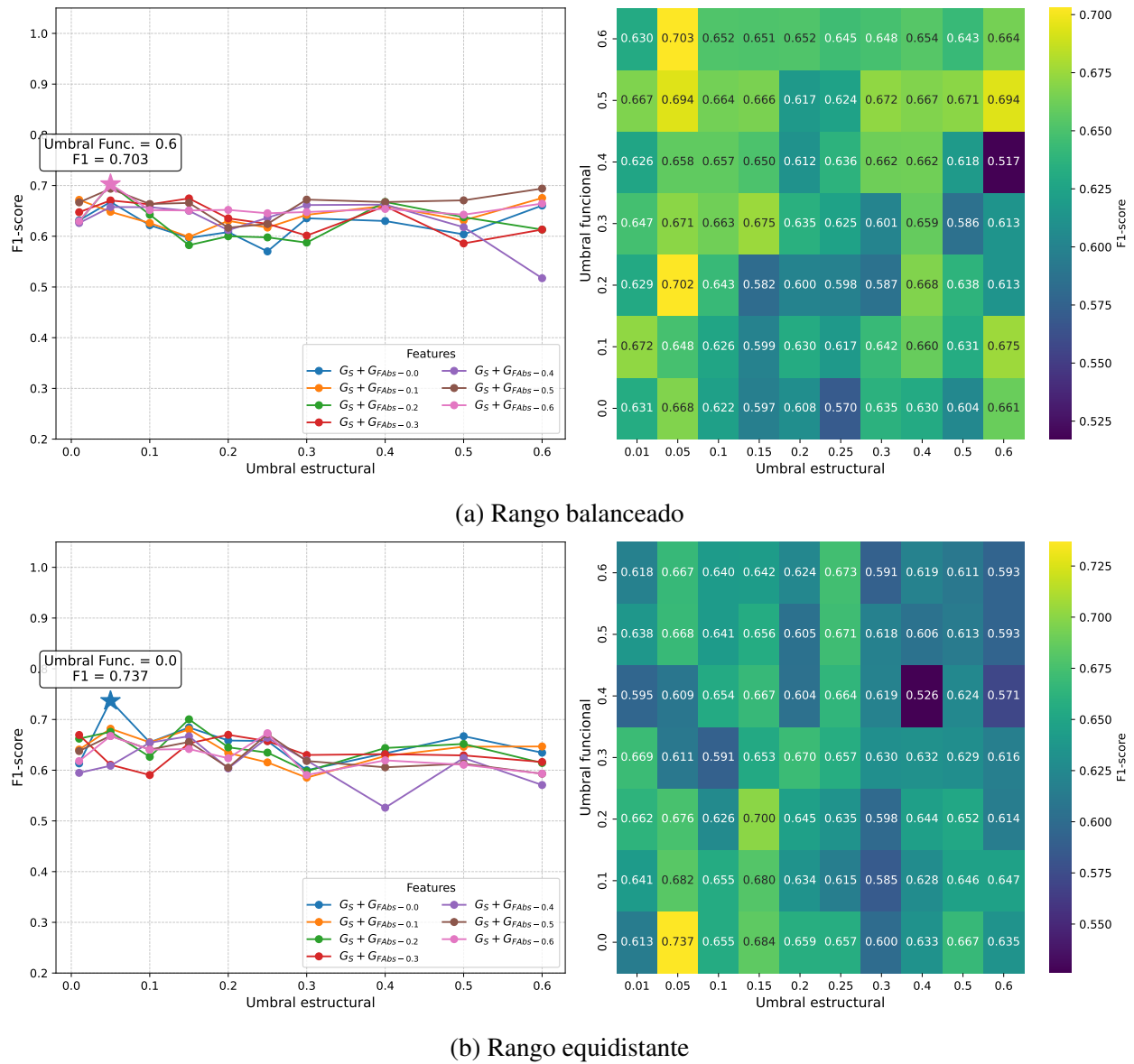


Figura 5.4: F1-score para los modelos de representación combinada basados en Support Vector Machine a través de diferentes umbrales estructurales.

5.3.4. Resumen de los resultados

La Tabla 5.2 agrupa los mejores resultados de cada modelo, agrupados en base al algoritmo en el que se basan y los rangos etarios que utilizan. La Tabla 5.1 complementa esta información, indicando las configuraciones usadas para cada uno de los resultados, respectivamente.

La distribución etaria balanceada muestra la clasificación mas baja, lo que está principalmente

