



UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA Y
DISEÑO INDUSTRIAL

Grado en Ingeniería Electrónica y Automática Industrial

TRABAJO FIN DE GRADO

**Detección precoz de la enfermedad de Alzheimer
en sujetos con deterioro cognitivo leve mediante
análisis longitudinal de imágenes de resonancia
magnética y de medidas cognitivas.**

Jaime Simarro Viana

Tutor: Carlos Platero Dueñas
Departamento: Ingeniería

Eléctrica, Electrónica,
Automática y Física aplicada

Madrid, Junio, 2018



UNIVERSIDAD POLITÉCNICA DE MADRID

ESCUELA TÉCNICA SUPERIOR DE INGENIERÍA Y
DISEÑO INDUSTRIAL

Grado en Ingeniería Electrónica y Automática Industrial

TRABAJO FIN DE GRADO

**Detección precoz de la enfermedad de Alzheimer
en sujetos con deterioro cognitivo leve mediante
análisis longitudinal de imágenes de resonancia
magnética y de medidas cognitivas**

Firma Autor

Firma Cotutor (si lo hay)

Firma Tutor

Copyright ©2018-Jaime Simarro Viana

Esta obra está licenciada bajo la licencia Creative Commons

Atribución-NoComercial-SinDerivadas 3.0 Unported (CC BY-NC-ND 3.0). Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-nc-nd/3.0/deed.es> o envíe una carta a Creative Commons, 444 Castro Street, Suite 900, Mountain View, California, 94041, EE.UU.

Todas las opiniones aquí expresadas son del autor, y no reflejan necesariamente las opiniones de la Universidad Politécnica de Madrid.

Título: Detección precoz de la enfermedad de Alzheimer en sujetos con deterioro cognitivo leve mediante análisis longitudinal de imágenes de resonancia magnética y de medidas cognitivas

Autor: Jaime Simarro Viana

Tutor: Carlos Platero Dueñas

EL TRIBUNAL

Presidente:

Vocal:

Secretario:

Realizado el acto de defensa y lectura del Trabajo Fin de Grado el día de de ... en, en la Escuela Técnica Superior de Ingeniería y Diseño Industrial de la Universidad Politécnica de Madrid, acuerda otorgarle la CALIFICACIÓN de:

VOCAL

SECRETARIO

PRESIDENTE

Agradecimientos

Debo hacer una mención especial a mi familia: Mama, Papa, Alba vosotros me habéis hecho ser la persona que soy.

A mis tíos, Paco y Maria Jesús por apoyarme y creer en mi.

A mi tutor Carlos Platero por mostrarme este maravilloso mundo de la neuroimagen.

A todos muchas gracias.

Resumen

En este proyecto se han buscado marcadores que permitan realizar un diagnóstico precoz de la enfermedad de Alzheimer. La investigación se ha focalizado en el sector de la población con mayor riesgo de padecer esta enfermedad, los sujetos con deterioro cognitivo leve, prediciendo si una persona de dicho grupo va a sufrir la enfermedad dentro de tres años. Para conseguir este ambicioso objetivo, se han propuesto marcadores provenientes de dos fuentes de información.

- La primera fuente se basa en la atrofia a nivel neuronal causada por esta enfermedad neurodegenerativa. Esta atrofia presenta un cambio en la morfología del cerebro que puede ser medido con Imágenes por Resonancia Magnética estructural. Se han utilizado las diferentes estructuras y subestructuras dadas por el atlas de Desikan-Killiany junto con un análisis masivo, el cual ha permitido detectar aquellas áreas corticales donde se sufre mayor deterioro.
- La segunda fuente de información se basa en el patente deterioro de las funciones cognitivas de los pacientes, las cuales son mostradas en las evaluaciones clínicas.

Estos marcadores son combinados y estudiados mediante un análisis longitudinal en dominio del tiempo. Para posteriormente ser utilizados por un clasificador de Análisis Discriminante Lineal.

La metodología propuesta permite obtener resultados tan prometedores como un 76,4% de precisión sobre el grupo de test con una sola visita, y hasta un 82,1% cuando el número de visitas aumenta a cuatro.

Palabras clave: Enfermedad de Alzheimer, Deterioro cognitivo leve, Imagen de resonancia magnética, Medidas cognitivas, Análisis longitudinal, Machine learning.

Abstract

In this project of investigation, we have been searching a group of markers which might be appropriate to make an early diagnosis of Alzheimer's Disease.

We have focused on the most risky group of population who suffer this disease, people with Mild Cognitive Impairment. We are able to predict if a person belonging to this group will be suffering Alzheimer during the next three years. This ambitious goal is achieved using two sources of data.

- The first source is based on the atrophy caused by this neurodegenerative disease. The atrophy shows a change in the human brain morphology that can be measured using Magnetic Resonance Imaging. The comparison of the different structures and substructures obtained from the Desikan-Killiany Atlas, as well as a mass-univariate analysis, have allowed us to allocate the cortical areas where the damage is more severe.
- The second source of data is based on the evident decline in cognitive functions, which are taken from the clinical assessment markers.

These markers are combined and studied using analysis of longitudinal data through-time domain, then have been processed by a classification algorithm named Linear Discriminant Analysis.

This approach can successfully achieve a classification accuracy of 76,4% with just one visit, while this accuracy rise to 82,1% when the number of visits increase to four.

Keywords: Alzheimer's Disease, Mild cognitive impairment, Magnetic Resonance Imaging, Clinical assessment markers, Analysis of longitudinal data, Machine learning.

Índice general

Agradecimientos	IX
Resumen	XI
Abstract	XIII
Índice	XVII
1. Introducción	1
1.1. Motivación del proyecto	1
1.2. Objetivos	2
1.3. Estructura del documento	3
2. Estado del arte	5
2.1. Enfermedad de Alzheimer	5
2.1.1. Neuropatología	5
2.1.2. Biomarcadores	6
2.2. Revisión problemática. Biomarcador MRI estructural	7
2.2.1. Método basado en mapas de densidad	7
2.2.2. Métodos basados en la posición de los vértices	8
2.2.3. Método basado en las regiones predefinidas	9
2.3. Problemática en los estudios longitudinales masivos	10
2.4. Revisión metodología de estudios similares	12
3. Materiales y herramientas utilizada	15
3.1. Base de datos	15
3.1.1. Datos demográficos de los sujetos pertenecientes en el ensayo .	17
3.2. Test neuropsicológicos	17
3.2.1. Evaluación neuropsicológica mínima	18
3.2.2. Evaluación neuropsicológica completa	20
3.3. Software	23
3.3.1. Matlab	23
3.3.2. Freesurfer	23
3.3.3. Filezilla	23
3.3.4. Conexión al servidor desde el bash de Ubuntu	24
3.3.5. Tmux Matlab -nodesktop	24
3.3.6. Archivos .sh	25

4. Métodos empleados	27
4.1. Etapas del desarrollo	27
4.2. Recogida de datos	27
4.3. Procesamiento de imágenes	27
4.3.1. Análisis Cross	29
4.3.2. Análisis Base	31
4.3.3. Análisis Long	32
4.4. Extracción de las características corticales y subcorticales	33
4.5. Control de calidad	33
4.5.1. Correlación entre pares de scans del mismo sujeto	33
4.5.2. Coeficiente de correlación de Pearson	33
4.5.3. Correlación de volumetría hipocampal y del valor de ICV	34
4.5.4. Comparación con los datos del grupo de investigación de Harvard	34
4.6. Obtención de los marcadores basados en MRI	34
4.6.1. Métodos de reducción basados en atlas	34
4.6.2. Métodos de reducción basados en el análisis masivo	34
4.7. Algoritmo de selección de características. Algoritmo mRMR, <i>minimum-Redundancy-Maximum-Relevance</i>	36
4.8. Obtención de los modelos y clasificación mediante residuos longitudinales	38
4.9. Liner Mixed Effects (LME)	38
4.9.1. Modelos LME basados en vóxel o vértices (V-LME)	39
4.9.2. Modelo LME espaciotemporal (ST-LME)	40
4.9.3. Residuos longitudinales	42
4.10. Realimentación entre selección de características y clasificación	43
4.10.1. Bootstrap	43
4.10.2. Validación cruzada (CV)	44
5. Resultados	49
5.1. Control de calidad	49
5.1.1. Correlación de volumetría hipocampal y del valor de ICV	49
5.1.2. Comparación con los datos del grupo de investigación de Harvard	50
5.2. Unión de los datos obtenidos este año con los procesamientos de otros años	53
5.3. Análisis masivo	57
5.3.1. sMCI vs pMCI	57
5.3.2. AD vs HC	60
5.3.3. Errores de tipo I y tipo II	62
5.4. Primera aproximación al modelo LME	62
5.5. Doble K-fold Cross Validation	67
5.5.1. Selección con MCI	67
5.5.2. Selección con HC y AD	70
5.6. Bootstap	70
5.6.1. Selección con HC y AD	72
5.6.2. Mejores marcadores de k-fold	72
5.7. DEMO del Software	72
5.7.1. DEMO extract features	73
5.7.2. DEMO machine learning	73

6. Discusiones y conclusiones	75
6.1. Contraste con otros estudios	78
6.2. Desarrollos futuros	79
A. Conceptos teóricos básicos	81
A.1. Imágenes de resonancia magnética	81
A.1.1. Glosario básico	81
A.1.2. Funcionamiento	82
A.1.3. Tipos de imágenes	83
A.1.4. Ventajas	85
A.2. Registro de imágenes	85
A.2.1. Transformaciones geométricas	85
A.2.2. Transformaciones espaciales	86
A.2.3. Transformaciones rígidas	86
A.2.4. Transformaciones no rígidas	86
A.3. Contraste de hipótesis	88
B. Acrónimos	91
Bibliografía	93

Índice de figuras

2.1. Atlas de Desikan-Killiany	9
2.2. Comparación entre métodos basados en la superficie (espesor) y métodos basados en densidad (volumen). Fuente Winkler(2010)	11
3.1. Subtipos de MCI. Fuente página web de ADNI	16
3.2. Flow-chart para la clasificación para sujetos MCI. Fuente http://adni.loni.usc.edu/	17
3.3. Tabla utilizada para el test CDR. Fuente https://www.hipocampo.org/hughes.asp	20
3.4. Tabla utilizada para el test FAQ. Fuente https://www.alz.org	21
3.5. Tabla utilizada para el test RAVTL. Fuente J Pediatr (Rio J). 2004;80(2 Suppl):S111-6	22
4.1. Resumen visual del procesado mediante recon all. Fuente https://surfer.nmr.mgh.harvard.edu	29
4.2. Cerebro humano representado donde se puede observar las coordenadas de Talairach	30
4.3. Resultado del proceso de skull strip	30
4.4. Diagrama de bloque de recon all. Fuente https://surfer.nmr.mgh.harvard.edu/	32
4.5. Esquema de la estrategia de doble k-fold CV	45
5.1. Coeficiente de correlación de Pearson	50
5.2. Coeficiente de correlación de los datos obtenidos con ADNIMERGE	50
5.3. Trayectorias de NVH de los sujetos analizados este año	51
5.4. Trayectorias de ETC de los sujetos analizados este año	51
5.5. Trayectorias de NHV separadas en función del número de scans. En todas las figuras el eje X representa los años, mientras que el eje Y representa la trayectoria del volumen del hipocampo normalizado	52
5.6. Trayectorias de NHV suavizadas en función del número de scans. El eje X representa los años, mientras que el eje Y representa la trayectoria del volumen del hipocampo normalizado	53
5.7. Distribución del NHV de los sujetos analizados este año	54
5.8. Distribución del ECT de los sujetos analizados este año	54
5.9. Trayectorias de NVH de todos los sujetos	55
5.10. Trayectorias de ETC de todos los sujetos	55
5.11. Distribución del NHV de todos los sujetos	56
5.12. Distribución del ECT de todos los sujetos	56

5.13. Mapa de q-valor entre sujetos sMCI y pMCI. De izquierda a derecha se muestra el hemisferio izquierdo y derecho	59
5.14. Mapa de los clusters realizados entre sujetos sMCI y pMCI.	59
5.15. Mapa de q-valor entre sujetos HC y AD. De izquierda a derecha se muestra el hemisferio izquierdo y derecho	61
5.16. Mapa de los clusters realizados entre sujetos HC y AD.	61
5.17. Error de tipo I de cada mascara. Se representa la media y las desviaciones típicas del error de tipo I de los 8 modelos. Los círculos representan el error del modelo final	62
5.18. Diferentes combinaciones de las covariables con características de MRI en todas las visitas	64
5.19. Diferentes combinaciones de las covariables con características de MRI junto con CAM en todas las visitas	65
5.20. Diferentes combinaciones de las covariables con características de MRI junto con CAM en el Baseline	66
5.21. Comparativa de dimensiones en doble k-fold Cross Validation seleccionando con MCI. Los marcadores incluyen datos de MRI basados en atlas, CAM y clusters de MCI. En todas las subfiguras el eje x representa el número de marcadores mientras que el eje y representa la precisión en el test.	68
5.22. Comparativa de visitas en doble k-fold Cross Validation seleccionando con MCI. Los marcadores incluyen datos de MRI basados en atlas, CAM y clusters de MCI. En todas las subfiguras el eje x representa el número de marcadores mientras que el eje y representa la precisión en el test	69
5.23. Comparativa de visitas en doble k-fold Cross Validation seleccionando con todos los sujetos. En todas las subfiguras el eje x representa el número de marcadores mientras que el eje y representa la precisión en el test	71
6.1. Trayectoria del marcador FAQ	75
6.2. Figura donde se muestran los cluster más significativos	76
A.1. Estructuras del cerebro humano	82
A.2. Forma de onda del seno cardinal	83
A.3. Resonancia magnética cerebral. Secuencias ponderadas en T2 (izquierda), T1 (centro) y T1 con contraste (derecha) en el plano transversal	84
A.4. Tipos de cortes	85
A.5. Ejemplo de una transformación proyectiva	88
A.6. Contraste entre dos distribuciones normales	89

Índice de tablas

3.1.	Protocolo de inclusión. Fuente http://adni.loni.usc.edu/	16
3.2.	Datos demográficos en el baseline de los sujetos sMCI y pMCI utilizados en el estudio	17
3.3.	Datos demográficos en el baseline de los sujetos AD y HC utilizados en el estudio	18
4.1.	Flow char del proyecto	28
5.1.	Tabla comparativa entre Bernal, Jaime y todos. Los datos se presentan junto con la desviación típica	57
5.2.	Tabla donde se muestra la relación entre los cluster formado a partir de los MCI y las diferentes estructuras del atlas de Desikan-Killiany	58
5.3.	Tabla donde se muestra la relación entre los cluster formado a partir de los AD y las diferentes estructuras del atlas de Desikan-Killiany	60
5.4.	Valor del poder estadístico en función de los diversos valores de α	62
5.5.	Tabla con los mejores marcadores seleccionado con MCI, utilizando solo una porción del grupo de MCI. Los resultados se muestran con un intervalo de confianza de 95 %	70
5.6.	Tabla con los mejores marcadores seleccionado con MCI con todos sujetos disponibles. Los resultados se muestran con un intervalo de confianza de 95 %	70
5.7.	Tabla de resultados de k-fold seleccionando con HC y AD. Los resultados se muestran con un intervalo de confianza de 95 %	72
5.8.	Tabla de resultados de bootstrap seleccionando con HC y AD. Los resultados se muestran con un intervalo de confianza de 95 %	72
5.9.	Tabla de resultados de bootstrap utilizando las mejores características del k-fold anterior, por ello se utiliza la otra fracción de los sujetos. Los resultados se muestran con un intervalo de confianza de 95 %	72
5.10.	Tabla de resultados reseñables de bootstrap utilizando características del k-fold anterior, por ello se utiliza la otra fracción de los sujetos. Los resultados se muestran con un intervalo de confianza de 95 %	73
6.1.	Tabla donde se muestra la relación entre los cluster más significativos y las diferentes estructuras del atlas de Desikan-Killiany	76
6.2.	Tabla de resultados con los marcadores ADAS13, FAQ y rh MCI el 1 y su evolución con el número de visitas. Los resultados se muestran con un intervalo de confianza de 95 %	78
6.3.	Tabla comparativa entre los estudios más reseñables en la literatura y este proyecto	79

A.1. Tabla comparativa de los tiempos de repetición y eco	83
A.2. Tabla comparativa entre las diferentes apariencias de diversos elementos según la ponderación utilizada	84
A.3. Tabla comparativa donde se muestran los tipos de errores	89
B.1. Tabla de acrónimos	91

Capítulo 1

Introducción

1.1. Motivación del proyecto

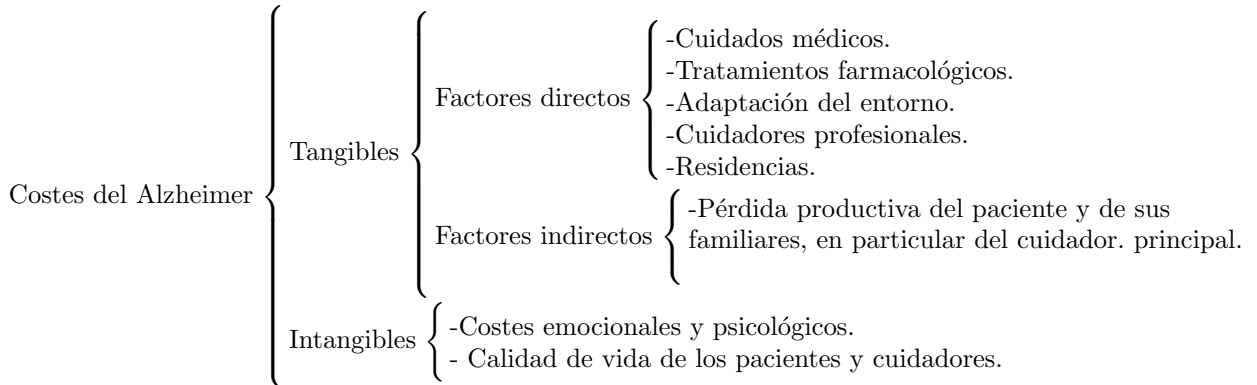
La motivación principal de este proyecto es minimizar el impacto económico y social causado por la enfermedad de Alzheimer. Para llevar esto a cabo, se propone un método preciso capaz de detectar el Alzheimer antes de que los síntomas aparezcan. Esta técnica podría ser utilizada en el seguimiento de la enfermedad y en la monitorización de los efectos causados por tratamientos, con especial importancia en ensayos clínicos con drogas.

La enfermedad de Alzheimer es una enfermedad neurodegenerativa que causa un deterioro progresivo de las funciones cognitivas. El inicio de la enfermedad tiene una gran variabilidad según el paciente, pero la mayoría de ellos permanecen asintomáticos hasta los 65 años. Desde los 65 años, el riesgo de padecer esta enfermedad se dobla cada 5 años[22]. Con 85 años el riesgo de sufrir un cognitivo significativo se acerca al 45 % [31]. La supervivencia aproximada desde el comienzo de esta enfermedad es de 8 años y el declive cognitivo común es de unos 3 puntos por año en el MMSE (*Mini-mental state examination* o Miniprueba de estado mental).

Esta enfermedad es la causa más frecuente de demencia, suponiendo entre un 50 % a 75 % del total [53]. La demencia provoca un aumento de la mortalidad y aumenta en gran medida la dependencia de los pacientes, los cuales ven disminuidas de manera considerable su calidad de vida.

La esperanza de vida de la población mundial esta aumentando, y por lo tanto la población anciana. Debido a que la edad es el principal factor de riesgo para esta enfermedad, el número de pacientes muestra una tendencia de crecimiento rápida. En 2006, el número de pacientes mundiales con Alzheimer era de 26.6 millones. La predicciones sugieren que este número va a aumentar a 106.8 millones en 2050 con 16.51 millones de pacientes en Europa [30]. Otra predicción muestra un incremento del 87 % de pacientes en Europa entre 2010 y 2050[40]. La enfermedad de Alzheimer se ha convertido en uno de los problemas más significativos de la salud pública, siendo la demencia la tercera enfermedad más costosa tras las enfermedades cardiovasculares y el cáncer. Desde el 2000, las muertes causadas por enfermedades cardiovasculares se han disminuido un 14 %, mientras que las muertes por Alzheimer aumentaron un 89 %.

El impacto social y económico de esta enfermedad es abrumador, tanto para los sistemas de salud como para los familiares y pacientes. Se aproxima el coste del tratamiento en España de 14 mil millones de euros al año, con un coste medio anual por paciente aproximado de 19.000 euros. Los costes de la demencia supone un 1% del PIB mundial. El siguiente ejemplo ayuda a escalar estos datos “*Si el cuidado de la demencia fuera un país ,sería la decimoctava economía más grande del mundo*”. De la misma manera existe un gran coste intangible; el siguiente diagrama 1.1 esquematiza los costes de esta enfermedad.



Se conoce que la principal barrera para implementar terapias y en consecuencia disminuir el coste de esta enfermedad es la dificultad en el diagnóstico, por ello en este proyecto se utilizan técnicas de vanguardia para realizar un **diagnóstico precoz**.

1.2. Objetivos

Desde 1990 se conoce que el hipocampo se ve afectado de manera temprana y severa en el proceso patológico del Alzheimer [6]. El objetivo de este proyecto es determinar un conjunto de marcadores que tenga un alto grado de sensibilidad y especificidad para la detección precoz del Alzheimer, dando así a los investigadores y a los clínicos una herramienta que permita monitorizar nuevos tratamientos. Estos marcadores deben de ser ampliamente accesibles, costo-eficientes y no ser invasivos. En consecuencia, los grupos de marcadores potenciales son:

1. Los factores de riesgo clínicos (*Clinical Risk Factors, CRF*) entre los que se incluye la edad, el género [32], los años de educación [23] y el gen APOE-4 [20].
2. Evaluaciones clínicas cognitivas y funcionales (*Clinical Assessments Markers, CAM*).
3. Medidas de morfología procedentes de imágenes de resonancia magnética estructural.