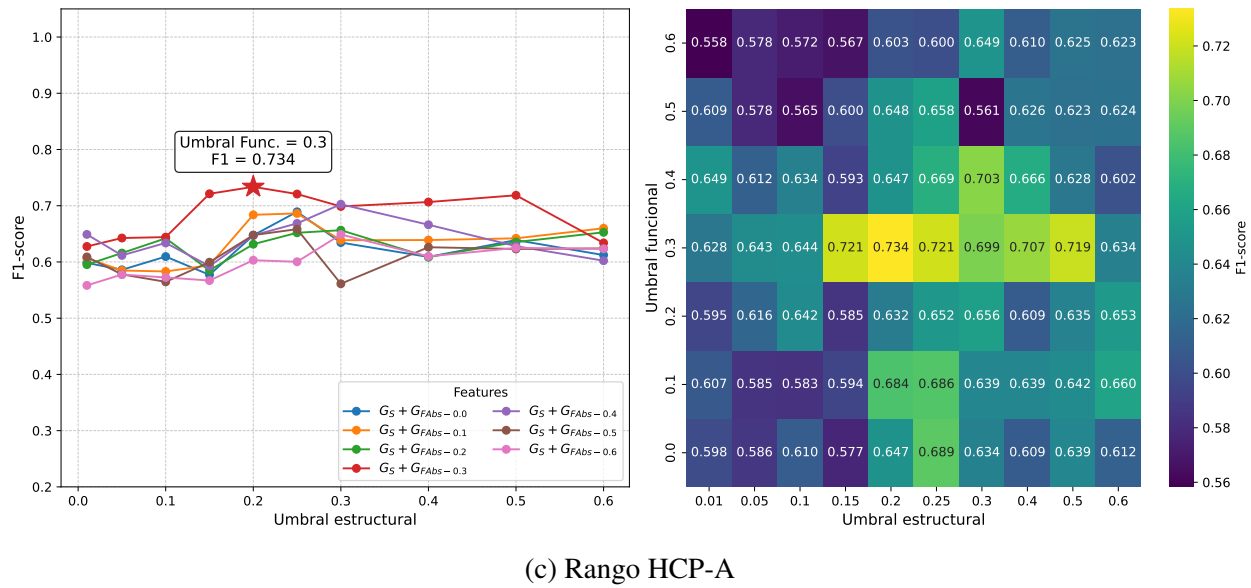


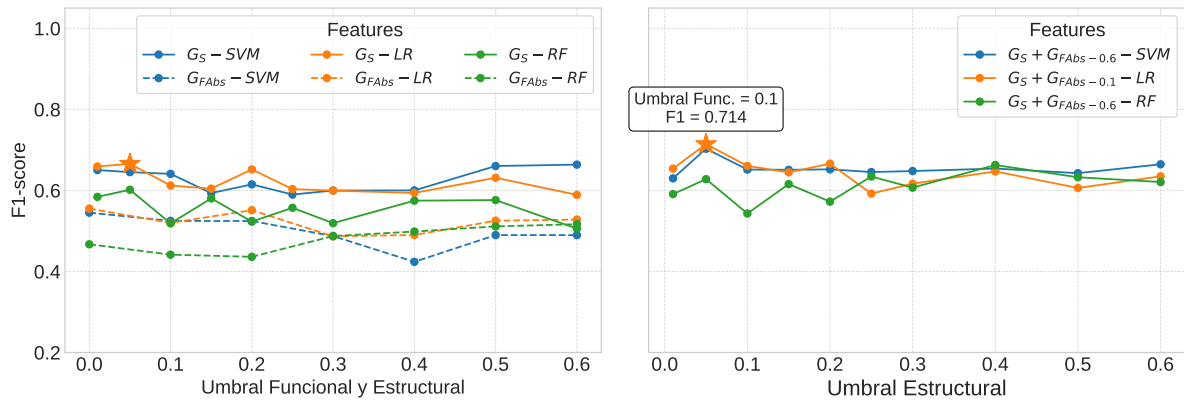
Figura 5.4: F1-score para los modelos de representación combinada basados en Support Vector Machine a través de diferentes umbrales estructurales.

determinado por la mala separación de los dos primeros grupos etarios. Se observa una reducción considerable en las medidas respecto a HCP-A y equidistante.

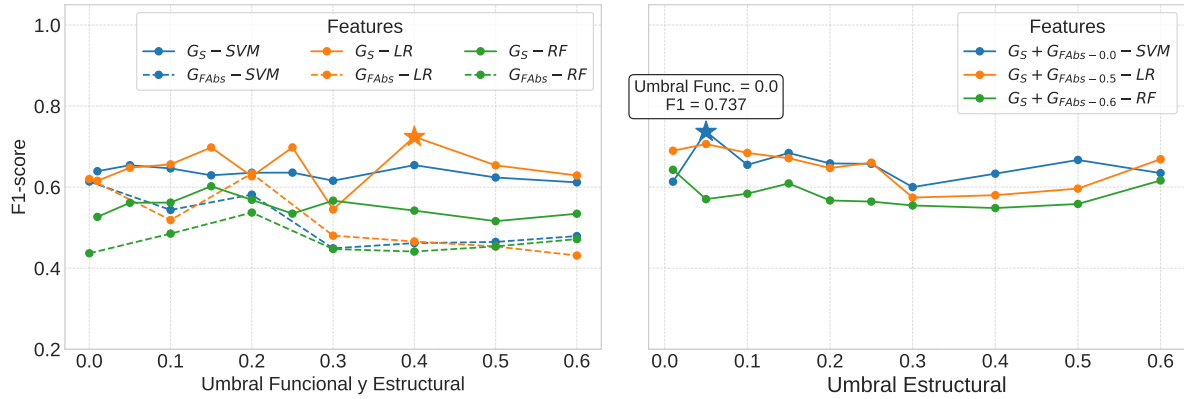
Finalmente, la Figura 5.5 muestra la comparativa más directa entre los mejores modelos de una única característica y aquellos que utilizan la combinación de estas.

Cuadro 5.1: Configuración de los resultados de la Tabla 5.2 en su orden respectivo.

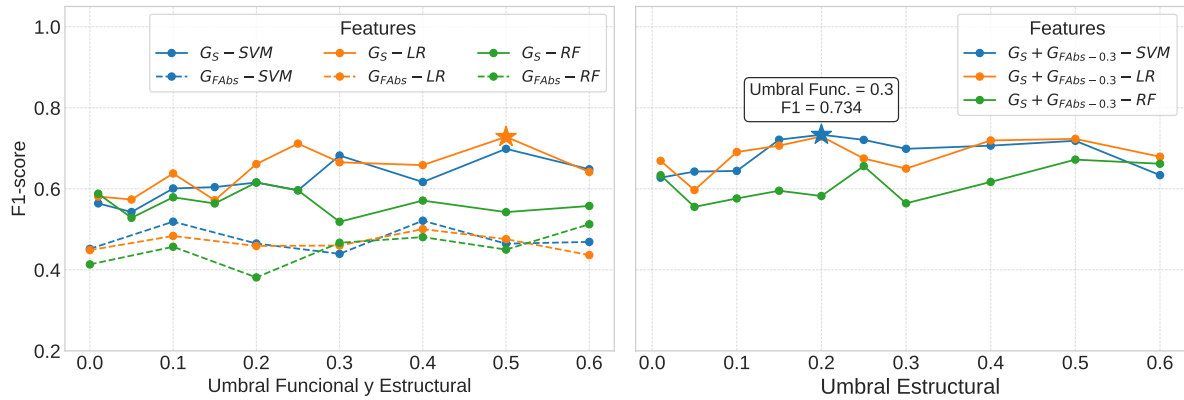
Modelo	Equidistante	HCP-A	Balanceado
RF	Umbral estructural: 0.60 Umbral funcional: 0.01	Umbral estructural: 0.50 Umbral funcional: 0.30	Umbral estructural: 0.40 Umbral funcional: 0.60
LR	Umbral estructural: 0.05 Umbral funcional: 0.50	Umbral estructural: 0.20 Umbral funcional: 0.30	Umbral estructural: 0.05 Umbral funcional: 0.10
SVM	Umbral estructural: 0.05 Umbral funcional: 0.00	Umbral estructural: 0.20 Umbral funcional: 0.30	Umbral estructural: 0.05 Umbral funcional: 0.60



(a) Rangos de edad balanceados.



(b) Rangos de edad equidistantes.



(c) Rangos de edad HCP-A.

Figura 5.5: Puntuaciones F1 para modelos SVM, RF y LR con diferentes umbrales aplicados a las características estructurales y funcionales. La columna izquierda muestra las puntuaciones F1 utilizando solo 34 características estructurales o 34 funcionales. La columna derecha muestra las puntuaciones F1 con las 68 características combinadas.

Cuadro 5.2: Mejores resultados en modelos tipo 2 (combinación de características estructurales y funcionales) desagregados por métrica de evaluación y clases. Se destacan en negrita los mejores resultados de cada modelo en cada rango.

Modelo	Equidistante				HCP-A				Balanced			
	Clase	Prec.	Rec.	F1	Clase	Prec.	Rec.	F1	Clase	Prec.	Rec.	F1
RF	[36, 57)	0.74	0.80	0.77	[36, 65)	0.74	0.89	0.81	[36, 48)	0.69	0.85	0.76
	[57, 78)	0.53	0.57	0.55	[65, 80)	0.52	0.37	0.43	[48, 70)	0.57	0.47	0.52
	[78, 101)	0.72	0.52	0.60	[80, 101)	0.83	0.71	0.76	[70, 101)	0.72	0.67	0.70
	macro avg	0.66	0.63	0.64	macro avg	0.70	0.65	0.67	macro avg	0.66	0.66	0.66
LR	[36, 57)	0.83	0.85	0.84	[36, 65)	0.86	0.86	0.86	[36, 48)	0.72	0.73	0.72
	[57, 78)	0.61	0.61	0.61	[65, 80)	0.58	0.57	0.57	[48, 70)	0.58	0.60	0.59
	[78, 101)	0.66	0.64	0.65	[80, 101)	0.72	0.76	0.74	[70, 101)	0.84	0.80	0.82
	macro avg	0.70	0.70	0.70	macro avg	0.72	0.73	0.72	macro avg	0.71	0.71	0.71
SVM	[36, 57)	0.80	0.89	0.84	[36, 65)	0.83	0.86	0.85	[36, 48)	0.69	0.80	0.74
	[57, 78)	0.66	0.61	0.64	[65, 80)	0.60	0.60	0.60	[48, 70)	0.56	0.52	0.54
	[78, 101)	0.77	0.68	0.72	[80, 101)	0.78	0.71	0.75	[70, 101)	0.86	0.77	0.81
	macro avg	0.74	0.72	0.73	macro avg	0.74	0.72	0.73	macro avg	0.70	0.70	0.70

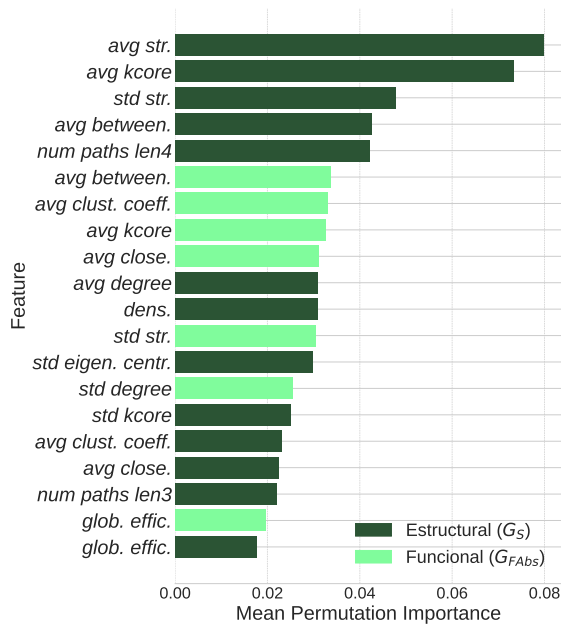
5.4. Reducción de características

Para cada modelo se calculó el *feature importance* como se plantea en la Sección 4.2.7. Este método resultó en un conjunto de 20 características para SVM, otro para LR y un último para RF. La Figura 5.6 muestra las 20 características más importantes para cada uno de los algoritmos. Finalmente, para el reentrenamiento se consideraron las 10, 15 y 20 características más importantes, con el objetivo de evaluar el comportamiento al reducir el *input* de los modelos a diferentes tamaños.

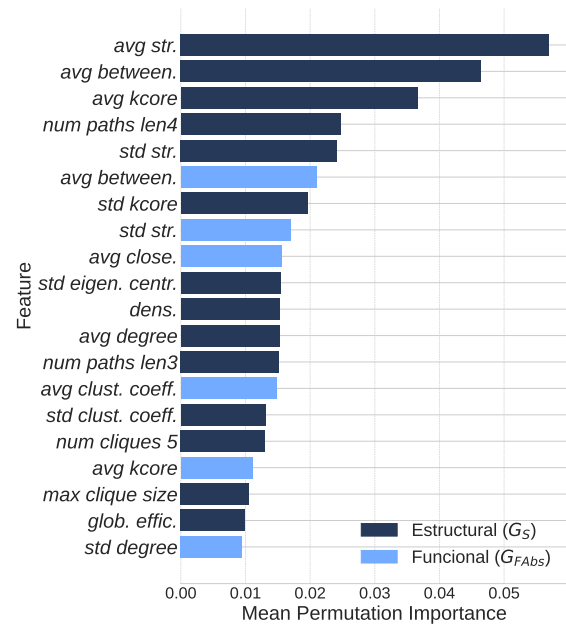
5.5. Re-entrenamiento de modelos

Con las características seleccionadas, se realizó un reentrenamiento para todos los modelos que combinaban medidas estructurales y funcionales. Luego estos modelos (a los que podemos llamar modelos tipo 3), fueron comparados con sus contrapartes que consideraban el total de características para evaluar el impacto de la reducción de características.