Para conseguir el fin previamente nombrado se han definido los siguientes subobjetivos:

1. Extracción de las características corticales y subcorticales desde imágenes de resonancia magnética ponderadas en T1.

2. Realizar un diagnóstico de pacientes en estas categorías:
 - Control (*Healthy Control, HC*). Sujeto sano.
 - Deterioro cognitivo (*Mild Cognitive Impairment, MCI*). Sujeto con problemas en la memoria u otros problemas mentales funcionales, pero sin que estos problemas intervengan de manera severa a sus actividades diarias. Dentro de este grupo se diferencian dos subgrupos:
 - Declive cognitivo estable (*stable MCI, sMCI*). Sujeto MCI que en el periodo de observación no se ha convertido en paciente de Alzheimer.
 - Declive cognitivo convertido a Alzheimer (*progressive MCI, pMCI*). Sujeto que en el periodo de observación se ha convertido en paciente de Alzheimer.
 - Enfermo de Alzheimer (*Alzheimer Disease, AD*): sujeto diagnosticado como paciente de Alzheimer.

Cabe destacar que este proyecto se ha centrado en la **diferenciación entre sujetos sMCI y pMCI**, debido al alto riesgo que tienen los sujetos MCI de convertir a AD. En consecuencia, es el sector de la población donde la prevención es de fundamental y el diagnóstico tiene de vital importancia. Sin lugar a dudas, los sujetos MCI son el principal objetivo terapéutico para poder controlar la enfermedad, puesto que se debe actuar antes de que esta aparezca. Entre los sujetos normales o control, la probabilidad de transformar a AD es muy baja, entre 1-2 %/ año, dependiendo de la edad. Por ello, el estudio de este grupo no tiene tanta relevancia como el conjunto de sujetos MCI.

3. Análisis longitudinal de pacientes en la enfermedad del Alzheimer.
4. Comprobación empírica de los factores de riesgo de la enfermedad. Género, educación y gen APOE-4.
5. Simplificar, automatizar y ampliar la DEMO utilizada, facilitando el aprendizaje en el campo a las personas que deseen investigar en el mismo y proveerles de herramientas eficaces y eficientes que puedan ser utilizadas posteriormente. Por ello, se ha subido a Internet todo el trabajo realizado bajo una licencia de código abierto. Este es el enlace donde descargar todo el trabajo realizado <https://bitbucket.org/JaimeSV/tfg-alzheimer-early-diagnosis/src/master/>.

Es necesario resaltar que los resultados obtenidos en este proyecto tienen un ruido implícito en el etiquetado de los diversos grupos, causado por el diagnóstico clínico debido a que el diagnóstico definitivo de la enfermedad solo es posible post-mortem. Se ha reportado que el grupo AD tiene de media una sensibilidad del 81 % y una especificidad del 70 % [24].

1.3. Estructura del documento

A continuación y para facilitar la lectura del documento, se detalla el contenido de cada capítulo.

- En el capítulo 1 se realiza una introducción, mostrando la motivación del proyecto y sus objetivos.

- En el capítulo 2 se hace un repaso a conceptos patológicos de la enfermedad de Alzheimer y los marcadores más utilizados en esta enfermedad. Seguidamente, se realiza una revisión problemática centrándose en investigaciones que utilizan marcadores de imágenes de resonancia magnética estructural y evaluaciones clínicas cognitivas y funcionales para la diferenciación entre sMCI y pMCI.
- En el capítulo 3 se muestran los materiales utilizados con especial detalle en la base de datos ADNI (*Alzheimer's Disease Neuroimaging Initiative*).
- En el capítulo 4 se explica detalladamente el flow-chart seguido en el proyecto. A continuación, se desarrolla el marco teórico utilizado junto con los algoritmos propuestos en base a este marco.
- En el capítulo 5 se muestran los resultados obtenidos. Mostrando el control de calidad realizado, los clusters realizados tras el análisis másivo y las mejores combinaciones de marcadores.
- En el capítulo 6 se realiza la conclusión y se proponen nuevos proyectos futuros.
- Anexos.
 - Conceptos teóricos básicos. Se repasan conceptos necesarios para poder llevar a cabo este proyecto como la creación de imágenes de resonancia magnética, el registro de imágenes y el contraste de hipótesis.
 - Acrónimos. Para amenizar la lectura del documento se han utilizado una gran cantidad de acrónimos, por ello se facilita una lista donde puedan ser consultados.

Capítulo 2

Estado del arte

En el comienzo de este capítulo se introduce la enfermedad de Alzheimer, explicando brevemente la historia de esta enfermedad junto con sus efectos neuropatológicos. Seguidamente se resaltan los diversos biomarcadores utilizados en esta enfermedad.

A continuación, se hará una revisión problemática de los métodos utilizados para medir la atrofia en imágenes de resonancia magnética estructurales.

Para concluir el capítulo, se resalta la problemática existente en los estudios longitudinales masivos, finalizando con la revisión de la metodología utilizada por un conjunto de artículos muy similares a este proyecto. De esta literatura se han extraído conclusiones muy beneficiosas que han formado la base para esta investigación.

2.1. Enfermedad de Alzheimer

La enfermedad del Alzheimer (Alzheimer Disease, AD) fue descubierta en 1901 por el psiquiatra y neurólogo alemán Alois Alzheimer. Cuando el neurólogo realizó una autopsia a una mujer que había tenido demencia descubrió dos tipos de lesiones patológicas: los ovillos neurofibrilares y las placas seniles amiloides¹. Estos descubrimientos fueron mostrados a la comunidad científica en 1906.

2.1.1. Neuropatología

Como se ha destacado anteriormente, la enfermedad del Alzheimer tiene dos lesiones patológicas asociadas. Seguidamente, se explicaran cada una de ellas:

- **Ovillos neurofibrilares.** Se localizan en el interior de la neurona. Son estructuras anómalas que provocan la pérdida en la capacidad de transmitir mensajes nerviosos. Estas estructuras están básicamente formadas por la proteína τ . En condiciones normales, esta proteína forma unos microtubos esenciales en la arquitectura de la neurona.
- **Placas neuríticas.** Estructuras esféricas que se localizan en el espacio extracelular. Están formadas por restos de dentritas neuronales degeneradas alrededor

¹Estos cambios patológicos aparecen en grado leve debido al envejecimiento; pero son mucho más intensos en los pacientes con AD

de un núcleo (con una longitud de 40 a 42 aminoácidos), estas estructuras son llamadas placas β -amiloide. Estas placas están formadas por un gran número de proteínas, siendo la proteína β -amiloide la más numerosa; por ello esta proteína da el nombre al conjunto.

Algunos estudios muy recientes muestran que primero se desarrolla el material amiloide y posteriormente los ovillos neurofibrilares, pero estos estudios deben de ser contrastados.

2.1.2. Biomarcadores

Según Consejo Nacional de Investigación de los Estados Unidos (*National Research Council*) un biomarcador es un evento que puede usarse como medida para la evaluación de sistemas biológicos. Los principales biomarcadores de AD son:

- **Proteína fluido cerebroespinal (*Cerebrospinal Fluid Proteins, CSF*).** El CSF es un líquido transparente que baña el cerebro y la médula espinal. Estudios muestran que los cambios en los niveles de τ y β -amiloide en el CSF pueden ser utilizados para la detección de esta enfermedad [1].
- **Perfiles genéticos de riesgo.** Estudios han demostrado que existen genes que incrementan el riesgo de padecer esta enfermedad aunque no garanticen el desarrollo de la misma. En el estudio realizado por Richard J. O'Brien [35] destaca la importancia de los siguientes genes: *Amyloid precursor protein (APP)*, *Presenilin-1 (PS-1)*, *Presenilin-2 (PS-2)*, *Apolipoprotein E-e4 (APOE4)*. Siendo el último el más significativo .
- **Neuroimágenes.** Es el campo más prometedor y en el que se basa esta investigación.
 - Neuroimágenes estructurales. Estas técnicas extraen de información sobre la forma, la posición, el volumen o el espesor del tejido cerebral. Se ha demostrado que el volumen del cerebro de los enfermos de Alzheimer disminuye a medida que progresa la enfermedad. Las técnicas utilizadas son: topografía axial computerizada (*Computed Tomography, CT*) e imágenes de resonancia magnética (*Magnetic Resonance Imaging, MRI*), esta investigación se va a centrar en la última técnica.
 - Neuroimágenes funcionales. Revela como están trabajando las diferentes regiones del cerebro gracias a la monitorización de moléculas como la glucosa o el oxígeno. Estas técnicas incluyen tomografía por emisión de positrones (*Positron Emission Tomography, PET*) o imagen de resonancia magnética funcional (*functional MRI, fMRI*). Se ha demostrado la disminución de glucosa en áreas del cerebro relacionadas con la memoria, el aprendizaje o la solución de problemas.
 - Neuroimágenes moleculares. Utilizan ondas de radio muy específicas para detectar cambios celulares o químicos relacionados con la enfermedad.

Este proyecto de investigación esta focalizado en el uso de neuroimágenes estructurales, debido a que son mucho más económica que fMRI y PET; y menos invasivas que CSF. Igualmente se han utilizado evaluaciones clínicas cognitivas y funcionales junto con factores de riesgo clínicos.

2.2. Revisión problemática. Biomarcador MRI estructural

El Alzheimer es una enfermedad de neurodegenerativa, la cual está caracterizada por un daño precoz en la sinapsis, seguido de degeneración en los axones y finalizando con la atrofia en las dendritas [45]. Este proceso neurodegenerativo es más severo en ciertas partes del cerebro. Por esas razones las imágenes de resonancia magnética tienen un potencial muy elevado como biomarcador de la enfermedad, puesto que permite medir esta atrofia. Los siguientes tres métodos son los más utilizados por la literatura en la obtención de marcadores:

2.2.1. Método basado en mapas de densidad

Este método cuantifica la atrofia utilizando la densidad en los mapas de la materia blanca (*White Matter, WM*), materia gris (*Grey Matter, GM*) y el fluido craneoencefálico². Dependiendo el modo en el que se use estos mapas de densidad podemos diferenciar dos categorías:

Todo el mapa de densidad como característica (*density maps as features*).

Se centra en el estudio de la densidad de materia gris y materia blanca. Se ha demostrado que la relativamente menor densidad en materia gris encontrada en el hipocampo es un buen marcador de la atrofia. Por ejemplo, el estudio realizado por Plant [39] utilizó el determinante Jacobiano calculado mediante los mapas de densidad como entrada a un clasificador de tipo SVM (*Support Vector Machine*) obteniendo un 62 % de precisión en la predicción de AD.

Reducciones de mapas de densidad.

Comúnmente es necesario reducir la dimensión del vector de características debido a que este es relativamente grande en comparación con el número de sujetos disponibles. Para llevar a cabo esta reducción se utilizan los siguientes métodos:

Métodos de reducción supervisados y no supervisados. Destaca el estudio realizado por Liu [29] donde métodos lineales embebidos (*Linear embedding method*) fueron utilizados para transformar regiones multivariantes del cerebro en datos localmente lineales en el espacio.

Métodos basados en atlas. Este método parcela la imagen cerebral en regiones predefinidas anatómicamente y seguidamente extrae las características de dichas regiones.

Métodos basados en Regiones de Interés (*Region of Interest, ROIs*) adaptativos. Debido a la complejidad de la enfermedad la atrofia no sigue los patrones de las diversas regiones definidas anatómicamente. Para solventar este problema se utilizan ROIs adaptativas. Dos metodologías son utilizadas en función del número de atlas utilizados: Un único set, cuando los sujetos son registrados utilizando un único atlas y

²Veasé el apéndice A, apartado Glosario para obtener más información, sobre las diferentes estructuras del cerebro

múltiples set adaptativos.

Otro estudio realizado por Liu [28] muestra la mejora al utilizar múltiples atlas, a pesar de que las características obtenidas son diferentes representaciones del mismo sujeto. Al utilizar diferentes atlas las características no deben ser concatenadas, para solventar este problema el algoritmo se entrenó con las características de cada uno de los atlas y el resultado de la clasificación fue la predicción más votada por los diversos atlas. Esta estrategia provocó un aumento del 70.06 % al 78.88 % en la clasificación entre sMCI y pMCI.

2.2.2. Métodos basados en la posición de los vértices

Los enfermos AD sufren cambios no solo en las estructuras subcorticales, sino también en zonas de la superficie cortical como el lóbulo temporal y parietal [12]. Estos cambios pueden ser detectados analizando la compleja superficie cortical. Los datos de la corteza son extraídos mediante los vértices de la superficie del cerebro. Las características pueden ser directamente las medidas de los vértices o se puede realizar una reducción de las mismas.

Vértices como base de las características.

Los vértices son utilizados en diferentes estudios [27] para calcular diferencias morfológicas entre las que se incluyen las volumétricas como el espesor de la corteza, el área de la superficie y el volumen de la materia gris; o geométricas como la profundidad del surco o la media de la curvatura.

Reducción de vértices como base de las características.

Otra aproximación es comprimir la información dada por todos los vértices en un vector de características, los métodos utilizados para llevar a cabo esta labor son:

Métodos de reducción supervisados y no supervisados. La estrategia utilizada por Cho [9] es convertir los datos del espesor al dominio de la frecuencia consiguiendo reducir la dimensión de las características y pudiendo así filtrar la alta frecuencia (el ruido). Una alternativa es la utilizada por Park [37], el cual extrae el espesor parametrizado utilizando una malla tridimensional previamente creada. A continuación, se entrenó al clasificador (tipo SVM) con sujetos MCI y control para ser testado con pacientes que convertían a AD. Usando las imágenes de sujetos hasta una visita antes de la conversión se obtuvo un 83 %.

Métodos basados en atlas. Las imágenes del cerebro son registradas con ciertos estándares para que encajen a la perfección con las regiones existentes en las plantillas de los atlas. A diferencia de las plantillas nombradas anteriormente las cuales se basaban en el volumen, estas se basan en la superficie. Desikan [11] busca marcadores entre las diversas medias de los espesores corticales obteniendo que la corteza entorhinal y el supramarginal gyrus son regiones de especial interés. Otros estudios como el realizado por Oliveira [36] utiliza la media del espesor del cerebro completo y posteriormente lo clasifica con SVM. Estudios como el llevado a cabo por Wee [51] utiliza el espesor de las estructuras corticales, y la correlación existente entre

características, basando por tanto su estudio en la similitud existente en el espesor de diversas regiones del cerebro. Las características fueron seleccionadas usando t-test en primer lugar y seguido de mRMR. Posteriormente se clasificaron con SVM recursivo (*recursive-feature elimination SVM, SVM-RFE*).

Eskildsen [15] intento predecir los conversores midiendo el espesor en varias etapas del desarrollo de la enfermedad. Así agrupó a los pMCI basándose en el tiempo de conversión a AD (6, 12, 24 o 36 meses). Cada una de estas subcategorías fueron enfrentadas en un clasificador LDA (*Linear discriminant analysis*) contra sMCI. Esta clasificación basada en el estado específico en el que se encuentra el sujeto pMCI demuestra una mejora en la clasificación en contraposición a la clasificación general entre pMCI y sMCI. La Figura 2.1 muestra el atlas utilizado por freesurfer, el atlas de Desikan-Killiany.

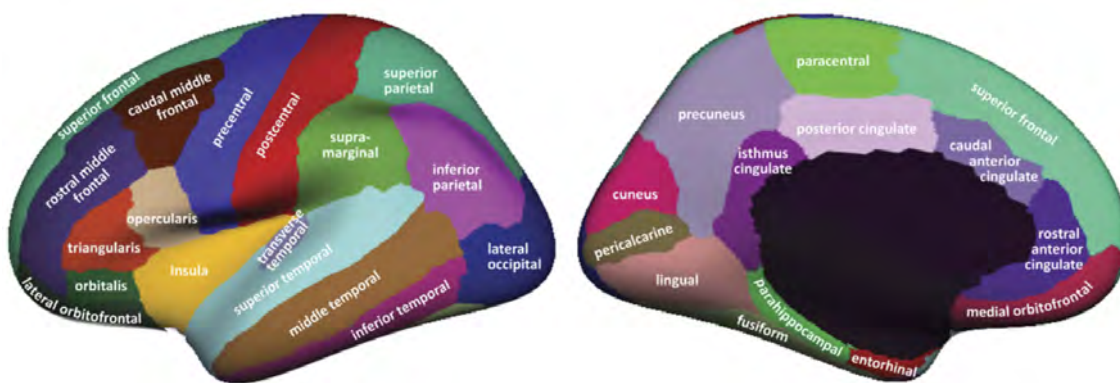


Figura 2.1: Atlas de Desikan-Killiany

Como juicio general, los métodos basados en las características procedentes de la reducción de vértices han muestra ser más precisas que los vértices por si mismos, mostrando un incremento entre el 10 %-13 % [37]. Además, métodos supervisados/no supervisados son métodos de reducción más efectivos que los métodos basados en atlas, en concreto se ha mejorado desde 65.22 % al 71.2 % en la diferenciación entre pMCI y sMCI.

2.2.3. Método basado en las regiones predefinidas

Este método se basa en los conocimientos previos de la enfermedad. Como se ha remarcado anteriormente existen numerosos estudios que prueban el deterioro del hipocampo, en consecuencia las propiedades geométricas del hipocampo han sido ampliamente exploradas para conseguir biomarcadores. Existen multitud de estrategias utilizadas, por ejemplo, en el estudio de Gardini [19] la forma del hipocampo es cuantificada mediante armónicos esféricos mientras que en el estudio realizado por Li [27] la superficie anatómica es modelada mediante una malla. Por otro lado, Sorensen [46] ha demostrado la importancia de la textura del hipocampo.

Características biológicamente seleccionadas.

La enfermedad de Alzheimer afecta a regiones del cerebro, atrofiando la corteza entorhinal, expandiendo los ventrículos y cambiando la volumétrica de núcleos del subcortex como la amígdala, el putamen, el caudato o el tálamo. Por ello, Chincharrini [8] usó la textura de la corteza entorhinal, del córtex perirhinal, del hipocampo y del parahipocampo. Estas características fueron analizadas utilizando el clasificador *random forest* para seleccionar las más adecuadas; posteriormente fueron clasificadas con SVM para predecir la conversión a AD. Otros estudios [47] usan las diferencias de forma existentes entre el lado izquierdo y derecho de regiones como la amígdala, hipocampo, tálamo, caudato, putamen, globo pálido y ventrículos laterales.

Para concluir, se debe resaltar las mejoras reportadas por la literatura al utilizar métodos de reducción tanto supervisados como no supervisados. Creando, por tanto, la necesidad de un algoritmo que seleccione las características más destacables junto con un clasificador relativamente complejo como puede ser LME o LDA.

En este proyecto los marcadores basados en imágenes de resonancia magnética estructurales utilizan tres métodos de los aquí expuestos:

1. Método basado en reducción de mapa de densidad mediante atlas. Obteniendo los volúmenes de las diferentes estructuras subcorticales dada por el atlas de Desikan-Killiany.
2. Método basado en la reducción de vértices mediante atlas. Obteniendo el espesor de las diferentes estructuras corticales dada por el atlas de Desikan-Killiany.
3. Método basado en la reducción de vértices mediante algoritmos supervisados y no supervisados. Primero extrayendo las zonas donde existe mayor discrepancia entre los diferentes grupos clínicos, para luego hallar el espesor medio de esta zona.

La Figura 2.2 muestra de manera muy visual como se ha obtenido tanto el espesor como el volumen en función del método utilizado.

2.3. Problemática en los estudios longitudinales masivos

En la última década ha surgido un tremendo crecimiento en los estudios longitudinales (LNI). La mayoría de ellos se han focalizado en un pequeño número de medidas provenientes de análisis univariantes y/o utilizan métodos que son subóptimos para la detección de efectos longitudinales [4]. Esta reducción del poder estadístico provoca un detrimento particularmente significativo cuando se explora la multitud de asociaciones en el cerebro realizando un análisis masivo univariante.

Desde hace unos pocos años, se han comenzado a realizar estudios con modelos longitudinales (e.g. modelo LME) a nivel de vóxel, realizando un análisis masivo sobre los datos longitudinales. Muchos de estos estudios tienen algunas de las siguientes carencias: los modelos seleccionados estudian cada vóxel de manera separada o los modelos presentes a nivel de vóxel no tienen en cuenta la estructura espacial de los datos. Como consecuencia de estas carencias, los estimadores son menos eficientes y

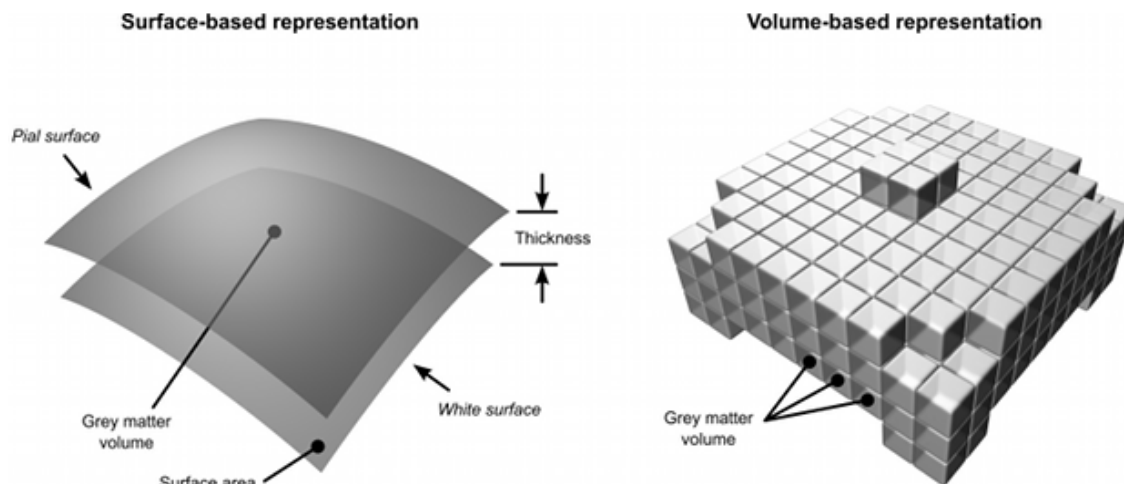


Figura 2.2: Comparación entre métodos basados en la superficie (espesor) y métodos basados en densidad (volumen). Fuente Winkler(2010)

el poder estadístico es reducido.