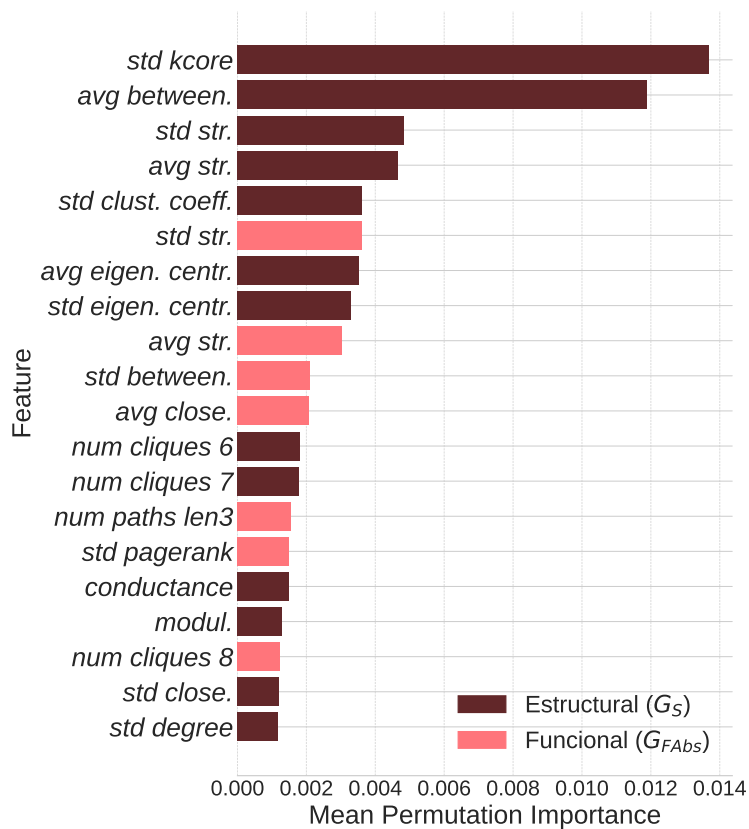
Los reentrenamientos consideraron en primer lugar 20 características, para luego avanzar a conjuntos menores de 15 y 10 características. La unión de estos resultados con la totalidad de experimentos confirmaron los malos resultados de los modelos basados en Random Forest. Por este motivo las siguientes secciones se centran principalmente en los modelos basados en LR y SVM.



(a) Top 20 características en LR.



(b) Top 20 características en SVM.



(c) Top 20 características en RF.

Figura 5.6: 20 características más relevantes en promedio para diferentes algoritmos, ordenadas por importancia.

En general para el resto de modelos se observan resultados variables, mejorando principalmente sobre rangos equidistantes.

5.5.1. Re-entrenamiento de LR

La Figura 5.7 muestra los mejores modelos basados en LR utilizando conjuntos seleccionados de características. El mejor resultado se obtiene sobre el rango equidistante utilizando 20 características, con un F1-Score de 0.759. Si bien se observa un mejor desempeño, esta mejora en los resultados varía de acuerdo al umbral aplicado. Las mejoras más notorias son sobre el rango equidistante.

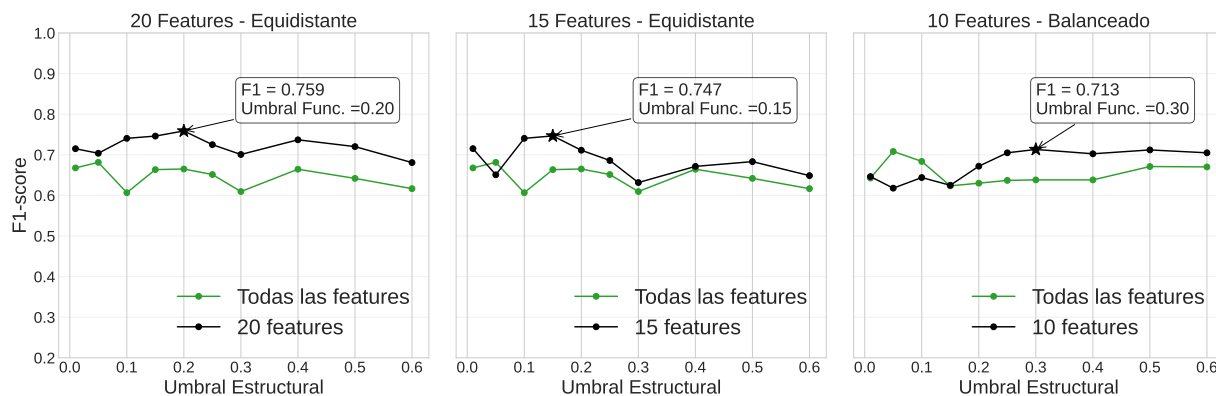


Figura 5.7: Mejores resultados de modelos reentrenados basados en LR.

5.5.2. Re-entrenamiento de SVM

La Figura 5.8 muestra los mejores modelos basados en SVM utilizando conjuntos seleccionados de características. El mejor resultado de todo el reentrenamiento se obtiene sobre el rango HCP-A utilizando 15 características, con un F1-Score de 0.764. Los modelos parecen encontrar el óptimo desempeño en 15 características, pues los resultados tanto con 10 como con 20 tienen peores métricas de evaluación.

5.5.3. Resumen de los resultados luego de la reducción

La Tabla 5.3 profundiza en los mejores resultados de cada modelo, para todos los sets reducidos de características. Como se mencionaba anteriormente, los modelos de SVM parecen llegar a un óptimo local de desempeño en 15 características. Por otro lado, los modelos de LR crecen

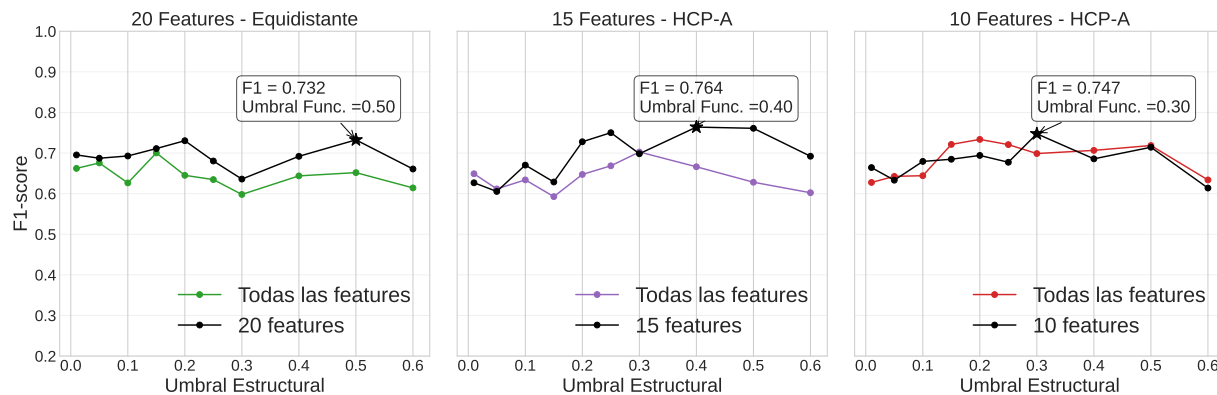


Figura 5.8: Mejores resultados de modelos reentrenados basados en SVM.

en desempeño a mediad que se aumentan las características, sugiriendo que la inclusión de más características podría mejorar los resultados.