Para solventar estos problemas son utilizados los modelos espaciotemporales. Los modelos espaciotemporales agrupan la covarianza temporal de las estructuras a través de la localización espacial, esta estrategia ha sido satisfactoriamente utilizada en MRI funcional [10]. Un ejemplo de esta metodología es la estrategia SPM donde las agrupaciones son realizadas sobre vóxeles receptivos. Considerando vóxeles receptivos aquellos que sobrepasan F-test para cualquier efecto de interés con un valor límite de p-valor de 0.001. El problema de esta estrategia es que ignora la correlación inter-vóxel asumiendo que es cero. Igualmente, este modelo no es adecuado para datos longitudinales por dos razones. En primer lugar, la covariación temporal en medidas longitudinales suele ser muy diferente entre regiones distantes del cerebro, reflejando que diferentes regiones del cerebro son afectadas en las diferentes etapas de la enfermedad. En segundo lugar, la correlación inter-vóxel tiende a ser bastante alta entre puntos próximos, puesto que raramente cambios estructurales se dan en un solo punto sino que afecta a la región del córtex completa.

Por ello, en este proyecto se utiliza una novedosa metodología propuesta por Bernal [5]. Las características de esta estrategia es el análisis masivo univariante de datos longitudinales basados en el modelo espaciotemporal (*linear mixed effects, ST-LME*). Esta metodología obtiene los beneficios provenientes del análisis masivo univariante, puesto que el análisis es realizado sobre un enorme número de localizaciones espaciales (vóxel o vértices) y además agrupa las estructuras a través de la covarianza temporal existente en las localizaciones vecinas. Adicionalmente, esta

metodología continúa con las ventajas obtenidas al utilizar el modelo LME, puesto que permite trabajar tanto con datos longitudinales irregulares en el tiempo como con visitas perdidas.

2.4. Revisión metodología de estudios similares

Debido al gran impacto de esta enfermedad se ha realizado un gran esfuerzo en la creación de modelos predictivos de AD utilizando patrones de clasificación basados en la clínica, las imágenes, la genética y proteína fluido cerebroespinal. Las primeras investigaciones en este campo se realizaron en la década de los ochenta y los noventa, donde los estudios tendían a utilizar modelos estadísticos convencionales utilizando un predictor univariante, además estaban limitados por pequeñas cantidades de la muestra. Unos ejemplos de los marcadores dados en estos estudios son: la medida neuropsicológica en el baseline, la atrofia en el baseline del hipocampo y la atrofia del lóbulo temporal medial.

En contraste con estos métodos, la literatura reciente utiliza estrategias mucho más sofisticadas que ha facilitado en gran medida el desarrollo de este proyecto, debido a que se han extraído conclusiones de gran utilidad que han sentado las bases para este estudio.

El artículo realizado por Sorensen [46] y su equipo, muestra las curvas de aprendizaje en función del número de ejemplos, sugieren que un número mayor de 500 ejemplos para el entrenamiento no hubiera mejorado los resultados obtenidos. Por lo tanto, podemos concluir que la dimensión del data set tiene un valor apropiado. Adicionalmente, se ofreció una posible manera de normalizar el volumen de las diversas estructuras subcorticales por el tamaño de la cabeza. Esta normalización se realiza mediante la división de las estructuras entre el volumen intracraneal (intracraneal volumen, ICV). En contraste, otro estudio ha demostrado que el espesor de las estructuras corticales no deben de ser normalizados [52].

En el estudio realizado por Korolev [25], las fuentes de datos incluía los factores de riesgo, evaluaciones cognitivas y funcionales, MRI estructural, y proteína fluido cerebroespinal. Ciento ochenta y seis características fueron incluidas como potenciales marcadores entre los que destacan factores de riesgo como edad, educación, sexo, genotipo APOE, historial familiar de demencia, riesgo de enfermedad cardiovascular, índice de masa corporal, historial de desórdenes psiquiátricos, abuso de alcohol, golpes en la cabeza y apnea del sueño. Además incluyeron información sobre si el sujeto está bajo el régimen de medicación de AD (memantina e inhibidor de colinesterasa). Ninguno de estos factores de riesgo son utilizadas en el modelo con mejor actuación. Concluyendo que solo evaluaciones clínicas y MRI pueden ser considerados marcadores mientras que los factores de riesgo clínicos y la proteína son marcadores inadecuados debido a que aportan información limitada o redundante sobre la progresión de la enfermedad.

La novedad propuesta por Korolev es realizar un clasificador binario probabilístico. La siguiente frase de Sir Willian Osler resume excepcionalmente el espíritu encontrado en esta investigación *“Medicine is a science of uncertainty and an art of*

probability”, es decir, “*la medicina es una ciencia incierta y un arte de probabilidad*”. Por ello, en el estudio se implementó un kernel probabilístico multiple con aprendizaje (*probabilistic multiple kernel learning, pMKL*) que generaba la probabilidad de transformar a AD usando una inferencia bayesiana. Este pMKL es clasificador basado en kernel muy similar a SVM, la diferencia es que SVM es un clasificador no probabilístico mientras que pMKL produce la probabilidad en la predicción. Esta aproximación tiene múltiples ventajas como la posibilidad de separar al grupo MCI en subgrupos en relación con el riesgo a la progresión o la posibilidad de utilizar la probabilidad asociada a cada individuo como una medida de confianza.

La selección de características realizada se llevó a cabo mediante un filtro wrapper, que utilizaba el criterio de *Joint Mutual Information (JMI)*. Este método es muy similar al utilizado en este proyecto puesto que maximiza la relevancia mientras que minimiza la redundancia y maximiza la complementariedad con las características previamente seleccionadas. La selección de características se realizó mediante *cross validation (CV)*, buscando las características que mejor equilibraran el rango de precisión (*balanced accuracy rate, BAR*) definido como $[\text{sensibilidad} + \text{especificidad}] / 2$. Como resultado a su investigación obtuvo que las evaluaciones cognitivas y funcionales eran las características que mejor predicen la enfermedad mientras que el plasma proteómico no era de gran utilidad. La mejor actuación combinaba marcadores cognitivos y morfológicos de MRI prediciendo con un 80 % de precisión. Los marcadores cognitivos seleccionados fueron *Alzheimer’s Disease Assessment Scale (ADAS)*, *Rey Auditory Verbal Learning Test (RAVLT)* y *Functional Activities Questionnaire (FAQ)* mientras que los marcadores de MRI fueron el volumen del hipocampo izquierdo, el espesor del gyro temporal medial y el espesor del lóbulo parietal inferior. Es importante destacar que realizó una análisis trasversal (*Cross*), por lo que podría haber mejorado esta actuación si hubiera realizado un análisis longitudinal.

Dentro del tema de análisis masivo, el estudio realizado por Bernal [5] ha servido como inspiración para este proyecto. Además de permitir el acceso a multitud de datos y algoritmos, testeó la metodología LME con un modelo cuadrático en el tiempo incluyendo como variables independientes a los efectos fijos: el tiempo desde el punto de partida del estudio conocido como baseline, el cuadrado del tiempo, el grupo clínico miembro, la interacción entre el grupo clínico y el tiempo, y el tiempo al cuadrado, la edad en el baseline, el sexo, el APOE-4 y los años de educación. Luego realizó multitud de pruebas y en el 80 % de las mejores, el tiempo al cuadrado no estaba dentro del grupo óptimo de efectos aleatorios. Por lo que el modelo final consideró un modelo lineal con dos efectos aleatorios, la interceptación y el tiempo. Basándonos en su experimentación el modelo LME llevado a cabo en este proyecto también va a ser lineal.

Para finalizar este capítulo se debe subrayar el potencial de medir la atrofia, es decir, la muerte neuronal a partir de imágenes de resonancia magnética estructural. Junto con las diversas metodología utilizadas para obtener dicha atrofia. Además, se ha mostrado la problemática existente en los análisis masivos longitudinales, permitiendo que al conocer el problema, se pongan los medios para solventarlo. En adicción, se ha resumido los estudios más significativos y novedosos en el campo que han facilitado en gran medida este proyecto.

Capítulo 3

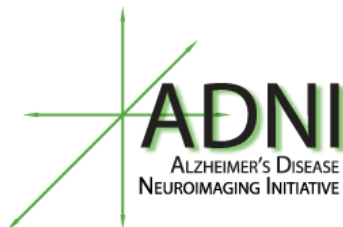
Materiales y herramientas utilizada

En este capítulo se muestra la importancia de la base de datos ADNI. Seguidamente, se explica en el protocolo de inclusión utilizado para la clasificación de los sujetos según el diagnóstico clínico, diferenciando entre tipos de MCI en función de los dominios afectados. Este punto tiene una tremenda importancia al poder ser el punto clave para comprender el desarrollo de la enfermedad.

A continuación, se repasan las evaluaciones neuropsicológicas utilizadas en este proyecto, con el objetivo de adquirir los conocimientos necesarios para obtener conclusiones reseñables cuando estos sean utilizados como marcadores.

Para finalizar, se muestran las herramientas imprescindibles para llevar a cabo el proyecto junto con algunas no estrictamente necesarias, pero si muy recomendables puesto que facilitan en gran medida el desarrollo del mismo. Estas herramientas son expuestas de modo de documento explicativo, siendo por este motivo de muy recomendable lectura para aquellas personas que deseen continuar con el desarrollo de esta investigación.

3.1. Base de datos



Las imágenes se obtienen de la base de datos creada por University of Southern California, ADNI (*Alzheimer's Disease Neuroimaging Initiative*). La página web de ADNI es: <http://adni.loni.usc.edu/>. En esta base de datos podemos encontrar imágenes de 1,5 Teslas o 3 Teslas de resolución, y dependiendo del sujeto pueden encontrarse imágenes de baseline (scan inicial), 6, 12, 24, y 36 meses.

ADNI fue creada en 2003 por *the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, the Food and Drug USA Administration*, compañías farmacéuticas privadas y organizaciones sin ánimo de lucro. Tuvo una inversión de 60 millones de dólares. Es necesario destacar la importancia de una **base de datos estandarizada** que permita comparar los diferentes algo-

ritmos utilizados por la comunidad.

El protocolo seguido por ADNI para la inclusión en los diversos grupos clínicos son las siguientes:

Tabla 3.1: Protocolo de inclusión. Fuente <http://adni.loni.usc.edu/>.

	Escala de Memoria de Wechsler	Mini-Mental State Exam (MMSE)	Clinical Dementia Rating (CDR)
HC	≥ 9 para ≥ 16 años de educación ≥ 5 para 8-16 años de educación ≥ 3 para 0-7 años de educación	24-30 (inclusive) ¹	0
MCI	≤ 8 para ≥ 16 años de educación ≥ 5 para 8-16 años de educación ≥ 3 para 0-7 años de educación	24-30 (inclusive)	0.5
AD	Igual que MCI	20-26 (inclusive)	0.5-1

Todos los sujetos incluidos en ADNI y por tanto en este estudio, tienen una edad comprendida entre 55 y 90 años, deben de tener una buena condición física sin ninguna enfermedad que pueda interferir en el estudio y su escala de Depresión Geriátrica (*Geriatric Depression Scale*) debe ser mayor que 6.

Los sujetos catalogados como MCI se pueden dividir en 4 subgrupos. En la Figura 3.1 se pueden observar los subgrupos. En la Figura 3.2 se puede ver el flow-chart seguido para la clasificación.

		Etiology				
		Degenerative	Vascular	Psychiatric	Trauma	
Clinical classification	Amnesic MCI	Single domain	AD		Depr	
		Multiple domain	AD	VaD	Depr	
	Non-amnesic MCI	Single domain	FTD			
		Multiple domain	DLB	VaD		

Figura 3.1: Subtipos de MCI. Fuente página web de ADNI

Todos los sujetos de incluidos en este estudio son MCI amnésicos afectados en un solo dominio. Se ha seleccionado este grupo dentro de los MCI para intentar tener un grupo homogéneo y de este modo poder dar respuestas mucho más precisas sobre la enfermedad. En adicción, este grupo de amnésicos son el grupo de mayor riesgo para trasformar a AD. Muy posiblemente los marcadores obtenidos pueden ser extrapolables a otros dominios pero esto debe ser contrastado.

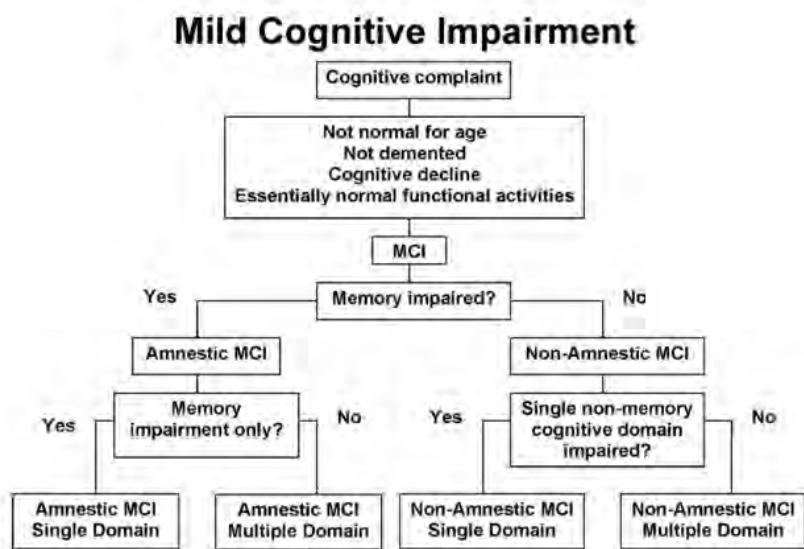


Figura 3.2: Flow-chart para la clasificación para sujetos MCI. Fuente <http://adni.loni.usc.edu/>.

3.1.1. Datos demográficos de los sujetos pertenecientes en el ensayo

Para este estudio se han utilizado dos conjuntos de grupos clínicos.

Grupo sMCI y pMCI Este grupo está formado por 534 scans de 143 sujetos. Se ha utilizado tanto para la selección de características como para la validación de las mismas. En la siguiente Tabla 3.2 se muestran las características en el baseline de los sujetos.

Tabla 3.2: Datos demográficos en el baseline de los sujetos sMCI y pMCI utilizados en el estudio

	Número de sujetos	Edad,años(media \pm std)	Genero Masculino (%)	MMSE ² (media \pm std)
sMCI	70	74.7 \pm 7.5	55.7	26.6 \pm 1.8
pMCI	73	74.5 \pm 6.8	71.2	27.7 \pm 1.5
Total	143	74.6 \pm 7.1	63.6	27.1 \pm 1.76

Cabe destacar que para la creación de los modelos en el análisis masivo se han utilizado los datos facilitados por Bernal de 50 sujetos sMCI y 50 sujetos pMCI, de los cuales no se tienen estos datos demográficos.

Grupo AD y HC Formado por 698 scans de 187 sujetos. Se ha utilizado tan solo para la selección de características. En la siguiente Tabla 3.3 se muestran las características en el baseline de los sujetos.

3.2. Test neuropsicológicos

A continuación, se explicarán brevemente los diferentes test neuropsicológicos utilizados en este proyecto con el objetivo de entender las diferencias existentes entre

Tabla 3.3: Datos demográficos en el baseline de los sujetos AD y HC utilizados en el estudio

	Número de sujetos	Edad,años(media \pm std)	Genero Masculino (%)	MMSE ³ (media \pm std)
AD	76	73.9 \pm 6.9	52.6	23.6 \pm 1.8
HC	111	75.5 \pm 5.2	53.1	29.1 \pm 0.9
Total	187	74.9 \pm 6.0	52.9	26.8 \pm 3.16

ellos y por lo tanto tener la capacidad de obtener conclusiones sobre la utilización de los mismos como marcadores.

3.2.1. Evaluación neuropsicológica mínima

Aquellos pacientes con Alzheimer tienen un deterioro en sus funciones cognitivas en particular en la memoria, atención, percepción, lenguaje, praxias (habilidades motoras adquiridas) y gnosias (capacidad de reconocer información previamente aprendida).

Los cuestionarios de evaluación neuropsicológica mínima han sido diseñados para cuantificar la severidad de la alteración en estas funciones cognitivas. Debido a la brevedad de los mismos y no de requerir personal especializado no son utilizados para establecer un diagnóstico.

Mini-mental state examination

La Mini prueba del estado mental (*Mini Examen del Estado Mental, MMSE*) es un método desarrollado por Marshal F. Folstein, Susan Folstein, and Paul R. McHugh en 1975 [49]. Es comúnmente utilizado para establecer el estado cognitivo del paciente y controlar su evolución. La principal ventaja de esta prueba es su corta duración (entre 5 a 15 minutos) por lo que suele utilizarse como evaluación inicial rápida. Al no detectar el dominio alterado ni conocer la causa del padecimiento nunca debe reemplazarse por una evaluación clínica completa.

El examen se basa en un cuestionario de 30 puntos agrupados en 10 secciones donde las características esenciales evaluadas son:

- Orientación espacio temporal.
- Capacidad de atención, concentración y memoria.
- Capacidad de abstracción (cálculo).
- Capacidad de lenguaje y percepción viso-espacial.
- Capacidad para seguir instrucciones básicas.

El formato utilizado es el siguiente:

1. Orientación en tiempo (5 puntos): Se pregunta el día de la semana, fecha, mes, año y estación del año. Se otorga un punto por cada respuesta correcta. Se puede considerar correcta la fecha con diferencia de 2 días.

2. Orientación en lugar (5 puntos): Se pregunta el lugar de la entrevista, hospital, ciudad, provincia y país. Se otorga un punto por cada respuesta correcta.
3. Registro de 3 palabras (3 puntos): Se pide al paciente que escuche atentamente tres palabras que se les mencionan (usualmente casa, zapato y papel) y tiene que repetirlas. Se le avisa que más tarde se le preguntarán de nuevo. Se otorga un punto por cada palabra correcta.
4. Atención y cálculo (5 puntos): Se pide que reste a 100 el número 7 y continúe restando de su anterior respuesta hasta que el explorador lo detenga. Por cada respuesta correcta se otorga un punto, y debe detenerse después de 5 repeticiones correctas.
5. Evocación (3 puntos): Se le pide que repita los objetos nombrados anteriormente. Por cada repetición correcta se otorga un punto.
6. Nominación (2 puntos): Se muestra un lapicero y un reloj, el paciente debe nombrarlos. Se otorga un punto por cada respuesta correcta.
7. Repetición (1 punto): Se pide al paciente que repita la siguiente oración: "Tres perros en un trigal", se otorga un punto si puede realizar la acción.
8. Comprensión (3 puntos): Se le indican tres órdenes simples que pueda realizar. Por ejemplo, "Tome el papel con su mano derecha, doblelo a la mitad y póngalo en el suelo". Se otorga un punto por cada acción correcta.
9. Lectura (1 punto) Se solicita al paciente que lea la orden "Cierre los ojos" (escrita previamente) y la obedezca. No debe decirlo en voz alta y sólo puede explicársele una vez.
10. Escritura (1 punto): Se pide al paciente que escriba una oración, que debe tener sujeto y predicado. Se otorga 1 punto si la oración tiene sentido.
11. Dibujo (1 punto): Debe copiar un dibujo simple de dos pentágonos cruzados. Se considera correcto si su respuesta tiene dos figuras de 5 lados y su cruce tiene 4 lados.

Teniendo en mente que el MMSE es análisis rápido los resultados informan de la siguiente manera:

- Entre 30 y 27: Sin Deterioro.
- Entre 26 y 25: Dudoso o Posible Deterioro.
- Entre 24 y 10: Demencia Leve a Moderada.
- Entre 9 y 6: Demencia Moderada a Severa.
- Menos de 6: Demencia Severa.

Valoración Clínica de Demencia (*Clinical Dementia Rating* ,CDR) de Hughes

Este método se basa en una entrevista estructurada mediante un protocolo desarrollado por Charles Hughes, Leonard Berg, John C. Morris [34]. La Tabla 3.3 sirve como referencia donde se le asigna a cada área la puntuación que le corresponda de acuerdo con la casilla que mejor encaje con el estado clínico del paciente. Se debe enfatizar que el ítem de la memoria es considerado como una categoría primaria mientras que las demás son categorizadas como secundarias. Por lo tanto se pueden dar las siguientes combinaciones:

1. Si al menos otras tres áreas son calificadas con la misma puntuación que la memoria está predomina.
2. Si más de tres categorías se gradúan por encima o por debajo a la memoria predomina estas puntuaciones.
3. Si dos categorías se clasifican en un lado y las otras tres por el otro lado la memoria, el CDR final será el de la memoria.

	Deterioro				
	Ninguno 0	Dudoso 0.5	Leve 1	Moderado 2	Grave 3
Memoria	Sin pérdida de memoria o leves olvidos inconstantes	Olvidos leves constantes; recolección parcial de eventos; olvidos "benignos"	Pérdida moderada de memoria; más marcada para eventos recientes; el defecto interfiere con las actividades diarias	Pérdida grave de memoria; sólo retiene materias con mucho aprendizaje; materias nuevas se pierden con rapidez	Pérdida grave de memoria; sólo retiene fragmentos
Orientación	Completamente orientado	Completamente orientado pero con leve dificultad para las relaciones temporales	Dificultad moderada con las relaciones temporales; orientado en el lugar del examen; puede tener algo de desorientación geográfica en otro lugar	Dificultad grave con las relaciones temporales; habitualmente desorientado en el tiempo; a menudo en el lugar	Orientado sólo en persona
Razonamiento y solución de problemas	Resuelve los problemas diarios y se encarga bien de los negocios y finanzas; razonamiento bueno con relación al comportamiento previo	Leve dificultad para resolver problemas, similitudes y diferencias	Dificultad moderada para hacer frente a problemas, similitudes y diferencias; razonamiento social habitual mantenido	Grandes dificultades para hacer frente a problemas, similitudes y diferencias; juicio social habitual limitado	Incapaz de razonar o resolver problemas
Actividades fuera de casa	Función independiente a su nivel habitual en el trabajo, compras, voluntariado y agrupaciones sociales	Leve dificultad en estas actividades	Incapaz de ser independiente en estas actividades, aunque aún puede participar en alguna; parece normal a primera vista	Incapaz de ser independiente fuera de casa Parece estar lo suficientemente bien como para realizar funciones fuera de casa	Parece demasiado enfermo/a como para realizar funciones fuera de su casa
Actividades domésticas y aficiones	Vida en casa, aficiones e intereses intelectuales bien conservados	Vida en casa, aficiones e intereses intelectuales algo limitados	Dificultad leve pero clara de su actividad doméstica; abandono de las tareas más difíciles; abandono de las aficiones e intereses más complicados	Sólo realiza tareas simples; intereses muy restringidos y mal mantenidos	Sin función significativa en casa
Cuidado personal	Completamente capaz de cuidarse por sí mismo/a		Necesita recordatorios	Requiere ayuda para vestirse, asearse y encargarse de sus efectos personales	Requiere mucha ayuda para su cuidado personal; incontinencia frecuente

Figura 3.3: Tabla utilizada para el test CDR. Fuente <https://www.hipocampo.org/hughes.asp>

3.2.2. Evaluación neuropsicológica completa

Estos test son realizados por especialistas por lo que permiten dar un diagnóstico. Se suelen combinar con el objetivo de poner en manifiesto aquellos perfiles propios de cada enfermedad.

Cuestionario de actividades funcionales (*Functional Activities Questionnaire, FAQ*)

Este test permite medir capacidad del paciente para realizar sus actividades de la vida diaria, como puede ser prepararse la comida o manejar su situación financiera personal [38]. Otros test necesitan una alta habilidad cognitiva mientras que el gracias a la sencillez de FAQ permite evaluar cambios en las funciones cognitivas a lo largo del tiempo. Véase la Figura 3.4 donde se muestran las preguntas realizadas en este test.

Functional Activities Questionnaire

Administration

Ask informant to rate patient's ability using the following scoring system:

- Dependent = 3
- Requires assistance = 2
- Has difficulty but does by self = 1
- Normal = 0
- Never did [the activity] but could do now = 0
- Never did and would have difficulty now = 1

1.	Writing checks, paying bills, balancing checkbook	
2.	Assembling tax records, business affairs, or papers	
3.	Shopping alone for clothes, household necessities, or groceries	
4.	Playing a game of skill, working on a hobby	
5.	Heating water, making a cup of coffee, turning off stove after use	
6.	Preparing a balanced meal	
7.	Keeping track of current events	
8.	Paying attention to, understanding, discussing TV, book, magazine	
9.	Remembering appointments, family occasions, holidays, medications	
10.	Traveling out of neighborhood, driving, arranging to take buses	
TOTAL SCORE:		

Figura 3.4: Tabla utilizada para el test FAQ. Fuente <https://www.alz.org>

Test de Aprendizaje Auditivo Verbal de Rey (*The Rey Auditory Verbal Learning Test, RAVLT*) [43]

Este test es utilizado para comprobar el deterioro de la memoria así como analizar la capacidad de aprendizaje. Se ha comprobado que la actuación en este puede estar influenciada por variables demográficas y culturales. Esta evaluación se basa en dos listas de palabras:

1. La primera lista consta de 15 palabras que no tienen relación entre sí y que el paciente debe de recordar. Cuando se le pide que las recuerde, se anota el orden con el objetivo de identificar el patrón de recuerdo. Este primer ensayo nos proporciona información de la memoria inmediata o de corto plazo. Posteriormente, se repite el experimento hasta 5 veces. Este último ensayo nos informa sobre el nivel de adquisición final. La puntuación otorgada es el número total de palabras recordadas en todos los ensayos.

Gracias a este test podemos obtener la *Curva de Aprendizaje*, es decir la diferencia de aciertos entre la primera y quinta vez. Las puntuaciones normales suelen ser de 6 acertadas en la primera vez y 12 en la última. Por último, tras 30 minutos se le realiza la prueba de *Recuerdo Demorado* puesto que se le pide recordar de nuevo todas las palabras, comúnmente suelen acordarse de 2 menos que el último ensayo.

2. La segunda lista es denominada de *Reconocimiento* y consta de 30 palabras, 15 de la primera lista y 15 nuevas. Se le pide al sujeto que indique si la palabra escuchada esta o no en la lista que ha debido memorizar anteriormente. Este ensayo es utilizado para determinar si las palabras no recordadas en el recuerdo demorado se debe a un problema de evocación o que no se han memorizado.

REY AUDITORY VERBAL LEARNING TEST (RAVLT) Phase I										
Patient: _____										
Age: _____										
LIST A	1	2	3	4	5	LIST B	1	LIST A	6	7
DRUM						DESK		DRUM		
CURTAIN						RANGER		CURTAIN		
BELL						BIRD		BELL		
COFFEE						SHOE		COFFEE		
SCHOOL						STOVE		SCHOOL		
PARENT						MOUNTAIN		PARENT		
MOON						GLASSES		MOON		
GARDEN						TOWEL		GARDEN		
HAT						CLOUD		HAT		
FARMER						BOAT		FARMER		
NOSE						LAMB		NOSE		
TURKEY						GUN		TURKEY		
COLOR						PENCIL		COLOR		
HOUSE						CHURCH		HOUSE		
RIVER						FISH		RIVER		
SCORE										

Figura 3.5: Tabla utilizada para el test RAVTL. Fuente J Pediatr (Rio J). 2004;80(2 Suppl):S111-6

Escala de Evaluación para Enfermedad de Alzheimer (*Alzheimer's Disease Assessment Scale, ADAS*) [44]

Es un test especialmente diseñado para la enfermedad de Alzheimer. Esta evaluación consta de dos subescalas: ADAS-Cog o cognitiva y ADAS-Noncog o conductual. Esta última no es muy popular debido a la existencia de escalas conductuales más amplias como Escala de Medición de Patología Conductual en la Enfermedad de

Alzheimer (BEHAVE-A). ADAS-Cog varia entre 0 puntos (sin deterioro) a 70 (demencia severa). La evaluación de las funciones cognitivas está repartida en:

- Memoria: 27 puntos
- Orientación: 8 puntos
- Lenguaje: 25 puntos
- Praxias: 10 puntos

Este test es utilizado tanto en el diagnóstico como para monitorizar la evolución de la enfermedad. Por ello, es ampliamente utilizado para la valoración de la eficacia de nuevos fármacos.

3.3. Software

3.3.1. Matlab



MATLAB, abreviatura de MATrix LABoratoryes, es una herramienta de software matemático que ofrece un entorno de desarrollo integrado con un lenguaje de alto nivel propio. La principal ventaja de este software es la facilidad de depuración y la multitud de toolbox presentes, de las cuales se han utilizado: Statistics and Machine Learning Toolbox (<https://es.mathworks.com/help/stats/index.html>) y LME (descargada en el paquete de Freesurfer).

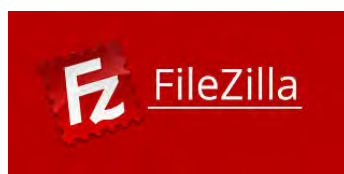
Teniendo en cuenta el nombre del programa se debe subrayar que matlab esta específicamente diseñado para operar con matrices, por lo que es extremadamente recomendable utilizar matrices en la implementación en lugar de bucles, disminuyendo así el tiempo de computación.

3.3.2. Freesurfer



Software desarrollado por Center for Biomedical Imaging at Massachusetts General Hospital. Su función es analizar imágenes de resonancia magnética. Entre la multitud de herramientas que incluye se debe destacar la reconstrucción topológica, los modelos geométricos precisos para la superficie de la sustancia gris y la sustancia blanca, las medidas del espesor cortical y el etiquetado de 35 regiones no corticales.

3.3.3. Filezilla



Software libre, multiplataforma y de código abierto. Gracias a su interfaz gráfica se puede transferir archivos entre el servidor y el cliente o viceversa, de manera visual e intuitiva.

3.3.4. Conexión al servidor desde el bash de Ubuntu

El comando necesario para acceder desde un terminal bash de ubuntu a un servidor es: `ssh -X -p 22 nombreUsuario@IpDelServidor` ⁴ Para desconectarse del ssh se utiliza el comando `exit`.

3.3.5. Tmux Matlab -nodesktop

Ejecutar freesurfer en Matlab suele durar varias horas. Al ejecutar este programa desde un servidor y desconectarnos se interrumpe la ejecución. Ante este problema existen varias soluciones:

- Dejar el cliente encendido mientras se ejecuta el programa en el servidor. Nada deseable.
- Iniciar sesión en otro servidor y que este sea el cliente del servidor donde se ejecuta el programa. Esta solución es eficaz pero nada eficiente, es la solución que se ha llevado a cabo en el pasado.

Por ello, se propone la utilización de la herramienta Tmux con el objetivo de solventar este problema.

Tmux es un multiplexador de terminales, es decir, permite lanzar múltiples terminales dentro de única pantalla (sesión de consola, sesión telnet/ssh, pestaña de terminal en un GUI). Cada terminal es totalmente gestionable y separada de las demás. Por ello, si se crea una sesión y se ejecuta matlab desde un servidor en otro, tan solo es necesario reconectarte a la sesión previamente creada, evitando en consecuencia el uso del servidor intermedio.

Comandos frecuentes

Es necesario destacar que para comunicarnos con tmux debemos de realizarlo de la siguiente manera: `CTRL - b < comando >`. Los comandos más frecuentes son:

- `&-` cierra la ventana actual
- `d-` cierra tmux
- `tmux new -s NombreDeLaSesion` -crea una sesión
- `tmux list-sessions-` lista las sesiones creadas
- `tmux attach -t NombreDeLaSesion-` permite entrar en una sesión creada

Matlab -nodesktop

Una vez iniciada nuestra sesión en Tmux, es necesario ejecutar matlab en la terminal, debido a que si se realiza de forma gráfica, al cerrar la pantalla de matlab se cierra el programa. Para ello ejecutamos:

```
/usr/local/MATLAB/R2017b/bin/matlab -nodesktop 5
```

Dentro de Matlab -nodesktop se ejecuta:

⁴se recomienda incluir este comando en un archivo `.sh`

⁵se recomienda incluir este comando en un archivo `.sh`

cd /home/jsimarro/Desktop/DEMO/scripts -con el objetivo de cambia a la carpeta donde se encuentran los scripts

pwd -matlab nos devuelve la carpeta en la que se encuentra actualmente

3.3.6. Archivos .sh

Primeramente un archivo .sh es un archivo de Developer creados en lenguaje Bash, es decir, no dejan de ser comandos de texto dentro de la interfaz de la shell. La utilidad de estos archivos es la simplicidad que ofrecen al no ser necesarios recordar ningún parámetro.

Creación

Se debe escribir en un editor de texto:

```
#!/bin/bash
ssh ComandosDeseados
exit
```

Ejemplo

Con el fin de no tener que recordar la IP del servidor Copernico, se ha creado este archivo llamado copernico.sh:

```
#!/bin/bash
ssh -X -p 22 jsimarro@138.100.100.136
exit
```

Gracias a este .sh la conexión al servidor se realiza con la simple ejecución de *sh copernico.sh*

En este capítulo se han expuesto los materiales junto con los protocolos de inclusión de ADNI y de este proyecto (solo sujetos MCI amnésicos de un solo dominio). Con el objetivo de entender el grupo clínico estudiado, se ha procurado eliminar cualquier otro sesgo que pudiera introducir ruido en los datos (e.g. sujetos con depresión).

Además se ha repasado las herramientas utilizadas en este proyecto, todas ellas son gratuitas excepto Matlab, cuya licencia se obtiene de manera gratuita para la comunidad universitaria. Esto tiene una alta relevancia puesto que al ser herramientas accesibles facilitan la replicación de la metodología llevada a cabo.

Capítulo 4

Métodos empleados

En este capítulo se explica de manera detallada las etapas de desarrollo llevadas a cabo en esta investigación; junto con los fundamentos matemáticos en los que se basan los algoritmos utilizados.

4.1. Etapas del desarrollo

El diagrama 4.1 muestra el flujo de etapas llevadas a cabo en esta investigación.

4.2. Recogida de datos

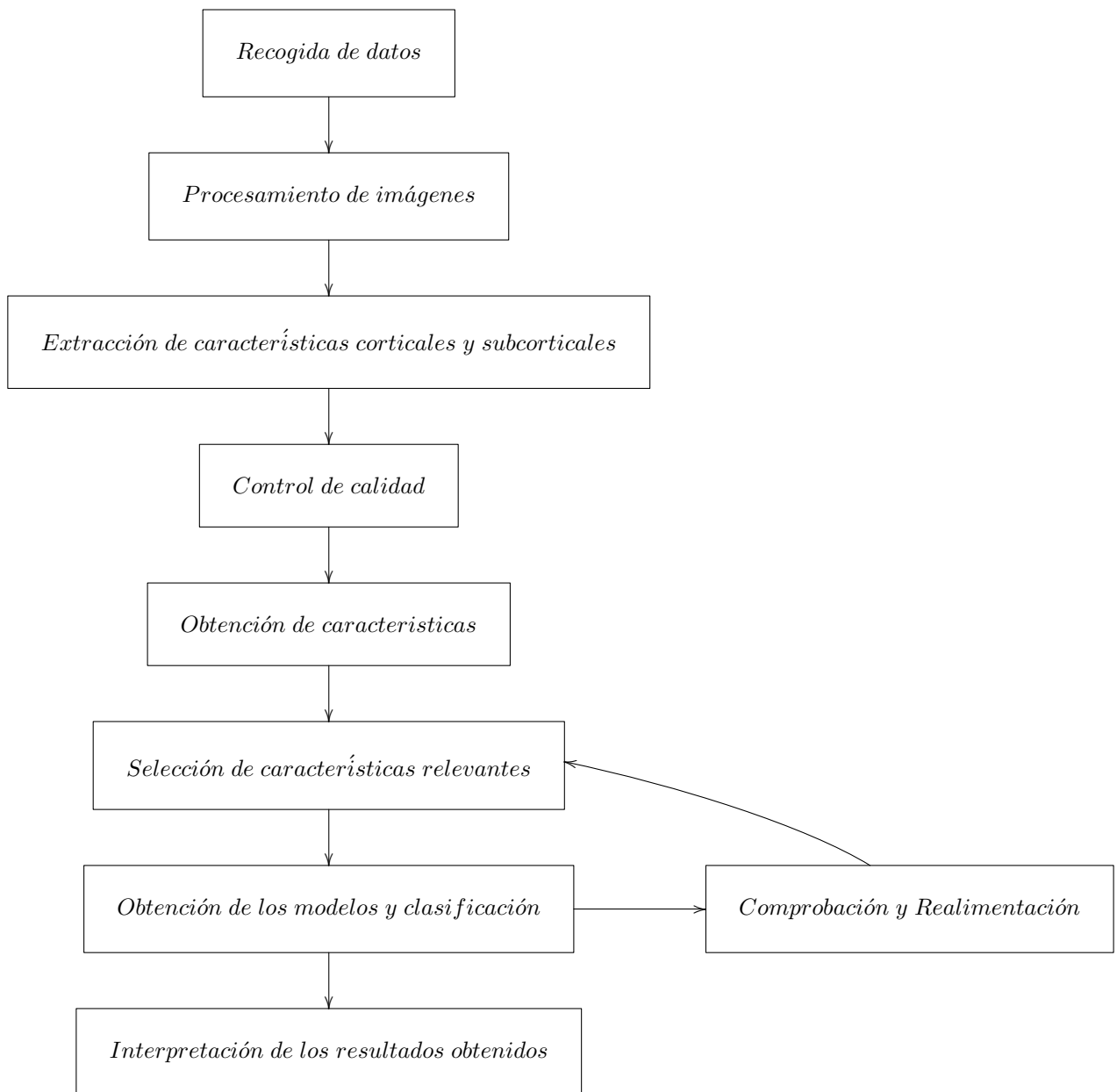
Desde la pagina inicial de ADNI <http://adni.loni.usc.edu/>, es necesario dirigirse a Data & Samples, a continuación Access Data. Antes de acceder a la base de datos es necesario la creación de una cuenta. Se debe especificar la universidad y el grupo de investigación junto con algunos datos personales. Una vez aceptada la cuenta ya se tiene acceso; por lo que se puede acceder a la carpeta *Download* → *ImageCollections*. Desde esta página se puede filtrar en Advanced Search, si es necesario encontrar a un sujeto con unas características determinadas para el estudio. De lo contrario, el siguiente paso es Data Collections, donde se selecciona *OtherSharedCollections* → *ADNI* más la carpeta deseada. Es necesario destacar que al elegir las imágenes a descargas del mismo sujeto, se debe elegir las mismas opciones de preprocesado que son mostradas en *Description*.¹

4.3. Procesamiento de imágenes

Al descargar una imagen desde ADNI, se obtiene junto a la misma un gran número de carpetas con los datos de identificación de cada sujeto, día, hora de la cita etc. Se desea obtener la imagen nii sin la necesidad de navegar manualmente por estos múltiples directorios. Por ello, la función `getNIIFile` automatiza este proceso. Además, `getNIIFile` acorta el nombre del sujeto dejando tan solo lo necesario (número de identificación del paciente y de la visita) transformando la imagen de nii. a .nii.gz.

¹ Es importante diferenciar entre Scaled/Scaled 2

Tabla 4.1: Flow char del proyecto



Para extraer las estimaciones del volumen y espesor de las subestructuras, las imágenes son automáticamente procesadas utilizando el análisis longitudinal [42] de FreeSurfer. En contraste con el estudio cross, el estudio longitudinal permite incrementar el poder estadístico debido a la reducción de los efectos producidos por la variación entre sujetos [48]. Aun así, el análisis cross es un paso previo para el estudio longitudinal.

Se deben de realizar las siguientes sub-etapas para realizar un procesamiento longitudinal de las imágenes.

4.3.1. Análisis Cross

Se debe ejecutar *FS_recon_all* para procesar imágenes T1 de tipo Cross. Este proceso es muy costoso computacionalmente, por lo que su duración temporal es bastante grande. En la Figura 4.3.1 se muestra de manera muy abreviada y visual las diversas transformaciones que se realizan junto con los diferentes atajos que se pueden realizar.

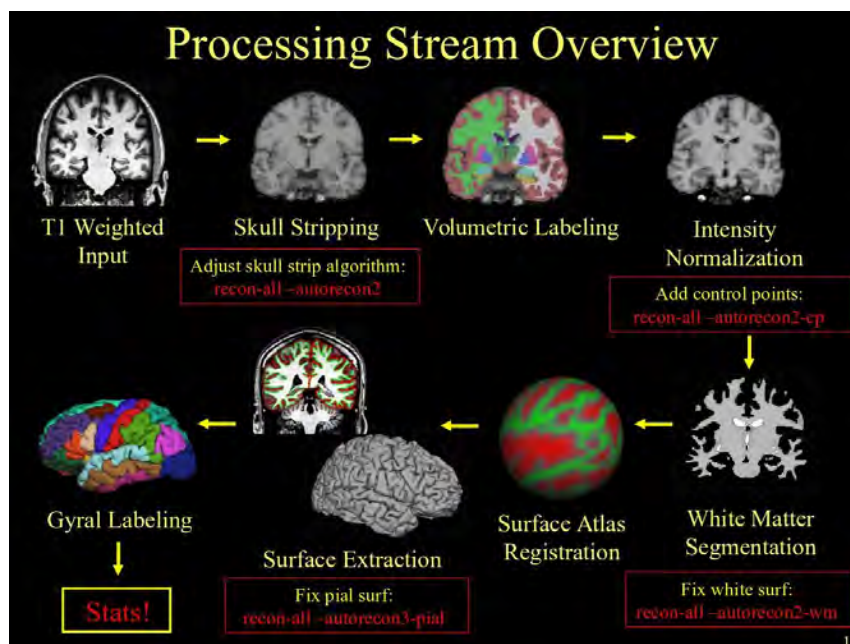


Figura 4.1: Resumen visual del procesado mediante recon all. Fuente <https://surfer.nmr.mgh.harvard.edu>.

A continuación, se explica brevemente los diversos pasos realizados en este proceso.

1. **Motion Correction.** Este paso realiza pequeñas correcciones de movimiento (Transformaciones geométricas y espaciales) que pueden haber entre diversas fuentes realizando una media entre ellas.
2. **Intensity Correction.** Realiza una corrección de intensidad no uniforme sobre las MRI, asumiendo algunas características sobre los datos.
3. **Talairach.** Antes de explicar que se realiza en en este paso es necesario comprender que es Talairach. El espacio de Talairach, es un sistema de coordenadas del cerebro humano utilizado para localizar las diferentes estructuras indiferentemente del tamaño y la forma del cerebro. Por ello, en este paso, se realiza una transformación afín utilizando la mascara 305 MRI.
4. **Normalization.** Se realiza una normalización de intensidad con el objetivo de facilitar la segmentación de base mediante intensidad. Se debe destacar que la intensidad esta escalada siendo el máximo 110, correspondiente con el color blanco.
5. **Skull Strip.** Se elimina el cráneo de la imagen.