Cuadro 5.3: Resultados de los modelos SVM y LR utilizando diferentes tamaños en los conjuntos de características, para distintos rangos etarios.

Modelo	Número de características	Rango	Macro-F1	Joven			Medio			Mayor		
				Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
SVM	10	HCP-A	0.747	0.826	0.939	0.879	0.645	0.571	0.606	0.875	0.666	0.756
SVM	15	HCP-A	0.764	0.828	0.954	0.887	0.739	0.485	0.586	0.782	0.857	0.818
SVM	20	Equidistant	0.732	0.803	0.818	0.810	0.634	0.619	0.626	0.760	0.760	0.760
LR	10	Balanced	0.713	0.717	0.785	0.750	0.600	0.600	0.600	0.833	0.750	0.789
LR	15	Equidistant	0.747	0.803	0.890	0.844	0.727	0.517	0.640	0.714	0.800	0.754
LR	20	Equidistant	0.759	0.818	0.818	0.818	0.674	0.690	0.682	0.791	0.760	0.775

5.6. Regresores

Con los resultados del reentrenamiento se utilizó la configuración del mejor modelo obtenido para entrenar dos modelos de regresión. El objetivo es verificar la capacidad de las características para entregar información acerca de la red de conectividad y obtener una métrica análoga del desempeño de los modelos entrenados. Esto se hace para poder realizar una comparación con los modelos de regresión presentes en la literatura.

El clasificador que obtuvo los mejores resultados fue el basado en SVM, por lo que el primer modelo considerado fue el regresor análogo de Support Vector Regressor. Para realizar una comparativa, se consideró también un regresor lineal con regularizador de LASSO, el cual fue elegido

por la naturaleza simple del modelo.

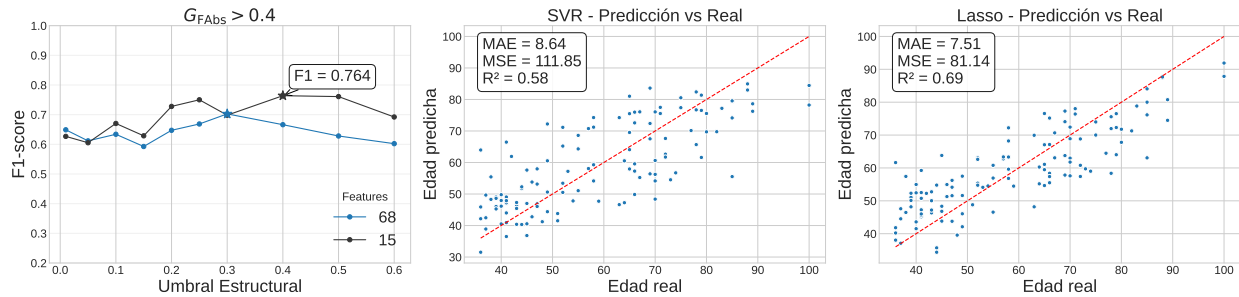


Figura 5.9: Resultados de modelos de regresión. A la izquierda, el mejor modelo de clasificación. Al centro, el modelo de regresión basado en SVR. A la derecha, el modelo de regresión basado en LASSO.

La Figura 5.9 muestra los resultados del mejor clasificador y ambos regresores junto a sus métricas de evaluación. Observamos un MAE de 7.51 en el modelo LASSO, estableciéndolo como el mejor regresor, mientras que su contraparte de SVR obtiene un MAE de 8.64. Esto indica que un buen desempeño en los modelos de clasificación no necesariamente se condice con un buen desempeño en regresión. Los valores obtenidos en el MAE de ambos modelos se pueden atribuir a la existencia del *brain age gap*, fenómeno que afecta a la predicción de la edad, distorsionando el *ground truth* utilizado.

Las características usadas para el mejor modelo de clasificación y los modelos de regresión se encuentran en la Figura 5.10. Se incluyen además las 3 métricas más importantes y su comportamiento en los sujetos de acuerdo a su edad con más detalle en la Figura 5.11.