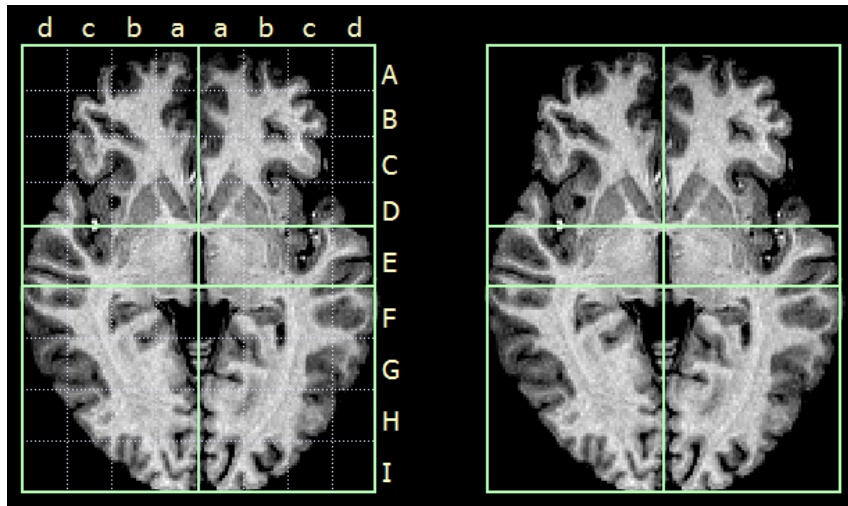


Figura 4.2: Cerebro humano representado donde se puede observar las coordenadas de Talairach

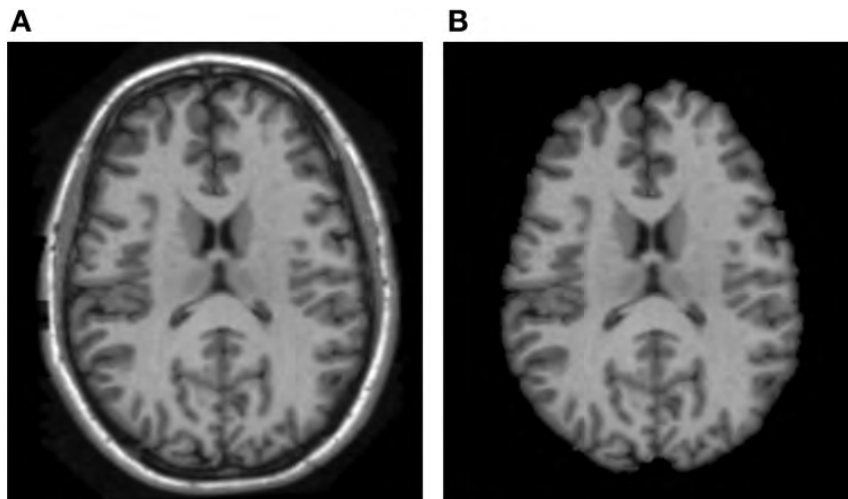


Figura 4.3: Resultado del proceso de skull strip

6. **EM (GCA) Registration.** Calcula la transformación para alinear la imagen con el atlas GCA² de freesurfer
7. **CA Normaliza.** Normalización basada en el modelo GCA
8. **CA Register.** Transformación no lineal para alinear con el atlas GCA
9. **Remove neck.** Se elimina el cuello de la imagen.
10. **EM Registration, with Skull.** Transformación donde se alinea la imagen sin cuello con la imagen GCA que posee cráneo.
11. **CA Label.** Se etiqueta las diferentes estructuras subcorticales, basándose en el modelo GCA.
12. **ASeg Stats.** Cuantificación estadística de las diversas estructuras subcorticales.

²Gaussian Classifier Atlas

13. **Normalization2.** Se realiza una segunda normalización de intensidad. Esta normalización es mejor que la anterior debido a que no existe el cráneo en este paso del proceso.
14. **WM Segmentation.** El objetivo es separar la materia blanca de todo lo demás. Se usa la intensidad del píxel, de los píxeles vecinos y constantes de uniformidad.
15. **Cut/Fill.** Se realiza un corte en el medio de la mente por el cerebelo. El hemisferio izquierdo es binarizado hasta 255, mientras que el derecho hasta 127
16. **Tessellation.** Se crea la superficie cubriendo los hemisferios con triángulos
17. **Orig Surface Smoothing.** Tras el paso anterior la superficie esta muy dentada (se ha rellenado con triángulos) por lo que se suaviza la superficie.
18. **Inflation.** Se minimiza la distorsión preservando el área y las distancias.
19. **QSphere** Primer paso para el reconocimiento automático. Se realiza una transformación esférica cuasi-homeomorfismo (quasi-homeomorphic spherical transformation) con el objetivo de localizar los defectos topológicos para corregirlos posteriormente.
20. **Automatic Topology Fixer.** Elimina los defectos, el usuario debe comprobar la superficie de la imagen.
21. **Final Surfaces.** Crea el archivo h.white y h.pial. El archivo white es realizado con un gradiente de intensidad de tipo blanco-gris mientras que el pial con un gradiente de intensidad de tipo gris-CSF.
22. **Cortical Ribbon Mask.** Crea una máscara con 1 o 0 .
23. **Spherical Inflation.** Transforma la superficie original en una esfera minimizando la distorsión métrica. Este paso es necesario para registrar la superficie en un atlas esférico.
24. **Ipsilateral Surface Registration.** Se registra la superficie en el atlas esférico.
25. **Contralateral Surface Registration.** Se registra la superficie en el atlas contralateral.
26. **Average Curvature.** Promedia la curvatura del sujeto.
27. **Cortical Parcellation.** Asigna etiquetas neuroanatómicas a las estructuras.
28. **Parcellation Statistics.** Crea una tabla resumen con los datos estadísticas de cada estructura. En estos datos se incluyen la superficie total, el volumen de la materia gris, la media del grosor cortical entre otros datos relevantes.

4.3.2. Análisis Base

Una plantilla para cada sujeto es creada utilizando registro robusto e inverso [41]. Se utilizan todos los scans del sujeto y por tanto todos los instantes de tiempo. Así se **estima la anatomía media/mediana del sujeto**. El comando necesario para ejecutar este paso en free surfer es *FS_recon_all_base*.

4.4. Extracción de las características corticales y subcorticales

El procesamiento de imágenes realizado anteriormente crea una carpeta llamada *stats* que se encuentra dentro de *Long*. Esta carpeta contiene una gran cantidad de datos correspondientes al volumen y el espesor de las subestructuras cerebrales, tanto corticales como subcorticales. Para la extracción y la clasificación de estas características se utiliza el script *getAsegAparcADNI*.

En este paso se añaden los test neuropsicológicos provenientes de ADNI juntando los mismos con el valor del espesor y el volumen de las diferentes estructuras.

4.5. Control de calidad

El control de calidad se encarga de comprobar si alguna de las imágenes no ha sido procesada adecuadamente. Para ello, se analiza la correlación existente entre pares de scans del mismo sujeto, obviamente esta correlación debe ser muy alta debido a que los cambios anatómicos presentes entre scans del mismo sujetos son muy bajos. Además, se crean las trayectorias del volumen del hipocampo y el espesor entorhinal, comparando las mismas con las trayectorias modelo del grupo analizado. Si el análisis se ha realizado correctamente la atrofia anual de un paciente debe ser similar a la esperada en ese grupo.

El control de calidad se realiza para asegurarse que todos los datos de entrada del clasificador son apropiados, es decir, son del grupo clínico buscado, sus valores están dentro de lo esperado para su grupo y no ha habido ningún error en la realización de la extracción de características. Para llevar a cabo esta labor se han utilizado las siguientes estrategias:

4.5.1. Correlación entre pares de scans del mismo sujeto

Como los cambios anatómicos que produce la enfermedad son relativamente pequeños, las diferencias entre los scans de un mismo sujeto no pueden ser muy grandes. Por ello, se realiza un plot donde se muestra la correlación entre dos variables, las cuales deben estar altamente correladas. Si algún sujeto es un outliers y por tanto no tiene esta alta correlación implica que se ha ocurrido un error en la extracción de características. Como por ejemplo en la alineación o la extracción del cráneo (*skull stripping*).

Cabe destacar que esta correlación se realiza sobre el hipocampo y el córtex, pero podría realizarse en cualquier otra región.

4.5.2. Coeficiente de correlación de Pearson

El coeficiente de correlación de Pearson mide la relación de linealidad existente en dos variables aleatorias. La principal diferencia con respecto a la covarianza es que el índice de Pearson es independiente de la escala las medias.

Para realizar este paso del control de calidad se busca la relación entre el volumen del hipocampo dado por freesurfer y el obtenido de una manera completamente independiente mediante segmentación. Por ello, la relación entre las dos variables deben de ser altamente lineales y tener un coeficiente de Pearson muy alto; puesto son representaciones de lo mismo obtenidas de dos modos diferentes.

4.5.3. Correlación de volumetría hipocampal y del valor de ICV

En la tabla data por ADNI llamada ADNIMERGE, se encuentran el valor del volumen del hipocampo y del ICV calculado mediante FSv4.4-cross. Es cierto que nuestro procesamiento ha sido realizado mediante la metodología longitudinal pero aun así debe existir una alta correlación entre estos valores. Si existe algún punto que salga de esta tendencia lineal deberá ser inspeccionado.

4.5.4. Comparación con los datos del grupo de investigación de Harvard

Bernal [4] ha publicado los datos del volumen del hipocampo normalizado y el espesor de la corteza entorhinal de 166 sujetos pMCI, siendo un total de 802 scans. Por ello, con el objetivo de verificar que nuestros datos son apropiados se realiza una comparación entre sus datos y los extraídos para este proyecto. La comparación se basa las trayectorias de los dos biomarcadores publicados por Bernal.

4.6. Obtención de los marcadores basados en MRI

La obtención de características ha seguido dos caminos. El primero de ellos basado en las regiones predefinidas por freesurfer y el segundo mucho más específico, extrayendo las características mediante análisis masivo de los vértices.

4.6.1. Métodos de reducción basados en atlas

En la Figura 2.1 puede verse el atlas de Desikan-Killiany, el cual es utilizado por Freesurfer. Un total de 35 ROIs son segmentadas y parceladas automáticamente por Freesurfer. En estas estructuras se incluye tanto el valor del **espesor de las estructuras corticales** como el **volumen de las estructuras subcorticales**.

4.6.2. Métodos de reducción basados en el análisis masivo

El primer paso llevado en el análisis masivo es la creación del modelo. Para ello, se han utilizado dos datos diferentes:

- El análisis longitudinal de 50 sMCI y 50 pMCI facilitado por Bernal [4].
- El análisis longitudinal de NC y AD facilitado por el trabajo de años anteriores.

Una vez cargados estos datos, se registra cada scan en la plantilla fsaverage. La plantilla utilizada tiene una elevada resolución al contar con 163.842 vértices por cada hemisferio. Posteriormente, se dividen el dataset y se crea el modelo LME masivo (el cual esta detalladamente explicado en el apartado: Modelo LME espaciotemporal ST-LME, de este mismo capítulo). El siguiente paso consiste en modelizar la trayectoria para cada vértice del cerebro de cada scan. Obteniendo las trayectorias de

cada grupo en cada vértice. El ratio de atrofia anual de cada grupo en cada vértice forma una distribución normal.

A continuación, se realiza un contraste de hipótesis entre las distribuciones de los grupos clínicos, siendo la hipótesis nula “las distribuciones del ratio de atrofia anual son iguales para ambos grupos clínicos”. Seguidamente se contrasta la hipótesis con un valor de alpha determinado obteniendo por tanto el p-valor en cada vértice y creado un modelo. Aquellos vértices cuyos p-valor sea muy bajo, tiene una muy baja probabilidad de que se cumpla la hipótesis nula y en consecuencia es un vértice potencialmente candidato para pertenecer a un cluster. Como la división se realiza n veces, se obtienen n modelos diferentes.

Algoritmo 1 Creación de los modelos de análisis masivo.

Entrada: Valores de espesor en cada vértice

Salida: Modelo masivo

- 1: Transformación al espacio del atlas fsaverage
 - 2: **para todo** Número de modelos deseados **hacer**
 - 3: División del dataset
 - 4: Creación del modelo ST-LME para cada vértice
 - 5: Obtención del ratio de atrofia anual para cada vértice
 - 6: Contraste de hipótesis y obtención del p-valor
 - 7: Creación de un modelo masivo
 - 8: **fin para**
 - 9: FDR y obtención de los mapas de q-valor
 - 10: Creación de clusters definiendo condiciones como el área mínima.
-

La veracidad de la hipótesis se basa en el cálculo de probabilidades. Al haberse calculado varios modelos y por tanto tratarse de repetir n veces hipótesis individuales puede darse el caso de que aunque no haya diferencias significativas entre ambos grupos tengan valores de p-valor menores que alpha. Aumentando de esta manera la probabilidad global de error de tipo I (falsos positivos). Esto acarrearía interpretaciones erróneas de las conclusiones obtenidas. Parafraseando a David Colquhoun miembro de la Royal Society “*It is well know that high false discovery rates occur when many outcomes of a single intervention are tested.*”, es decir “*Es ampliamente conocido que altos rangos de falsos positivos ocurren cuando muchos resultados de un mismo suceso son testeados.*”

En consecuencia, es necesario ajustar el valor del p-valor mediante contraste múltiple. Una posible solución es la corrección de Bonferroni que utiliza el FWER (*Family-Wise Error Rate*), este método es demasiado estricto por lo que a veces provoca pérdida de información. Una alternativa a este método es el propuesto por Benjamini and Hochberges [3] donde se optimiza mediante FDR (*False Discovery Rate*). A diferencia del FWER, este método tolera una baja proporción de falsos positivos, incrementando por tanto el poder estadístico mientras que se conservan muy bajos los errores de tipo I.

False Discovery Rate (FDR)

La definición de FDR tiene dos interpretaciones de la misma realidad:

- Número de falsos positivos entre la suma de falsos positivos más positivos verdaderos.
- Rechazo de la H_0 cuando ésta es correcta (Error de tipo I). Es decir, se obtiene una diferencia significativa entre ambos grupos aunque en la realidad esta no existe. Este valor depende de alpha puesto que es la posibilidad de tomar la decisión errónea.

Al ajuste de los p-valores usando el método de optimización de FDR se le llama q-valor. Estas son las diferentes interpretaciones del p-valor y el q-valor.

- **p-valor** de 5 % implica el 5 % de todos los tests van a ser falsos positivos.
- **q-valor** de 5 % implica que 5 % del resultado significativo va a ser falsos positivos.

4.7. Algoritmo de selección de características. Algoritmo mRMR, *minimum-Redundancy-Maximum-Relevance*.

La selección de características es una estrategia donde se reduce la dimensión de estas; con el objetivo de identificar subconjuntos pequeños pero informativos de las características originales. Esto puede eliminar el overfitting, mejorar la ejecución del modelo y además produce modelos más fáciles de interpretar y potencialmente más eficientes en el tiempo.

En contraste con otros métodos de selección de características utilizados en la literatura como secuencial *Forward Feature Selection (SFS)* [46] o *Joint Mutual Information (JMI)* [25], se ha decidido utilizar **mRMR** [13].

Un concepto fundamental en este algoritmo es la información mutua I . La información mutua existente en dos variables se basa en la probabilidad de distribución $p(x,y)$ y sus respectivas probabilidades marginales $p(x)$ y $p(y)$. Viene dada por la siguiente ecuación:

$$I(x, y) = \sum_{i,j} p(x_i, y_j) \log \frac{p(x_i, y_j)}{p(x)_i p(y)_j} \quad (4.1)$$

Cuando la información mutua es cero, representa que las variables tienen datos aleatorios. Por otro lado, cuando la variable está diferenciada en las diferentes clases, la información mutua es alta. Como el propio nombre del algoritmo indica, para encontrar la solución óptima tenemos que tener en cuenta:

Mínima redundancia. Se mide el nivel de similitud entre diversos genes (en nuestro caso, estructuras corticales y subcorticales), la idea es encontrar la característica con menor similitud y por ello la más representativa.

$$\min W_1, W_1 = \frac{1}{s^2} \sum_{i,j \in S} I(i, j) \quad (4.2)$$

donde $I(i,j)$ representa la similitud entre los genes $I(g1,g2)$ mientras que S el número de características.

Máxima relevancia. Mide el poder discriminante de los genes objetivo en relación con las variables de clasificación h (en nuestro caso HC, AD o MCI ; sMCI o pMCI).

$$\max V_1, V_1 = \frac{1}{s} \sum_{i \in S} I(h, i) \quad (4.3)$$

Donde $I(h,i)$ hace referencia a $I(h, g_i)$.

Para optimizar estas dos condiciones representadas en Eq: 4.2 y Eq: 4.3 se pueden utilizar dos criterios:

- MID (Mutual Information Differences). Se busca el máximo de $(V1 - W1)$
- MIQ (Mutual Information Quotients). Se busca el máximo de $(V1/W1)$

Existen dos scripts donde se encuentran estas dos estrategias. Ambas tienen estos parámetros de entrada:

D. Matriz de $m \times n$ de las características de nuestros candidatos. Las filas deben de ser los ejemplos, mientras que las columnas los atributos. Deben ser de tipo integer o double.

F. Vector de $m \times 1$ donde se muestra el grupo de clasificación.

K. Número máximo de características que deseamos seleccionar.

Para encontrar la solución exacta es necesario operar con N^s , por ello, la toolbox utiliza un algoritmo heurístico para la optimización de este problema.

Es necesario normalizar mediante **Z-score**³ para facilitar que el algoritmo converja.

La selección de características se ha implementado mediante la metodología de validación cruzada mostrada en el Algoritmo 2

Algoritmo 2 Selección de las características más relevantes.

Entrada: Dimensión de las características deseadas en cada combinación y número de las n mejores combinaciones

Salida: Mejores características ordenadas función del número de repeticiones

- 1: **mientras** No se hayan cumplido todos las pruebas **hacer**
 - 2: Separación de los datos en training y test
 - 3: Selección de las mejores combinaciones con los datos del training, utilizando $mRMR_q$ (Se puede modificar a $mRMR_d$)
 - 4: **fin mientras**
 - 5: Obtención de las combinaciones que aparecen con mayor frecuencia en las primeras posiciones.
-

Nota importante: Si se descarga de la página web es necesario modificar la toolbox, cambiando:

1. En `miinclude.h` :cambiar "`#include < math.h >`" por "`#include < cmath >`"
2. En "`estjointentropy.cpp`", "`estcondentropy.cpp`" (filas 56 — filas 65) cambiar "`muInf/ = log(2);`" por "`muInf/ = log(2,0)`"

³restar la media y dividir entre la desviación típica

Una vez realizado estos pasos es necesario ejecutar `makeosmex.m` en Matlab para volver a compilar los archivos con tu versión de Matlab.

4.8. Obtención de los modelos y clasificación mediante residuos longitudinales

La clasificación se realiza mediante la metodología propuesta por C.Platero en [7]. Primero se calculan la trayectoria de cada grupo mediante LME. En segundo lugar se “predice” mediante la esperanza marginal el valor del marcador. A continuación, se calcula el error producido por cada característica y se realiza la media de los errores de cada sujeto a lo largo de los diferentes scans, de ahí el termino residuos longitudinales. Con estos errores medios de cada característica, se crea un clasificador.

A continuación, se explicará con detalle las estrategias utilizadas en esta etapa.

4.9. Liner Mixed Effects (LME)

Existen dos aspectos del análisis longitudinal de neuroimagenes que deben ser correctamente modelizados. La tendencia en relación con el tiempo y la covarianza existente entre los diversos scans del mismo sujeto.

Linear Mixed Effects (LME) utiliza el paradigma de la regresión lineal, es decir, aproxima la medida como una combinación lineal de un grupo de variables independientes. La trayectoria temporal es determinada por la contribución del tiempo y/o las variables dependientes del tiempo, también llamadas covariables.

Existen tres causas potenciales de variedad que influyen en la correlación de las estructuras en los análisis longitudinales:

1. Variación natural existente entre sujetos.
2. Inherentes cambios biológicos.
3. Errores de medida.

El método LME propone la introducción de estos efectos dependientes del sujeto, llamados efectos aleatorios. Una característica que hace único a este método es la **posibilidad de analizar de manera diferenciada la relación presente entre los sujetos y las variaciones intra-sujeto.**

La ec. 4.4 resume el modelo LME:

$$Y_i = X_i\beta + Z_ib_i + e_i \quad (4.4)$$

Donde:

- Y_i es un vector de $n_i \times 1$, que representa la cuantificación de los marcadores para i sujetos. Siendo n_i el número específico de marcadores.
- X_i es una matriz de $n_i \times p$ que representa los efectos fijos.
- β es un vector $p \times 1$ de coeficientes de regresión para los efectos fijos. **Los parámetros β son iguales para todos los individuos**, por lo que representan la media de la población.
- Z_i es una matriz de $n_i \times q$ donde $q \leq p$. Esta matriz es llamada matriz de efectos aleatorios, las columnas de la misma son subgrupos de la X_i .
- b_i es un vector de $q \times 1$ que representa los coeficientes de los efectos aleatorios. Estos parámetros de la regresión reflejan como el subgrupo de características Z_i alejan el sujeto i -ésimo de la población. **Los coeficientes de b_i junto con los valores fijos, constituyen los coeficientes de regresión que representan trayectoria de un sujeto específico**
- e_i es un vector de $n_i \times 1$ que representa los errores de medida.

La esperanza matemática formaliza la idea de valor medio de un fenómeno aleatorio. Es notable la distinción que se debe realizar entre la esperanza condicional y la esperanza marginal dadas en el modelo LME:

La **esperanza condicionada** de una variable aleatoria es el valor esperado de dicha variable respecto a una distribución de probabilidad condicional.

$$E(Y_i|b_i) = X_i\beta + Z_ib_i \quad (4.5)$$

Mientras que la **esperanza marginal**:

$$E(Y_i) = X_i\beta \quad (4.6)$$

Existen alternativas al LME como ANOVA (Análisis de varianza) y modelo lineal general (Análisis de resumen de las medidas). Pero esos métodos no modelan adecuadamente la covarianza entre los scans del mismo sujeto y no permiten la introducción de imperfecciones en los datos como variabilidad en el tiempo de adquisición de los scans. Por lo tanto, la principal ventaja de LME es que no es requerido (aunque por supuesto recomendable) que los sujetos tengan en común el número de scan y la periodicidad de los mismos. Consecuentemente, este método posee una robustez mucho mayor a los nombrados anteriormente.

4.9.1. Modelos LME basados en vóxel o vértices (V-LME)

Una aproximación simple al estudio masivo univariante de datos longitudinales es aplicar el modelo LME a cada localización espacial (vóxel o vértice) por separado. Al hacer uso del modelo LME, cada trayectoria será una combinación de los efectos fijos determinados por la población junto con los efectos aleatorios provenientes de

cada individuo.

El modelo utilizado continua haciendo referencia a la ecuación 4.4, pero en este caso el vector columna Y_i son las medidas a lo largo del tiempo de una simple localización espacial mientras que el nombrado anteriormente son regiones anatómicas o regiones de interés (ROIs). Continuando con las semejanzas, el modelo V-LME realiza una simplificación al asumir que las trayectorias son lineales. Esto se puede realizar debido a que los análisis longitudinales están limitados en el tiempo, por lo que son solo capaces de exponer tendencias simples.

4.9.2. Modelo LME espaciotemporal (ST-LME)

Se parte de la asunción que la covarianza temporal de las estructuras en el modelo LME es compartida por todos los puntos (vóxeles o vértices) dentro de una región homogénea de interés. Ante esta asunción se realizan dos preguntas:

- ¿Cómo se puede dividir la imagen en regiones de interés homogéneas, considerando como homogéneo, que la covarianza en las medidas de cada sujeto dentro de cada región tenga una componente temporal similar y una relativamente suave componente espacial?
- ¿Cuál es el modelo para la dependencia espacial?

Segmentación de la imagen en regiones homogéneas

En cada región se asume que la covariación temporal es compartida a través de todos los vóxeles o vértices. El objetivo es encontrar un algoritmo que segmente automáticamente los datos en estas regiones. Denotando como R el dominio de la imagen completa, el objetivo de este algoritmo es dividir R en r regiones homogéneas. Cabe destacar que r no esta predefinida. A continuación se define las condiciones lógicas del problema:

1. $\cup_{i=0}^r R_i = R$
2. R_i es una región conectada $\forall i=1,2,\dots,r$
3. $R_i \cap R_j = \emptyset, \forall i \neq j$
4. $H(R_i) = true, \forall i=1,2,\dots,r$
5. $H(R_i \cup R_j) = false, \forall i \neq j$

donde $H(R_i)$ es una condición lógica de homogeneidad y \emptyset representa la región vacía.

El algoritmo de segmentación tiene dos estados. El primero, la imagen completa R es recursivamente dividida en regiones más pequeñas hasta que todas las regiones satisfazan la condición de $H(R_i) = true$. Si en alguna etapa no se satisface la condición, la subregión se vuelve a dividir hasta que satisfazca la condición. El segundo estado es combinar cada subregión siempre que la región resultante se mantenga aun homogénea, si no es así estas dos regiones no pueden ser combinadas.

Existe multitud de condiciones lógicas de homogeneidad, pero en este caso se considera $H(R_i) = true$ cuando se cumplen estos dos criterios

- Más del 95 % de los vértices de la región contribuyen en menos de k (k es un parámetro predefinido) desviaciones estándar de la media de la región.
- La correlación de los mínimos cuadrados ordinarios (*Ordinary Least Squares*, *OLS*) dentro de R_i debe ser mayor que 0.5.

El parámetro k determina como de similar debe ser la covarianza dentro de los componentes de una región para que se pueda asumir que sus valores son iguales. Un valor relativamente alto (e.g. $k=2.5$) repercutirá en grandes regiones, reduciendo la precisión del modelo. Si el valor de $k=0$ se reduce el ST-LME a V-LME puesto que cada vértice va a considerarse una región independiente. En general, un aumento de k supone un aumento en el poder estadístico, pero también aumenta el error de tipo 1. Bernal [5] recomienda la utilización de valores de k entre 1 y 2, siendo su valor por defecto de 2, al controlar el error de tipo 1 mientras que se obtiene una alta sensibilidad. El algoritmo de separación en la segmentación puede realizarse de diversas maneras. Por ejemplo, imágenes Euclídeas pueden separarse recursivamente en cuadrículas. Para el análisis de la superficie se utilizan el sistema de coordenadas esféricas.

Modelo para la dependencia espacial

Partiendo de una región homogénea nos focalizaremos en el modelo de esta, puesto que cada región será modelizada de manera independiente. Suponiendo g la región de interés y v_g el número de vóxeles o vértices de esa región. Y_i será un vector columna ($n_i v_g$) $\times 1$ que representa las medidas para una región específica g de un

$$\text{sujeto } i. Y_{ig} = \begin{pmatrix} Y_{ig1} \\ Y_{ig2} \\ \vdots \\ Y_{igv_g} \end{pmatrix}$$

El modelo de covarianza de Y_{ig} viene dado por:

$$COV(Y_{ig}) = W_{ig} - G_g \otimes \Sigma_{ig} = \begin{pmatrix} G_{g11}\Sigma_{ig} & G_{g12}\Sigma_{ig} & \cdots & G_{g1v_g}\Sigma_{ig} \\ G_{g21}\Sigma_{ig} & G_{g22}\Sigma_{ig} & \cdots & G_{g2v_g}\Sigma_{ig} \\ \vdots & \vdots & \ddots & \vdots \\ G_{gv_g1}\Sigma_{ig} & G_{gv_g2}\Sigma_{ig} & \cdots & G_{gv_gv_g}\Sigma_{ig} \end{pmatrix} \quad (4.7)$$

donde \otimes denota el tensor de Kronecker. En una región de un sujeto específico la matriz G_g es una matriz $v_g \times v_g$ que modeliza la correlación espacial. Un ejemplo de G_g particularmente útil debido a que ajusta adecuadamente a las medidas de MRI estructurales y que ofrece un buen control de los errores de tipo 1 viene determinada por la siguiente matriz:

$$G_g = \begin{pmatrix} 1 & e^{-a_g d_{12} - b_g d_{12}^2} & \cdots & e^{-a_g d_{1v_g} - b_g d_{1v_g}^2} \\ e^{-a_g d_{21} - b_g d_{21}^2} & 1 & \cdots & e^{-a_g d_{2v_g} - b_g d_{2v_g}^2} \\ \vdots & \vdots & \ddots & \vdots \\ e^{-a_g d_{v_g1} - b_g d_{v_g1}^2} & e^{-a_g d_{1v_g} - b_g d_{1v_g}^2} & \cdots & 1 \end{pmatrix} \quad (4.8)$$

donde $a_g, b_g \geq 0$ son parámetros del modelo desconocidos y $d_{jk} \geq 0$ representa la distancia entre los vóxeles/vértices j y k de la región g . En definitiva, la distribución de una serie de medidas dentro de una región es:

$$Y_{ig} \sim N(X_{ig}\beta_g, W_{ig}) \quad (4.9)$$

Una vez definido el modelo, se utiliza probabilidad máxima restringida (*Restricted Maximum Likelihood, REML*) para estimar los valores de los parámetros del modelo asociados con cada región g i.e D_g, a_g . Mientras que b_g se obtiene mediante maximización.

Finalmente, un modelo Satterthwaite es utilizado para computerizar el p-valor para la hipótesis nula para cada vóxel/vértice usando las estimaciones temporales de los parámetros.

Cabe subrayar que los atributos utilizados para el paso de segmentación no depende de la hipótesis testada que sigue al paso de la parcelización.

Datos longitudinales

El mapa de espesor cortical es computerizado automáticamente por FreeSurfer para cada sujeto en un momento determinado, el cual es transferido a una plantilla mediante transformaciones no lineales. Por último, cada espesor en el mapa es suavizado aplicando un método iterativo promediando a los vecinos más cercanos, así se aproxima al kernel de suavizado Gaussiano de la plantilla de alta resolución de FreeSurfer, fsaverage.

4.9.3. Residuos longitudinales

El modelo LME creado es de primer orden (lineal) y viene dado por:

$$y_{ij} = (\beta_r + b_{ri}) + (\beta_s + b_{si})t_{ij} + e_{ij} \quad (4.10)$$

Donde y_{ij} representa el valor del biomarcador de un sujeto i en su scan número j , t_{ij} es el tiempo respecto al baseline (en años), $\beta_r + b_{ri}$ representa el corte con el eje vertical (donde el tiempo es cero, por lo que nos encontramos en el baseline), y $\beta_s + b_{si}$ es la pendiente.

La trayectoria media del grupo clínico se calcula eliminando los efectos aleatorios (variaciones de cada individuo), quedando solo los efectos fijos. Por tanto la trayectoria viene dada por:

$$y_j = (\beta_r) + (\beta_s)t_j + e_j \quad (4.11)$$

El grado de atrofia del sujeto i viene dado por la división de la pendiente entre la intersección:

$$\Lambda_i = \frac{\beta_s + b_{si}}{\beta_r + b_{ri}} \quad (4.12)$$

La entrada para el clasificador va a ser los residuos longitudinales de la trayectoria, los cuales son la media del valor real del biomarcador menos la esperanza marginal del mismo (Ec. 4.6)[7]. En otras palabras, el error medio cometido al intentar predecir

el valor del biomarcador utilizando las características fijas de un individuo en la trayectoria de un grupo clínico. Viene definida por:

$$l_i = \frac{1}{n_i} \sum_{j \in S} (y_{ij} - X_{ij}^k \beta) \quad (4.13)$$

Algoritmo 3 Algoritmo de Residuos Longitudinales

Entrada: Valor de las características de cada individuo y su grupo clínico

Salida: Residuos del training set y del test set

- 1: **para todo** Las características **hacer**
 - 2: **si** Se tiene el modelo LME para cada característica **entonces**
 - 3: **devolver** Error medio longitudinal del test, al afirmar que el paciente esta en un grupo determinado
 - 4: **fin si**
 - 5: **fin para**
-

Algoritmo de clasificación

Los residuos longitudinales son analizados mediante Análisis Discriminante Lineal (*Linear discriminant analysis, LDA*)

Algoritmo 4 Algoritmo de clasificación

Entrada: Residuos longitudinales del training y del test

Salida: Valores de precisión, sensibilidad y especificidad del modelo

- 1: Obtención del modelo con los errores del training set
 - 2: **para todo** Los pacientes del test **hacer**
 - 3: Creación de los modelos con los residuos longitudinales del training.
 - 4: Clasificación del test mediante los residuos longitudinales.
 - 5: **fin para**
-

4.10. Realimentación entre selección de características y clasificación

Ante la complejidad que supone implementar de manera automática tanto la selección de características como la clasificación de las mismas se han implementado estas técnicas:

4.10.1. Bootstrap

Bootstrap es un método de remuestreo propuesto por Bradley Efron en 1979 [14]. Este método pretende estimar la distribución de una población mediante un conjunto de remuestras de dicha población. El remuestreo o el muestreo con reposición observa una población finita generando muestras aleatorias para a partir de estas, estimar características poblacionales y hacer inferencia sobre la población muestreada. Estas muestras tomadas son devueltas a dicha población, por lo tanto tienen las mismas posibilidades de ser elegidas que el resto de muestras. En contraste con adquirir la variable estadística deseada directamente de las muestras, esta técnica permite disminuir en gran medida la desviación típica y en consecuencia obtener

intervalos de confianza mejores.

Algoritmo 5 Algoritmo de bootstrap

Entrada: Datos corticales y subcorticales

Salida: Valores de sensibilidad, especificidad, precisión y ACU con Intervalos de Confianza

- 1: Creación de las covariables del modelo (X)
 - 2: Selección las mejores características
 - 3: **para todo** Numero de remuestras **hacer**
 - 4: Separa los sujetos en training y test
 - 5: Crear el modelo con training
 - 6: Testear el modelo con test
 - 7: **fin para**
 - 8: Crear las distribuciones infriendo en los datos de las muestras
-

En la primera implementación se cometió un error al utilizar la misma información para seleccionar las características y para realizar el análisis, lo cual supone un error de double dipping. Nikolaus Kriegeskorte destaca que a pesar de que el double dipping es ampliamente conocido por la comunidad científica, es una práctica comúnmente utilizada en sistemas de neuroimagen [26].

Para solventar este problema es necesario utilizar información independiente en el análisis y en la selección de características. Se proponen las siguientes soluciones que evitan este problema de double dipping:

- Seleccionar las características utilizando los sujetos HC y AD. Aunque obviamente son grupos diferentes a los buscados, la anatomía del cerebro de un sujeto sMCI se asemeja bastante a un HC mientras que las características anatómicas de un sujeto pMCI se asemeja a un AD. Por lo tanto, es esperable encontrar que las diferencias entre HC y AD sean similares a las existentes entre sMCI y pMCI.
- Seleccionar las características previamente sin utilizar información en el testeo del modelo.
- Metodología de validación cruzada (*Cross Validation, CV*)

4.10.2. Validación cruzada (CV)

La validación cruzada es un modelo utilizado para evaluar la actuación de algoritmos de machine learning. Se basa en realizar predicciones con nuevas fracciones de información, los cuales no habían sido utilizados para entrenar o seleccionar las características del modelo (evitando el double dipping).

Cada ronda de CV involucra una partición aleatoria, separando entre training set y test set. El training set es utilizado para seleccionar las mejores características, entrenar y supervisar el aprendizaje del algoritmo. Mientras que el testing set es utilizado para evaluar la actuación del modelo. Este proceso se repite k veces, finalmente el indicador que muestra la actuación del modelo es la media del error. Las técnicas de CV más utilizadas son:

- k-fold. Se divide la información en k grupos creados aleatoriamente de las mismas dimensiones. Un subset es usado para validar el modelo mientras que todos los otros para crearlo. Este proceso se repite k veces para que todos los grupos de subset sean utilizados para validar el modelo.
- Holdout (Retener). Dividir los datos en dos grupos con un ratio específico entre training y test.
- Leaveout. Las particiones se divide de la misma manera que k-fold siendo k el número de observaciones. También se le conoce como *Leave one out cross-validation*.

Otra ventaja de la estrategia k-fold CV es eliminar el overfitting en el modelo y optimizar parcialmente la estimación de la actuación del modelo. En contraste, el principal inconveniente del CV es que es muy costoso computacionalmente puesto que tanto el entrenamiento como la validación se realizan multitud de veces. Como cada partición es independiente este análisis se puede realizar en paralelo para aumentar la velocidad del proceso.

El procedimiento utilizado en este proyecto se basa en **dos bucles de k-fold CV**. El bucle interior se encarga de la selección de características y la creación del modelo, mientras que el exterior está diseñado para obtener la actuación del modelo, la Figura 4.10.2 facilita en gran medida la comprensión de esta estrategia. Esta aproximación viene inspirada por la metodología de Korolev [25] que utilizaba el bucle interno para seleccionar el número de características óptimo para su modelo. Se ha utilizado esta metodología debido a que Korolev reporta que la diferencia de BAR entre validación y test es menor 3% para todos los modelos, por lo que este procedimiento de k-fold con dos bucles prevenía al modelo de overfitting.

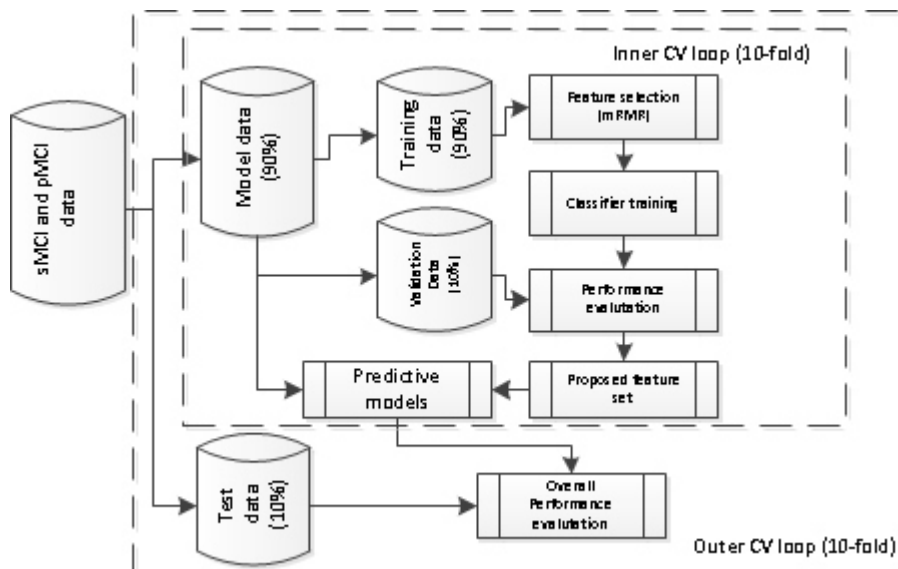


Figura 4.5: Esquema de la estrategia de doble k-fold CV

En el algoritmo 6 se muestra la implementación del doble k-fold CV, mientras que la función Choose Best utilizado en esa implementación se encuentra desarrollado

Algoritmo 6 Método doble k-fold CV

Entrada: Datos corticales y subcorticales**Salida:** Valores de sensibilidad, especificidad, precisión y ACU de los mejores marcadores

- 1: Creación de las covariables del modelo (X)
 - 2: **para todo** Número de loops exteriores **hacer**
 - 3: División de los datos entre selection-set y test-set
 - 4: **para todo** Número de loops interiores **hacer**
 - 5: División del selection-set entre selection-train y selection-test
 - 6: Selección de las 5 mejores características para cada dimensión con los datos de selection-train
 - 7: Creación de los modelos LME con las características seleccionadas con todas las visitas y evaluación de los mismos solo con los datos del selection-test en el baseline.
 - 8: **fin para**
 - 9: Selección de la mejor característica del loop interior mediante el algoritmo Choose Best
 - 10: Creación del modelo LME con selection-set y evaluación con test-set
 - 11: **fin para**
 - 12: Obtención de las frecuencias más utilizadas y realización de la media de sus evaluaciones.
-

en en el algoritmo 7.

Como se ha comentado anteriormente el algoritmo de k-fold CV es muy costoso computacionalmente, en consecuencia, la metodología de doble k-fold CV es exponencialmente más costosa. Por lo tanto se ha buscado optimizar al máximo este algoritmo intentando minimizar el tiempo de computación. Entre las diversas mejoras realizadas se debe destacar las siguientes:

- Solo se calcula el modelo LME de las características seleccionadas.
- Se agrupan las características de todas dimensiones calculando tan solo una vez por cada loop el modelo LME y los residuos, suprimiendo repetir innecesariamente el calculo de las mismas características para cada dimensión.
- Una vez calculados los residuos para cada característica en cada visita se agrupan en función del número de visitas. Esto se realiza de manera muy simple tan solo variando el valor de máximo de i en la ecuación 4.13.

Algoritmo 7 Algoritmo de Choose Best

Entrada: Valores de sensibilidad, especificidad, precisión y frecuencia de los marcadores**Salida:** Marcador óptimo

- 1: Ordenar los marcadores por la precisión
 - 2: Suponer el 1º como la solución óptima
 - 3: Siguiendo el orden anteriormente definido.
 - 4: **mientras** No se ha encontrado un marcador idóneo **hacer**
 - 5: **si** Sensibilidad > 70 % Y Especificidad > 70 % Y Frecuencia > 2 **entonces**
 - 6: **devolver** Marcador idóneo encontrado
 - 7: **fin si**
 - 8: **fin mientras**
-

Para finalizar este capítulo se debe resaltar que el flow-char seguido en el proyecto 4.1, puede ser conceptualmente fácilmente extrapolable a cualquier otro proyecto de

investigación. Además, cabe subrayar las grandes ventajas expuestas de un estudio longitudinal respecto a un análisis cross. Por último, se han expuesto la base teórica donde se sustenta toda la investigación junto con los algoritmos abstractos que ejemplifican la implementación de estos conceptos teóricos.

Capítulo 5

Resultados

En este capítulo se muestran los resultados obtenidos tras los diversos ensayos realizados. Los principales objetivos buscados son:

1. Determinar si el análisis longitudinal realizado ha sido adecuado, esto se comprobará mediante un proceso de control de calidad.
2. Mostrar los resultados obtenidos en el análisis masivo, donde destacan los mapas de q-valor junto con los clusters determinados.
3. Obtener las mejores covariables para el modelo LME.
4. Precisar la dimensión óptima de marcadores.
5. Definir la relación entre los puntos en el tiempo (número de visitas) y la fiabilidad de la predicción.

Una vez determinados estos objetivos se mostrará las mejores combinaciones obtenidas.

5.1. Control de calidad

En esta sección se muestran los resultados obtenidos tras realizar el control de calidad.

Correlación entre pares de scans del mismo sujeto

Como se puede observar en la Figura 5.1 existen dos outliers (133 S 0913 y 068 S 0476). Esto representa que se ha producido un error en la extracción de características, por lo que estos sujetos han sido revisados manualmente.

5.1.1. Correlación de volumetría hipocampal y del valor de ICV

La Figura 5.2 muestra una clara tendencia lineal, lo cual indica que los valores obtenidos mediante el análisis longitudinal es satisfactorio. Además, los valores de correlación entre ADNIMERGE y el procesamiento realizado son de:

- Correlación del volumen del hipocampo es de 0.981
- Correlación del volumen intracraneal (ICV) es de 0.998

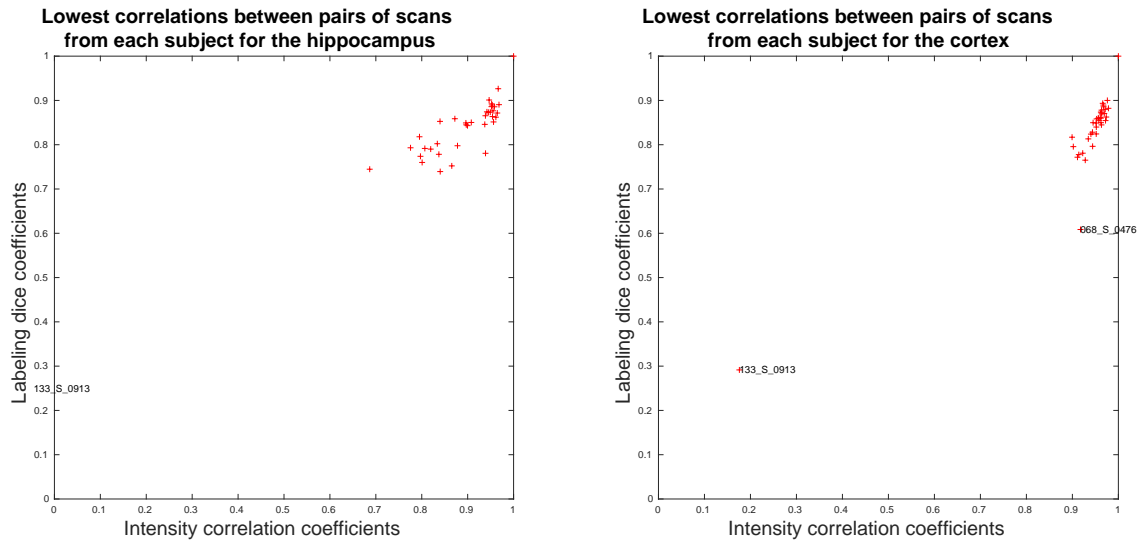


Figura 5.1: Coeficiente de correlación de Pearson

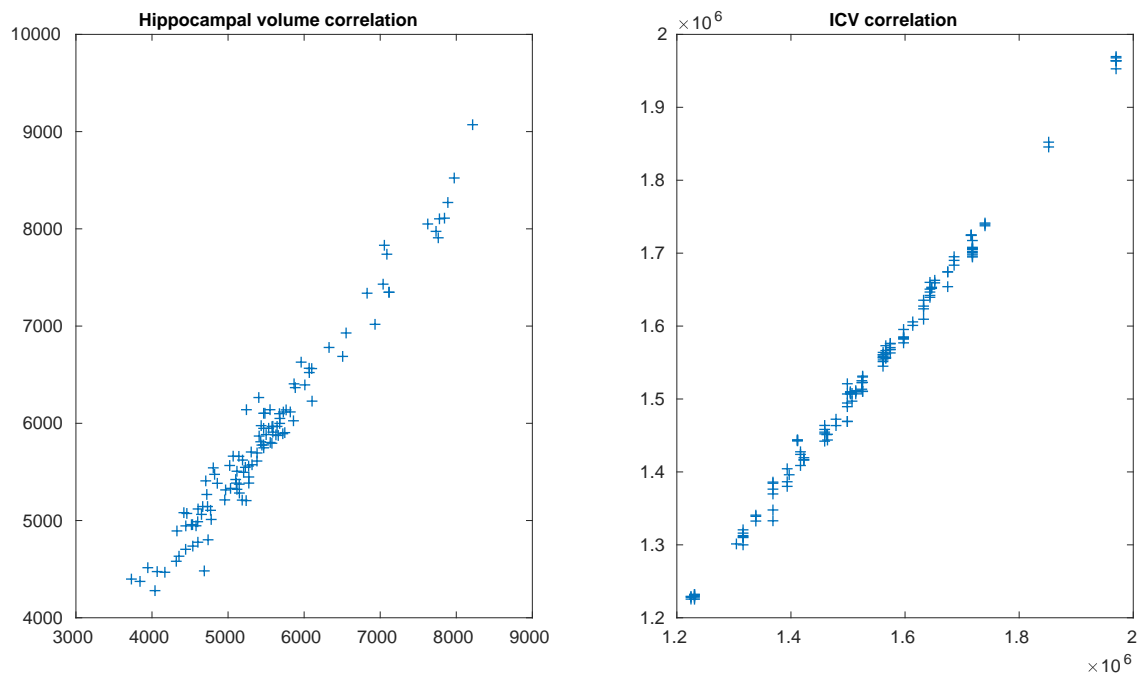


Figura 5.2: Coeficiente de correlación de los datos obtenidos con ADNIMERGE

5.1.2. Comparación con los datos del grupo de investigación de Harvard

Trayectorias de NVH y ETC

La Figura 5.3 junto con la Figura 5.4 muestran las trayectorias de del volumen hipocampal normalizado y el espesor de la corteza entorrhinal, comparando los valores obtenidos en este proyecto con los grupo de Bernal.