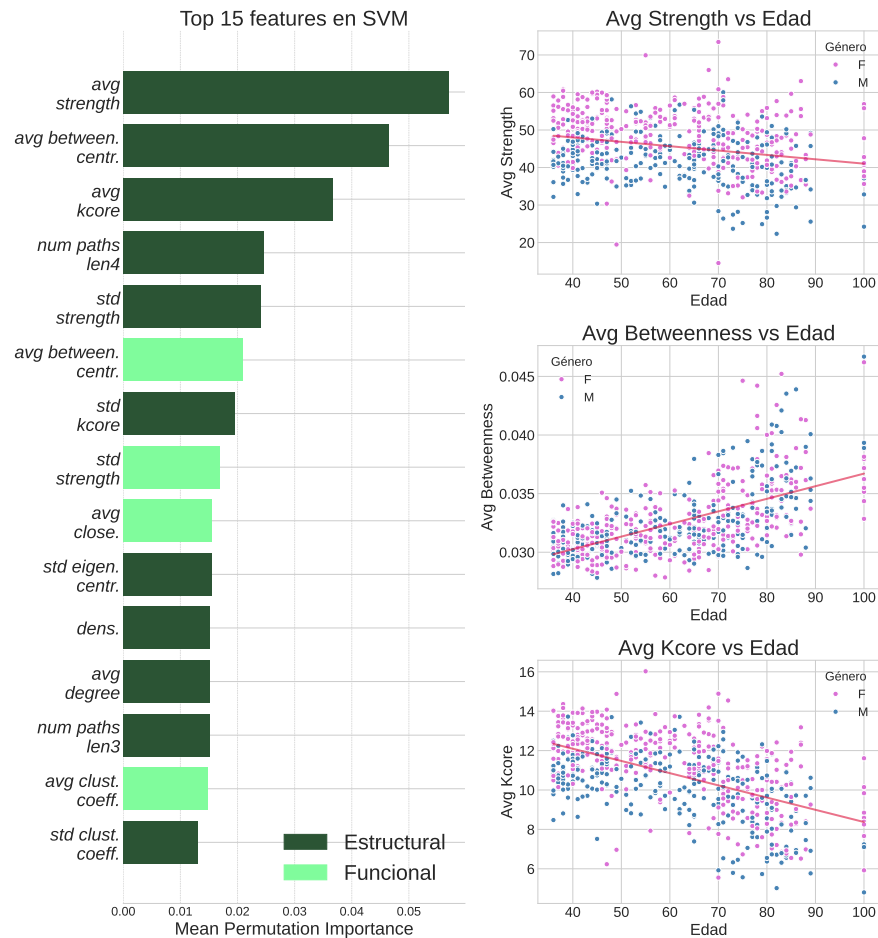


Figura 5.10: Características utilizadas por el modelo de regresión y su importancia previamente calculada. A la derecha, mejores 3 métricas y su comportamiento en los sujetos junto a su género.

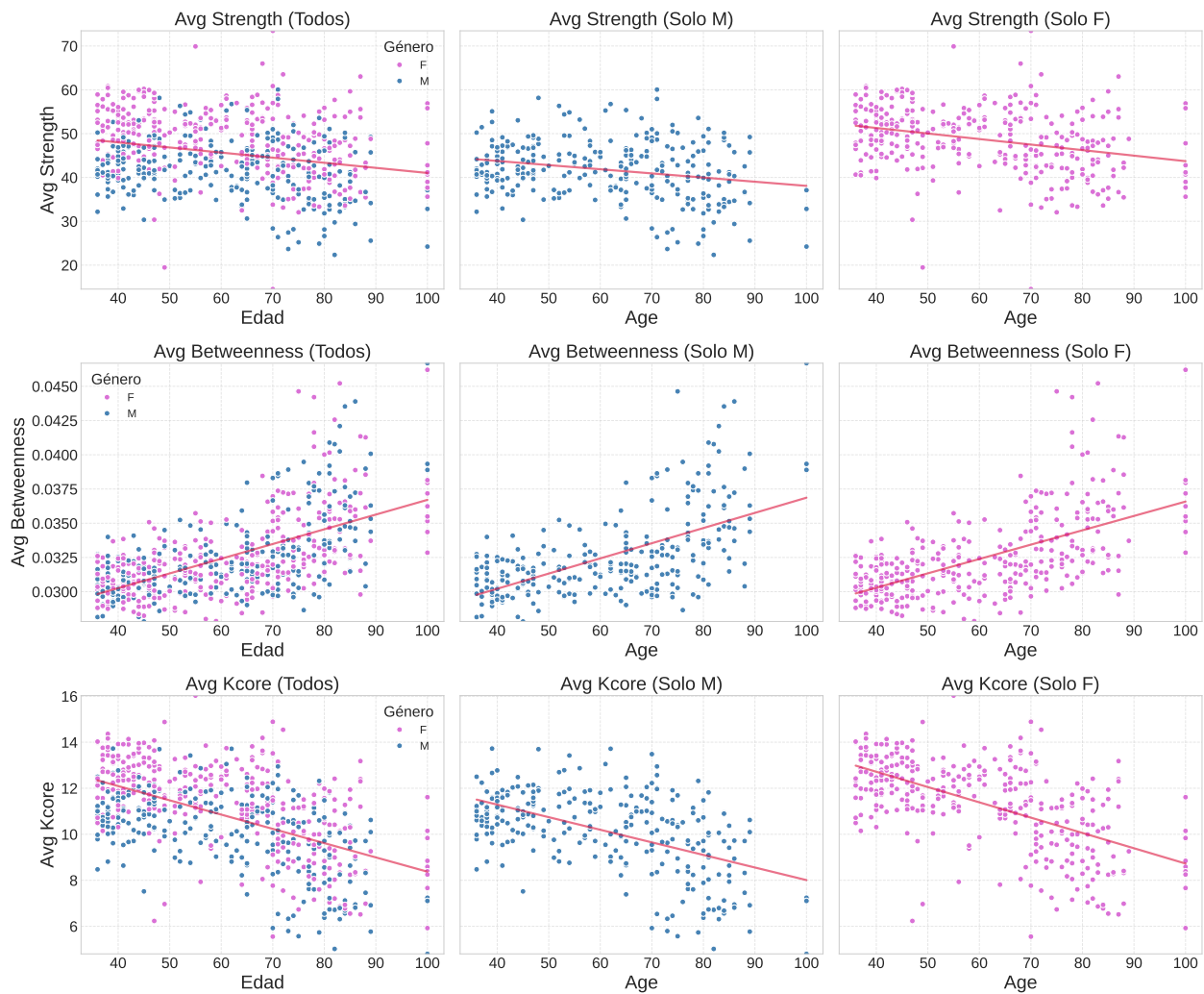


Figura 5.11: Comportamiento de las 3 métricas más importantes utilizadas por los regresores, separadas por género.

Capítulo 6

Conclusiones

La evaluación experimental expuesta en el capítulo anterior confirma la mejora en el rendimiento de los modelos de clasificación al combinar medidas de conectividad que caractericen a las redes estructural y funcional. Particularmente, se observa un comportamiento estable en la capacidad de las métricas estructurales de capturar información relevante, mientras que las medidas funcionales toman un papel secundario, brindando información adicional que mejora los resultados.

Notamos una relación significativa entre los resultados de clasificación y la división etaria utilizada para clasificar. Utilizar una división que mantuviese un balance de clases, es decir, la distribución etaria balanceada resultó en peores clasificaciones a nivel general para todos los modelos. Esto puede atribuirse a la arbitrariedad en la división, con rangos que fallan en representar ciertas etapas del desarrollo humano. Como ejemplo, se tiene el rango que comprende a los ‘adultos medios’ en la Figura 4.5 que termina comprendiendo 22 años, pero que carece de significado. En cambio, los mejores resultados son obtenidos por los rangos equidistantes y HCP-A, que toman una división más natural de la vida humana. Ambos rangos son similares a la hora de definir los grupos etarios, con la principal diferencia en el grupo más joven. Resulta entonces natural que ambos rangos compartan buenos resultados.