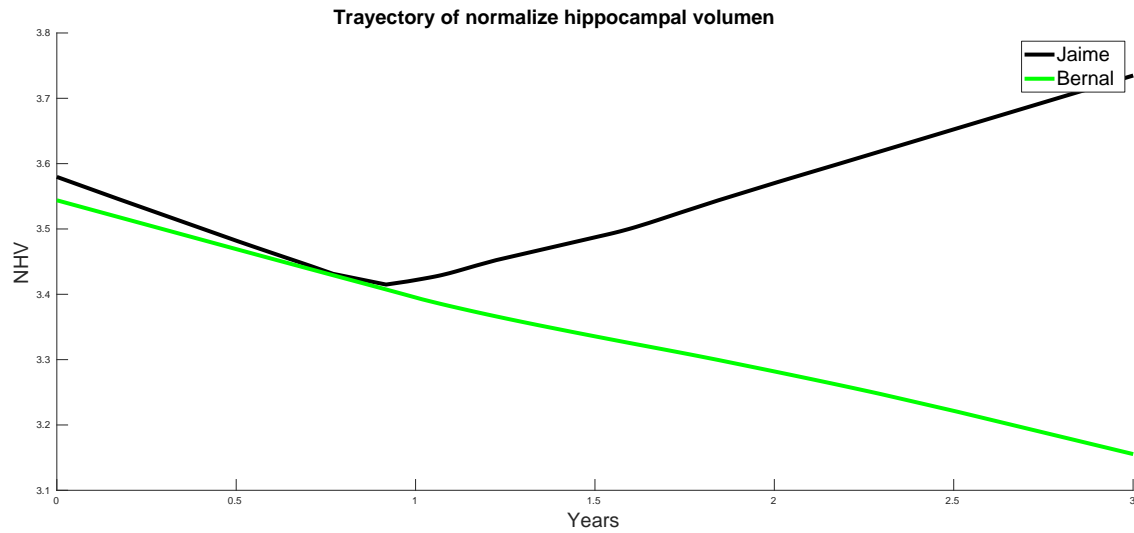


Figura 5.3: Trayectorias de NVH de los sujetos analizados este año

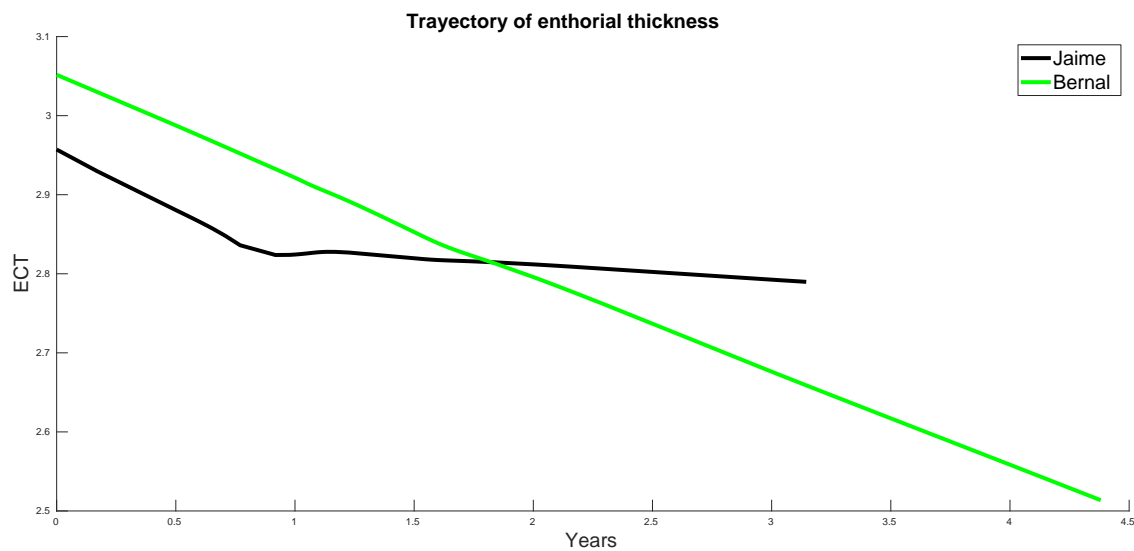


Figura 5.4: Trayectorias de ETC de los sujetos analizados este año

Como se puede apreciar en la figura, desde el baseline hasta 1 año las trayectorias son muy similares a las obtenidas por Bernal, esto implica que las trayectorias en este periodo son adecuadas. Pero a partir de 1 año, estas trayectorias carecen de sentido debido a que muestra una atrofia positiva, en concreto en el caso del volumen hipocampal normalizado. Por ello, se intentará buscar una explicación a este suceso.

Una estrategia muy útil a la hora de afrontar un problema es dividirlo en tareas más simples, comenzando por lo más sencillo hasta llegar a relaciones más complejas. Siguiendo con esta filosofía se han separado todas las trayectorias de los sujetos en

función de del número de scans.

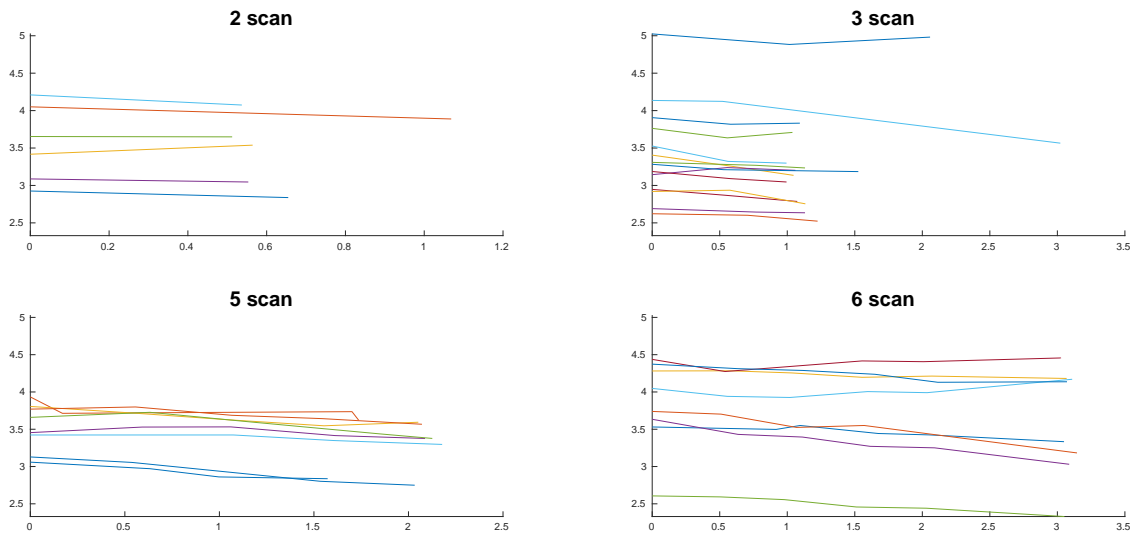


Figura 5.5: Trayectorias de NHV separadas en función del número de scans. En todas las figuras el eje X representa los años, mientras que el eje Y representa la trayectoria del volumen del hipocampo normalizado

La Figura 5.5 muestra resultados muy esperanzadores puesto que no hay ningún sujeto con una atrofia positiva notable, en contraste con lo que se mostraba en la Figura 5.3. Esto quiere decir que el problema no viene de ningún sujeto en concreto que esta contaminando el resultado del grupo con una atrofia muy positiva.

Siguiendo con la filosofía dicha previamente se ha realizado el plot suavizado de las trayectorias en función del número de visitas.

La Figura 5.6 muestra multitud de información. En primer lugar que las trayectorias de 5 y 6 visitas son adecuadas, mientras que la trayectoria de 2 y 3 visitas no tiene sentido. Estos picos no pueden ser representativos, pero ¿cuál es el origen de esta variación si como se ha visto previamente en la Figura 5.5 no existen ninguna trayectoria con estos picos?

Focalizandonos en el grupo formado por 3 visitas, se puede observar en la Figura 5.6 que se tiene una atrofia positiva a partir de 2 años, cuando solo existe un sujeto que sobre pasa esta duración y su atrofia es claramente negativa. Por lo que el problema tiene que estar en la representación. En consecuencia se va a proceder a observar la función dada en la toolbox LME de Bernal `lowess plot`. Esta función realiza un smooth basado en realizar la media de una ventada dada como parámetro de entrada. Por ello, al observar la trayectoria de 3 visitas se considera una atrofia positiva a partir de 2 años puesto que al realizar la media a partir de esta fecha empieza a tener mucha importancia los valores de la trayectoria del único sujeto disponible, la trayectoria tiende a ir estos valores pero como están por encima de la media parece que la trayectoria del conjunto tiene una atrofia positiva. En definitiva

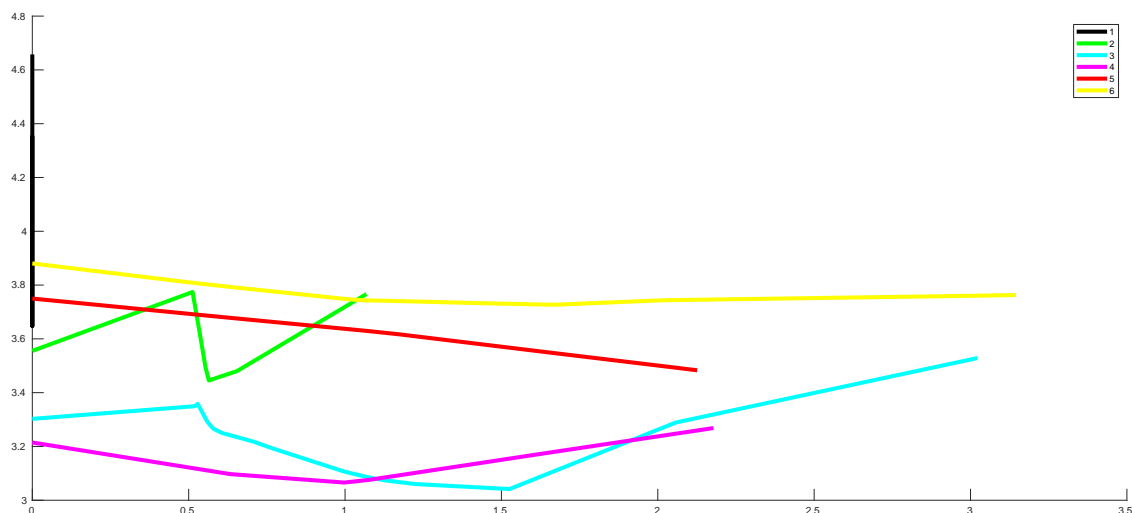


Figura 5.6: Trayectorias de NHV suavizadas en función del número de scans. El eje X representa los años, mientras que el eje Y representa la trayectoria del volumen del hipocampo normalizado

esta representación no muestra las trayectorias adecuadamente.

Una vez definido el problema de manera local (tan solo un sujeto) se puede extrapolar; debido a que como se muestra en la Figura 5.6 y en la Figura 5.5, las trayectorias de los sujetos con 6 visitas tienen valores más altos que la media. En definitiva, sucede exactamente el mismo problema que el mostrado anteriormente, pero en lugar de ser un solo sujeto es un conjunto de ellos. Esto explica la supuesta atrofia positiva mostrada en la Figura 5.3.

Distribuciones de NVH y ECT

Tanto las muestras de Bernal como la de este proyecto son muestras del mismo grupo de población, sujetos pMCI. Por ello, sus distribuciones normales deben de ser similares tanto en el valor del baseline como en la atrofia.

Las Figuras 5.7 y 5.8 muestran distribuciones muy similares a las de Bernal, por lo que pasan esta etapa del control de calidad.

5.2. Unión de los datos obtenidos este año con los procesamientos de otros años

Debido a la tremenda importancia que tiene que los datos sean fiables para que sobre ellos se pueda construir los algoritmos de machine learning, se ha juntado los datos extraídos en este proyecto junto con los obtenidos en años anteriores. Pudiendo así comprobar si la introducción de estas trayectorias inusuales van a causar un efecto indeseado al grupo.

La Figura 5.9 junto con la Figura 5.10, afirma el razonamiento previamente expuesto debido a que la introducción de nuevos sujetos pMCI que no tienen una alto

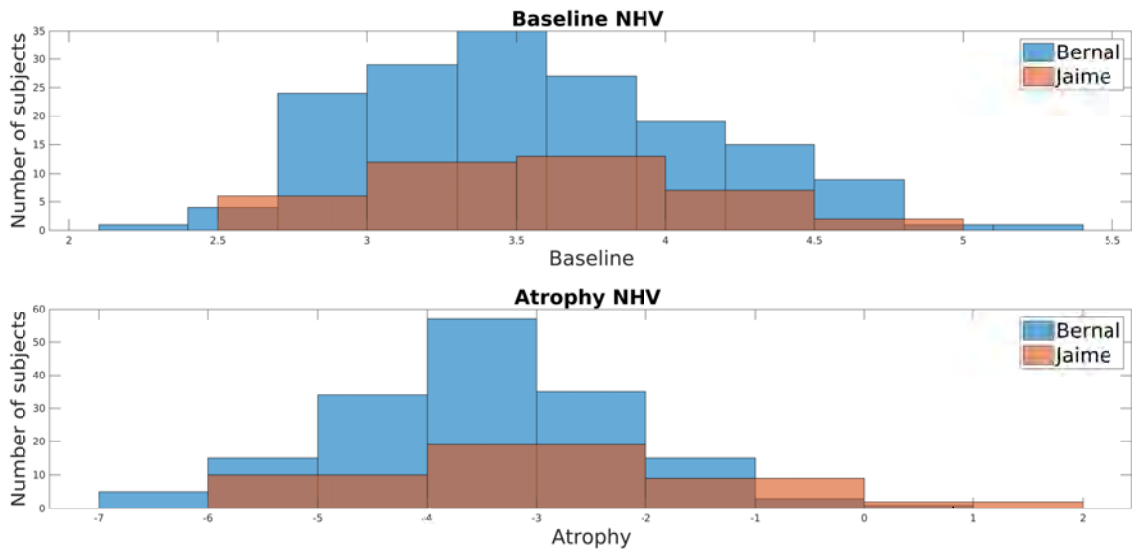


Figura 5.7: Distribución del NHV de los sujetos analizados este año

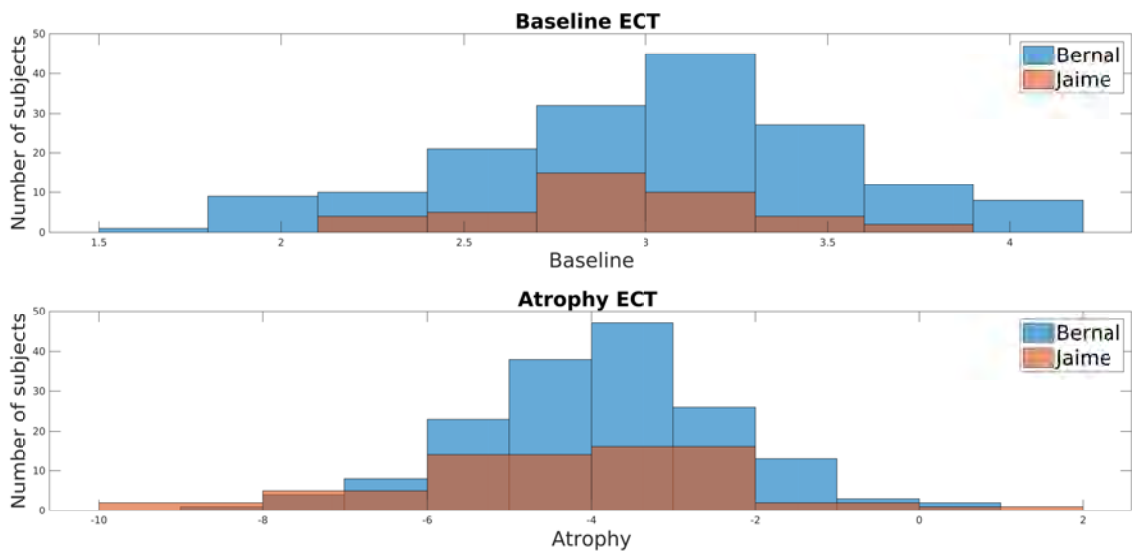


Figura 5.8: Distribución del ECT de los sujetos analizados este año

valor en el baseline suaviza la influencia del grupo de scans de 3 años con un valor alto.

Se debe subrayar la gran diferencia tanto en baseline como en atrofia anual entre los sujetos sMCI y los pMCI, siendo la trayectoria unida muy similar a la dada por Bernal.

Para concluir el control de calidad es necesario observar las nuevas distribuciones, las cuales son muy similares a las obtenidas por Bernal.