Ahora, en lo que a modelos respecta, también existe una notoria diferencia entre los modelos Random Forest y SVM o Logistic Regression. Observando los resultados expuestos en la Tabla 5.2, se concluye que las clasificaciones realizadas por Random Forest son deficientes comparadas con los otros modelos en un mismo rango. Esto se atribuye a que la estructura de Random Forest es poco eficiente para modelar los límites de decisión lineales que parecen caracterizar a los datos.

Para el caso particular del rango equidistante, se observaron mejorías sostenidas al reducir el número de características en todos los modelos. Esto indica la existencia de características más relevantes que otras, y que la identificación de las mismas puede ayudar a capturar información

igual de relevante en menos espacio. La utilización de estas características permite que los modelos clasifiquen de mejor manera a los sujetos para este rango. Trabajos posteriores deberían centrarse en el entendimiento de estas características y la posibilidad de complementarlas con características de otra naturaleza.

Los resultados de los regresores nos permiten asegurar que nuestros modelos de clasificación no se encuentran sobreajustados. Podemos interpretar el error presente en los regresores como el efecto del *brain age gap*. Aún con esa distorsión presente nuestros resultados superan el MAE de 8.89 presentado en el trabajo de Feng et al. [37]. Esta comparativa es relevante pues existen similitudes en la metodología utilizada en ambos trabajos.

Aun con todo, no se ha superado el umbral de 0.80 en la métrica de F1-Score (valor sobre el los resultados se consideran buenos), existiendo un gran margen de mejora para lograr un resultado importante. La clave para dar el paso faltante puede residir, como se ha mencionado antes, en la inclusión de nuevas características, como pueden ser aquellas más ligadas a la anatomía cerebral (espesor cortical, superficie del área cortical, entre otras) [20], relacionadas con la estadística (media, varianza, mediana de las señales rs-fMRI) [25].

Finalmente, se destaca que se cumplieron todos los objetivos planteados y además se incorporó el uso de regresores para estimar la edad en base a las características mas relevantes. Destacando el alcance de esta memoria de título, los resultados del trabajo fueron expuestos en la Jornada Técnica 2025 del AC3E en la Universidad Técnica Federico Santa María y el envío de un artículo a la conferencia International Symposium on Medical Information Processing and Analysis (SIPAIM).

Bibliografía

- [1] G. R. Boss y J. E. Seegmiller. “Age-Related Physiological Changes and Their Clinical Significance”. En: *Western Journal of Medicine* 135 (6 1981), pág. 434. ISSN: 00930415.
- [2] Ralf Th Krampe. “Aging, expertise and fine motor movement”. En: *Neuroscience & Bio-behavioral Reviews* 26 (7 nov. de 2002), págs. 769-776. ISSN: 0149-7634.
- [3] Cheryl Grady. “The cognitive neuroscience of ageing”. En: *Nature Reviews Neuroscience* 2012 13:7 13 (7 jun. de 2012), págs. 491-505. ISSN: 1471-0048.
- [4] Katja Franke et al. “Advanced BrainAGE in older adults with type 2 diabetes mellitus”. En: *Frontiers in Aging Neuroscience* 5 (DEC 2013). ISSN: 16634365.
- [5] Dennis M. Hedderich et al. “Increased Brain Age Gap Estimate (BrainAGE) in Young Adults After Premature Birth”. En: *Frontiers in Aging Neuroscience* 13 (2021). ISSN: 16634365.
- [6] Gary H. Glover. *Overview of functional magnetic resonance imaging*. 2011.
- [7] Susumu Mori. “Diffusion magnetic resonance imaging: Its principle and applications”. En: *Anatomical Record* 257 (3 1999). ISSN: 0003276X.
- [8] Casey C. Armstrong et al. “Graph-theoretical analysis of resting-state fMRI in pediatric obsessive-compulsive disorder”. En: *Journal of Affective Disorders* 193 (2016). ISSN: 15732517.
- [9] Ali Khazaei, Ata Ebrahimzadeh y Abbas Babajani-Feremi. “Identifying patients with Alzheimer’s disease using resting-state fMRI and graph theory”. En: *Clinical Neurophysiology* 126 (11 2015). ISSN: 18728952.
- [10] Jiayuan Huang et al. “Multimodal Magnetic Resonance Imaging Reveals Aberrant Brain Age Trajectory During Youth in Schizophrenia Patients”. En: *Frontiers in Aging Neuroscience* 14 (2022). ISSN: 16634365.
- [11] Lucina Q. Uddin et al. “Network homogeneity reveals decreased integrity of default-mode network in ADHD”. En: *Journal of Neuroscience Methods* 169 (1 mar. de 2008), págs. 249-254. ISSN: 0165-0270.
- [12] Katja Franke y Christian Gaser. *Ten years of brainage as a neuroimaging biomarker of brain aging: What insights have we gained?* 2019.

- [13] L. K. Soumya Kumari y R. Sundarajan. “A review on brain age prediction models”. En: *Brain Research* 1823 (ene. de 2024), pág. 148668. ISSN: 0006-8993.
- [14] Clifford R. Jack et al. *The Alzheimer’s Disease Neuroimaging Initiative (ADNI): MRI methods*. 2008.
- [15] Lea Baecker et al. “Machine learning for brain age prediction: Introduction to methods and clinical applications”. En: *eBioMedicine* 72 (oct. de 2021). Usa MRI estructural (T1), la idea es que use dMRI o fMRI., pág. 103600. ISSN: 2352-3964.
- [16] Shammi More et al. “Brain-age prediction: A systematic comparison of machine learning workflows”. En: *NeuroImage* 270 (abr. de 2023), pág. 119947. ISSN: 1053-8119.
- [17] Chang Le Chen et al. “Generalization of diffusion magnetic resonance imaging–based brain age prediction model through transfer learning”. En: *NeuroImage* 217 (2020). ISSN: 10959572.
- [18] Jose Ramon Chang et al. “Age prediction using resting-state functional MRI”. En: *medRxiv* (dic. de 2023), pág. 2023.12.26.23300530.
- [19] Sepideh Baghernezhad y Mohammad Reza Daliri. “Age-related changes in human brain functional connectivity using graph theory and machine learning techniques in resting-state fMRI data”. En: *GeroScience* 46 (5 oct. de 2024), págs. 5303-5320. ISSN: 25092723.
- [20] Denis Alexander Engemann et al. “Combining magnetoencephalography with magnetic resonance imaging enhances learning of surrogate-biomarkers”. En: *eLife* 9 (mayo de 2020), págs. 1-33. ISSN: 2050084X.
- [21] Kristine B. Walhovd, Martin Lövden y Anders M. Fjell. *Timing of lifespan influences on brain and cognition*. 2023.
- [22] Melis Anatürk et al. “Prediction of brain age and cognitive age: Quantifying brain and cognitive maintenance in aging”. En: *Human Brain Mapping* 42 (6 2021). ISSN: 10970193.
- [23] Jessica S. Damoiseaux. “Effects of aging on functional and structural brain connectivity”. En: *NeuroImage* 160 (2017). ISSN: 10959572.
- [24] Franziskus Liem et al. “Predicting brain-age from multimodal imaging data captures cognitive impairment”. En: *NeuroImage* 148 (2017). ISSN: 10959572.
- [25] Siamak K. Sorooshyari. “Beyond network connectivity: A classification approach to brain age prediction with resting-state fMRI”. En: *NeuroImage* 290 (abr. de 2024). ISSN: 10959572.
- [26] Colin R. Buchanan et al. “The effect of network thresholding and weighting on structural brain networks in the UK Biobank”. En: *NeuroImage* 211 (2020). ISSN: 10959572.

- [27] Hongfang Han, Sheng Ge y Haixian Wang. “Prediction of brain age based on the community structure of functional networks”. En: *Biomedical Signal Processing and Control* 79 (2023). ISSN: 17468108.
- [28] Derek L. Hansen et al. “Calculating and visualizing network metrics”. En: *Analyzing Social Media Networks with NodeXL: Insights from a Connected World* (ene. de 2020), págs. 79-94.
- [29] Jennifer Golbeck. “Network Structure and Measures”. En: *Analyzing the Social Web* (ene. de 2013), págs. 25-44.
- [30] *NetworkX — NetworkX documentation*.
- [31] Vito Latora y Massimo Marchiori. “Efficient Behavior of Small-World Networks”. En: *Physical Review Letters* 87 (19 oct. de 2001), pág. 198701. ISSN: 10797114.
- [32] David Eppstein, Maarten Löffler y Darren Strash. “Listing all maximal cliques in large sparse real-world graphs in near-optimal time”. En: *ACM Journal of Experimental Algorithmics* 18 (nov. de 2013). ISSN: 10846654.
- [33] Jean Paul Rodrigue. “The geography of transport systems”. En: *The Geography of Transport Systems* (abr. de 2024), págs. 1-402.
- [34] Tasmay Pankaj Tibrewal. *Support Vector Machines (SVM): An Intuitive Explanation*. Jul. de 2023.
- [35] *Scikit-Learn - 1.7.0 documentation*.
- [36] Will Koehrsen. *Random Forest Simple Explanation*. Dic. de 2017.
- [37] Guozheng Feng et al. “Methodological evaluation of individual cognitive prediction based on the brain white matter structural connectome”. En: *Human Brain Mapping* 43 (12 2022). ISSN: 10970193.
- [38] Eunji Kim et al. “Connectome-based predictive models using resting-state fMRI for studying brain aging”. En: *Experimental Brain Research* 240 (9 sep. de 2022), págs. 2389-2400. ISSN: 14321106.
- [39] Julie Gonneaud et al. “Accelerated functional brain aging in pre-clinical familial Alzheimer’s disease”. En: *Nature Communications* 2021 12:1 12 (1 sep. de 2021), págs. 1-17. ISSN: 2041-1723.
- [40] Susan Y. Bookheimer et al. “The Lifespan Human Connectome Project in Aging: An overview”. En: *NeuroImage* 185 (2019). ISSN: 10959572.

- [41] Michael P. Harms et al. "Extending the Human Connectome Project across ages: Imaging protocols for the Lifespan Development and Aging projects". En: *NeuroImage* 183 (2018). ISSN: 10959572.
- [42] Alexander Schaefer et al. "Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI". En: *Cerebral Cortex* 28 (9 2018). ISSN: 1047-3211.
- [43] Ye Tian et al. "Topographic organization of the human subcortex unveiled with functional connectivity gradients". En: *Nature Neuroscience* 23 (11 2020). ISSN: 15461726.
- [44] Matthew F. Glasser et al. "The minimal preprocessing pipelines for the Human Connectome Project". En: *NeuroImage* 80 (2013). ISSN: 10538119.
- [45] Mark Jenkinson et al. "Review FSL". En: *NeuroImage* 62 (2012).
- [46] Bruce Fischl. "FreeSurfer". En: *NeuroImage* 62 (2 ago. de 2012), págs. 774-781. ISSN: 1053-8119.
- [47] François Tadel et al. "MEG/EEG group analysis with brainstorm". En: *Frontiers in Neuroscience* 13 (FEB 2019). ISSN: 1662453X.
- [48] Michele Larobina y Loredana Murino. *Medical image file formats*. 2014.
- [49] J. Donald Tournier, Fernando Calamante y Alan Connelly. "MRtrix: Diffusion tractography in crossing fiber regions". En: *International Journal of Imaging Systems and Technology* 22 (1 2012). ISSN: 08999457.
- [50] Robert E. Smith et al. "SIFT2: Enabling dense quantitative assessment of brain white matter connectivity using streamlines tractography". En: *NeuroImage* 119 (2015). ISSN: 10959572.
- [51] Daniel S. Marcus et al. "Informatics and data mining tools and strategies for the human connectome project". En: *Frontiers in Neuroinformatics* 5 (2011). ISSN: 16625196.
- [52] Adam J. Schwarz y John McGonigle. "Negative edges and soft thresholding in complex network analysis of resting state functional connectivity data". En: *NeuroImage* 55 (3 abr. de 2011), págs. 1132-1146. ISSN: 1095-9572.
- [53] Obada Al Zoubi et al. "Predicting age from brain EEG signals-a machine learning approach". En: *Frontiers in Aging Neuroscience* 10 (JUL 2018). ISSN: 16634365.