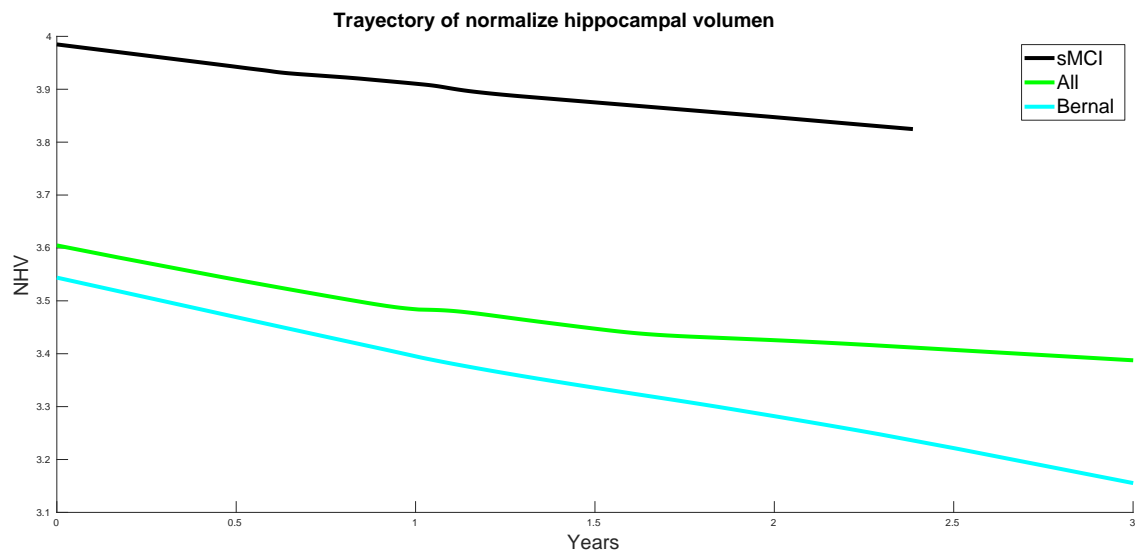


Figura 5.9: Trayectorias de NVH de todos los sujetos

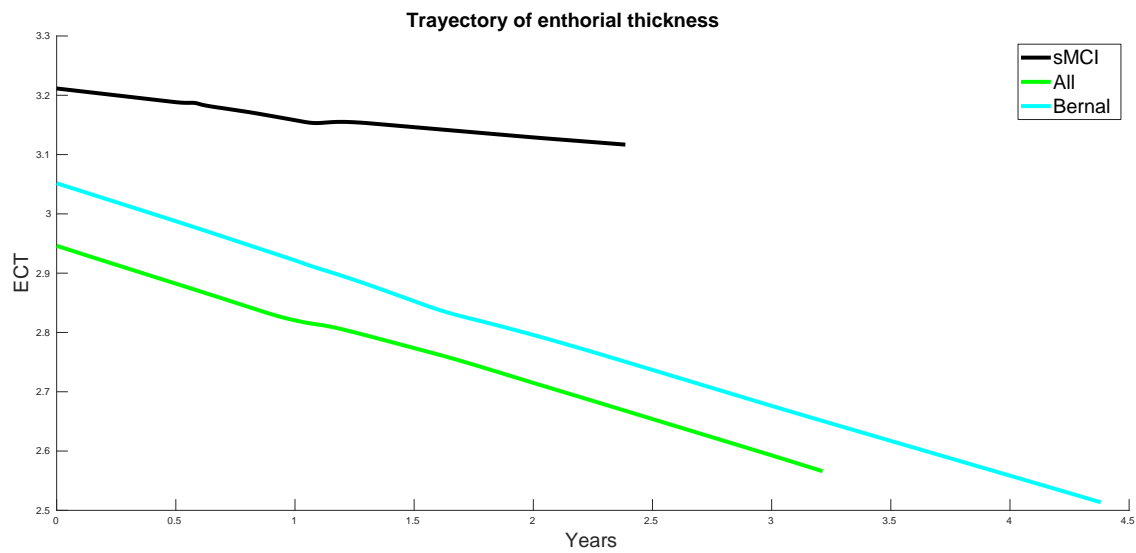


Figura 5.10: Trayectorias de ETC de todos los sujetos

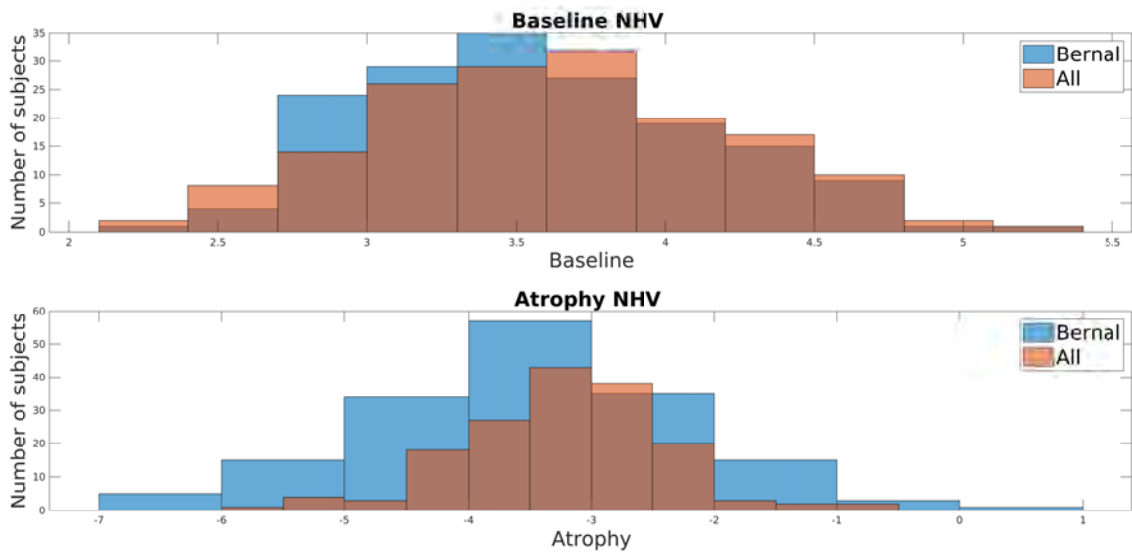


Figura 5.11: Distribución del NHV de todos los sujetos

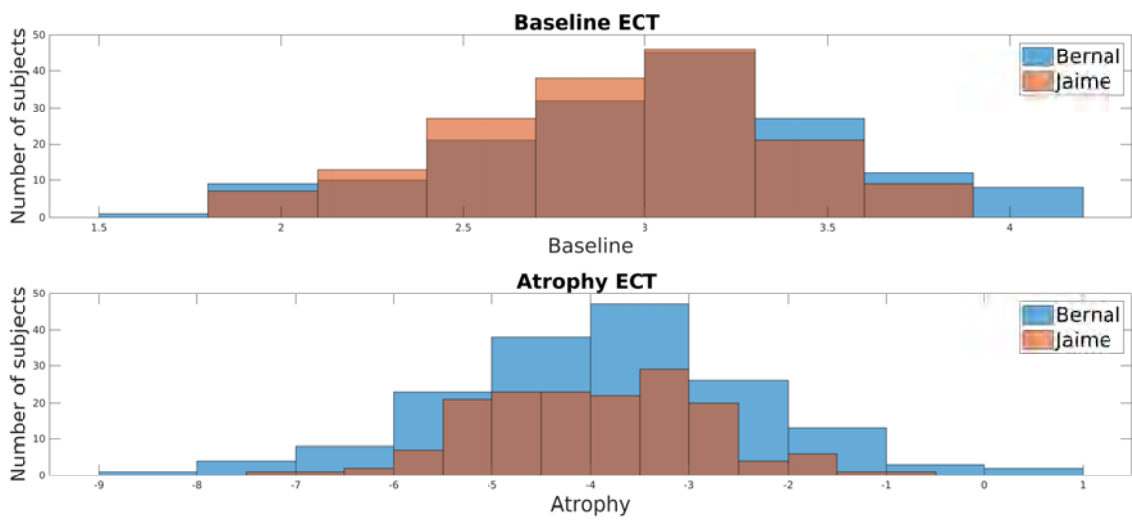


Figura 5.12: Distribución del ECT de todos los sujetos

Tabla 5.1: Tabla comparativa entre Bernal, Jaime y todos. Los datos se presentan junto con la desviación típica

		Bernal	Jaime	All
NHV	Baseline	3.57 (0.56)	3.61 (0.56)	3.62 (0.60)
	Atrophy	-3.52 (1.27)	-2.84 (1.27)	-3.21 (0.81)
ETC	Baseline	3.02 (0.50)	2.95 (0.38)	2.93 (0.43)
	Atrophy	-3.85 (1.54)	-4.19 (2.03)	-3.95 (1.11)

La conclusión obtenida tras el control de calidad es que a pesar de la primera impresión obtenida, **la extracción de características ha sido realizada satisfactoriamente**. La Tabla 5.1 resume los valores obtenidos en el control de calidad.

5.3. Análisis masivo

El análisis masivo se ha realizado en dos grupos:

- El análisis longitudinal de 50 sMCI y 50 pMCI facilitado por Bernal.
- El análisis longitudinal de NC y AD facilitado por el trabajo de años anteriores.

5.3.1. sMCI vs pMCI

Se han creado 8 mascararas mediante la metodología de k-fold obteniendo el p-valor para cada vértice utilizando el modelo ST-LME. Estas máscaras se han unido mediante FDR obteniendo el siguiente mapa de q-valor(véase Figura 5.13). La leyenda se encuentra en escala logarítmica, mostrando el $-\log_{10}(\text{q-valor})$. En consecuencia, cuanto más frío sea el color, más significancia existe entre ambos grupos.

Una vez obtenido estos valores es necesario crear grupos o clusters, agrupando así a los vértices que con características similares. Se ha decidido poner un límite de 2, es decir, $10^{-2} = 0,01$ como q-valor mínimo para considerar significancia y un área mínima de 0.50mm^2 . Se ha obtenido un total de 30 clusters, siendo 19 de ellos pertenecientes al hemisferio izquierdo y 11 al hemisferio derecho, la Figura 5.14 muestra los clusters realizados mientras Tabla 5.2 muestra la relación entre los cluster formados y las diferentes estructuras del atlas de Desikan-Killiany.

Tabla 5.2: Tabla donde se muestra la relación entre los cluster formado a partir de los MCI y las diferentes estructuras del atlas de Desikan-Killiany

Cluster	Estructuras
lh_MCI.cl.1	superiorparietal, inferiorparietal
lh_MCI.cl.2	supramarginal, bankssts, inferiorparietal, superiortemporal
lh_MCI.cl.3	insula
lh_MCI.cl.4	parstriangularis, rostralmiddlefrontal, parsopercularis
lh_MCI.cl.5	insula, lateralorbitofrontal
lh_MCI.cl.6	superiorfrontal, precentral
lh_MCI.cl.7	bankssts, superiortemporal
lh_MCI.cl.8	superiorparietal, inferiorparietal
lh_MCI.cl.9	lateraloccipital, fusiform
lh_MCI.cl.10	postcentral, superiorparietal
lh_MCI.cl.11	supramarginal, insula, transversetemporal, superiortemporal
lh_MCI.cl.12	superiorparietal, precuneus
lh_MCI.cl.13	postcentral, precentral
lh_MCI.cl.14	middletemporal
lh_MCI.cl.15	parsopercularis, precentral
lh_MCI.cl.16	medialorbitofrontal, superiorfrontal
lh_MCI.cl.17	inferiorparietal
lh_MCI.cl.18	inferiortemporal
lh_MCI.cl.19	medialorbitofrontal
rh_MCI.cl.1	middletemporal, inferiortemporal
rh_MCI.cl.2	parstriangularis, insula, lateralorbitofrontal, parsorbitalis
rh_MCI.cl.3	isthmuscingulate
rh_MCI.cl.4	entorhinal, insula, temporalpole, superiortemporal
rh_MCI.cl.5	insula, transversetemporal, superiortemporal
rh_MCI.cl.6	insula, superiortemporal
rh_MCI.cl.7	bankssts, middletemporal, superiortemporal
rh_MCI.cl.8	precuneus
rh_MCI.cl.9	medialorbitofrontal, superiorfrontal
rh_MCI.cl.10	isthmuscingulate, precuneus
rh_MCI.cl.11	parsopercularis

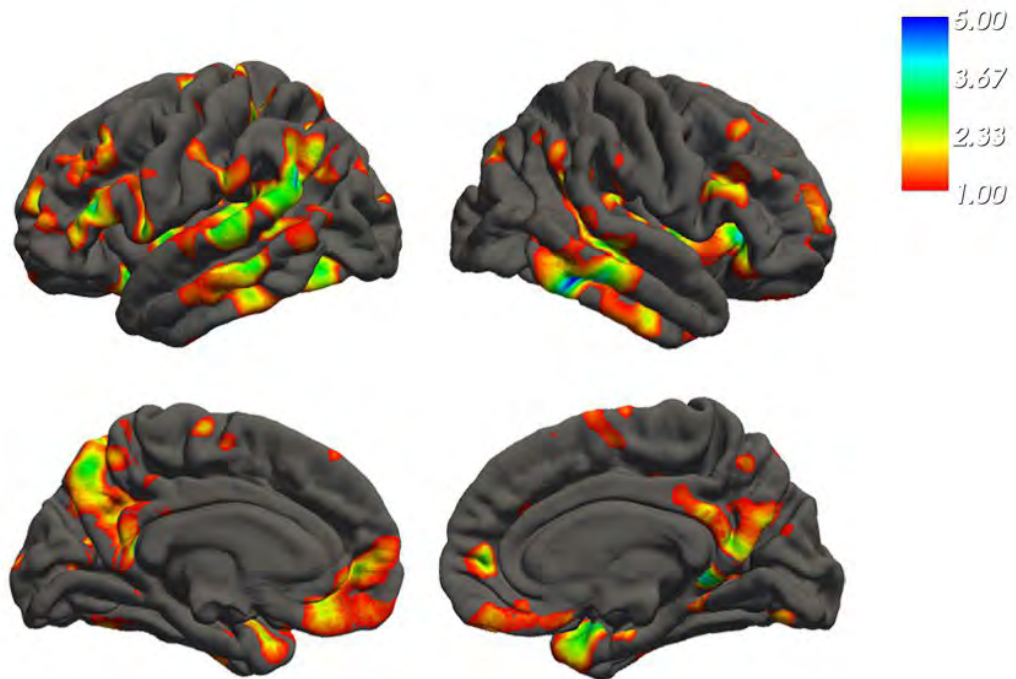


Figura 5.13: Mapa de q-valor entre sujetos sMCI y pMCI. De izquierda a derecha se muestra el hemisferio izquierdo y derecho

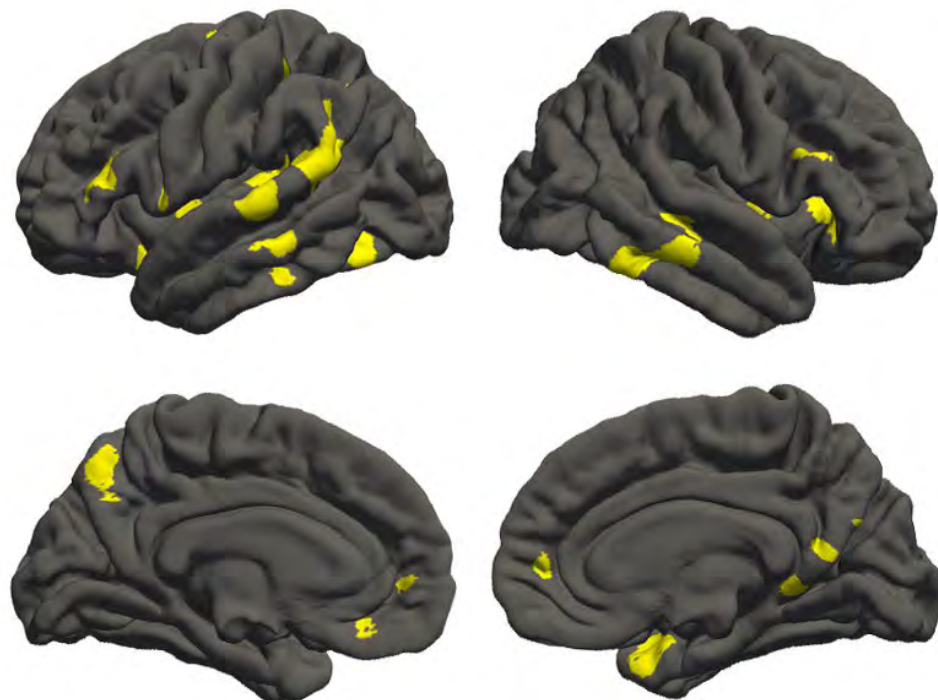


Figura 5.14: Mapa de los clusters realizados entre sujetos sMCI y pMCI.

5.3.2. AD vs HC

De manera análoga a la expuesta anteriormente se han creado 14 mascarar de p-valor entre los grupos de AD vs HC. Estas máscaras se han unido mediante FDR obteniendo el siguiente mapa de q-valor, véase Figura 5.15. Como en la figura anterior a leyenda se encuentra en escala logarítmica, mostrando el $-\log_{10}(q\text{-valor})$. En consecuencia, cuanto más frío sea el color, más significancia existe entre ambos grupos.

Se han creado los clusters con un valor limite de 8, es decir, $10^{-8} = 0,00000001$ como q-valor mínimo para considerar significancia y un área mínima de $0.50mm^2$. Obviamente existe mucha más diferencias entre AD y HC que entre sMCI y pMCI por lo que se ha decidido que el valor sea tan bajo. Finalmente se ha obtenido un total de 25 clusters, siendo 14 de ellos pertenecientes al hemisferio izquierdo y 11 al hemisferio derecho, la Figura 5.16 muestra los clusters realizados mientras que la Tabla 5.3 muestra la relación entre los cluster formados y las diferentes estructuras del atlas de Desikan-Killiany.

Tabla 5.3: Tabla donde se muestra la relación entre los cluster formado a partir de los AD y las diferentes estructuras del atlas de Desikan-Killiany

Cluster	Estructuras
lh_AD.cl.1	entorhinal, parahippocampal, fusiform, temporalpole
lh_AD.cl.2	lingual, fusiform
lh_AD.cl.3	inferiorparietal
lh_AD.cl.4	isthmuscingulate, precuneus, posteriorcingulate
lh_AD.cl.5	precuneus
lh_AD.cl.6	bankssts, middletemporal, superiortemporal
lh_AD.cl.7	inferiortemporal
lh_AD.cl.8	superiorfrontal
lh_AD.cl.9	lateraloccipital
lh_AD.cl.10	inferiortemporal
lh_AD.cl.11	middletemporal, inferiortemporal
lh_AD.cl.12	bankssts, superiortemporal
lh_AD.cl.13	caudalmiddlefrontal
lh_AD.cl.14	superiortemporal
rh_AD.cl.1	entorhinal, insula, parahippocampal, inferiortemporal fusiform, temporalpole, superiortemporal
rh_AD.cl.2	middletemporal, superiortemporal
rh_AD.cl.3	fusiform
rh_AD.cl.4	inferiorparietal
rh_AD.cl.5	middletemporal, inferiortemporal
rh_AD.cl.6	isthmuscingulate, precuneus, posteriorcingulate
rh_AD.cl.7	bankssts, middletemporal
rh_AD.cl.8	superiortemporal
rh_AD.cl.9	superiorfrontal
rh_AD.cl.10	caudalmiddlefrontal
rh_AD.cl.11	inferiorparietal

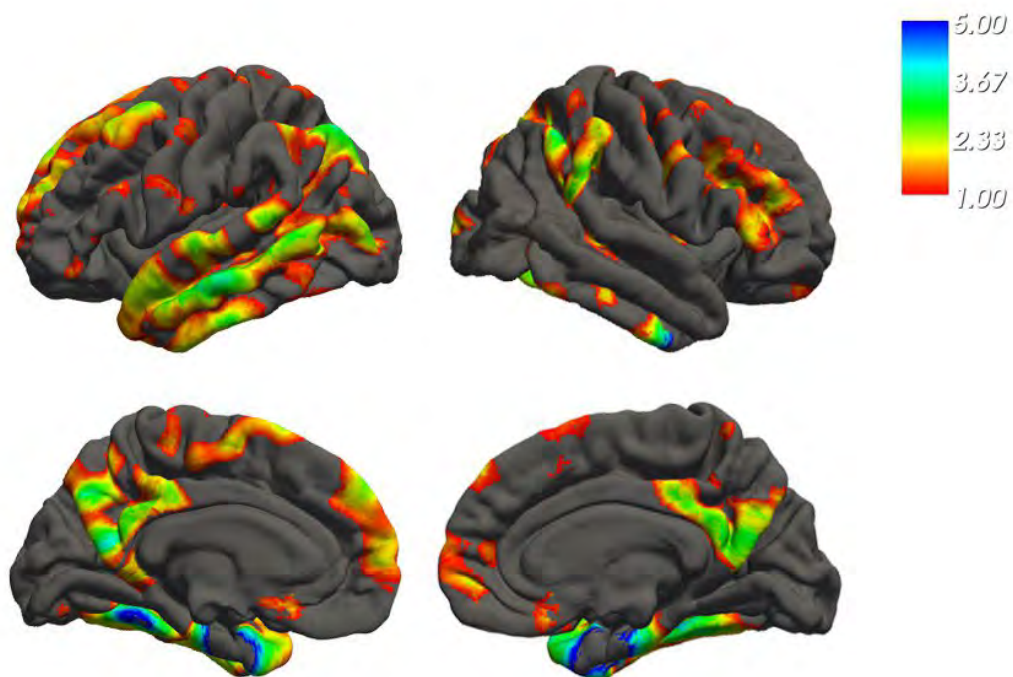


Figura 5.15: Mapa de q-valor entre sujetos HC y AD. De izquierda a derecha se muestra el hemisferio izquierdo y derecho

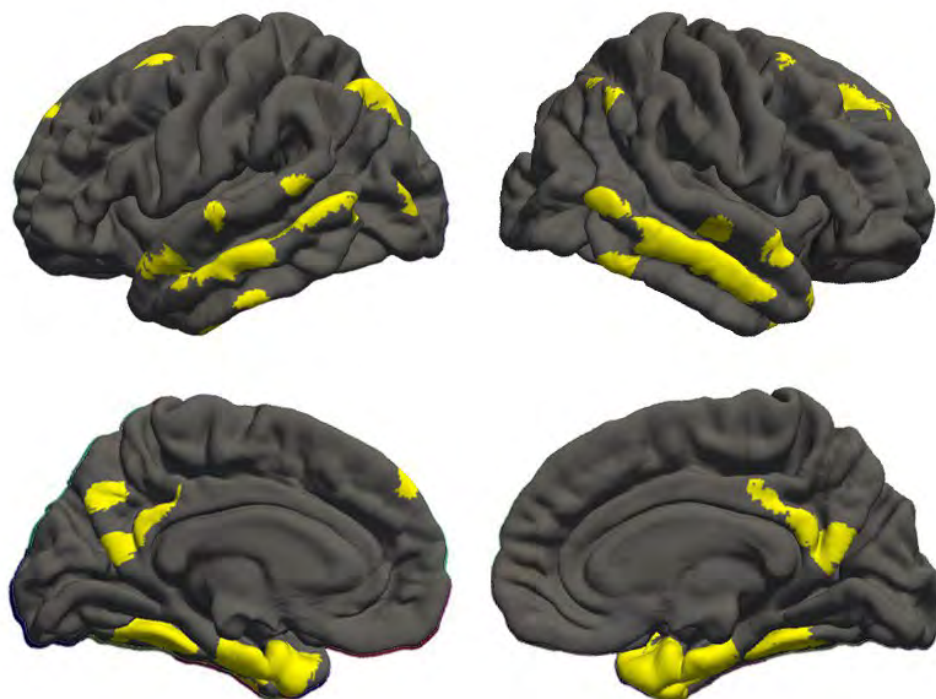


Figura 5.16: Mapa de los clusters realizados entre sujetos HC y AD.

5.3.3. Errores de tipo I y tipo II

El poder estadístico (rechazar la hipótesis nula cuando ésta es incorrecta) es evaluado con diferentes valores de α con el objetivo de buscar el óptimo. No debe extrañar resultados tan bajos en sMCI/pMCI puesto que el número de vértices pertenecientes a los clusters son muy bajos en relación al total. Por ello, es complejo rechazar la hipótesis nula cuando es incorrecta, es decir, afirmar que dos vértices no son iguales cuando realmente sí lo son.

Tabla 5.4: Valor del poder estadístico en función de los diversos valores de α

Valor de α	0.0010	0.0100	0.0500	0.1000
sMCI/pMCI	0.0102	0.0470	0.1406	0.2253
HC/AD	0.4563	0.6386	0.7649	0.8092

En la siguiente Figura 5.17 se muestra mediante plotbox el error de tipo I de cada modelo previamente creado (8 mascaradas creadas con bootstrap). Los círculos representan el error del modelo final. El error de tipo I es rechazar la hipótesis nula cuando esta es cierta, en consecuencia con valores muy bajos de α este error también lo es, puesto que se acepta la hipótesis nula en la mayoría de los casos.

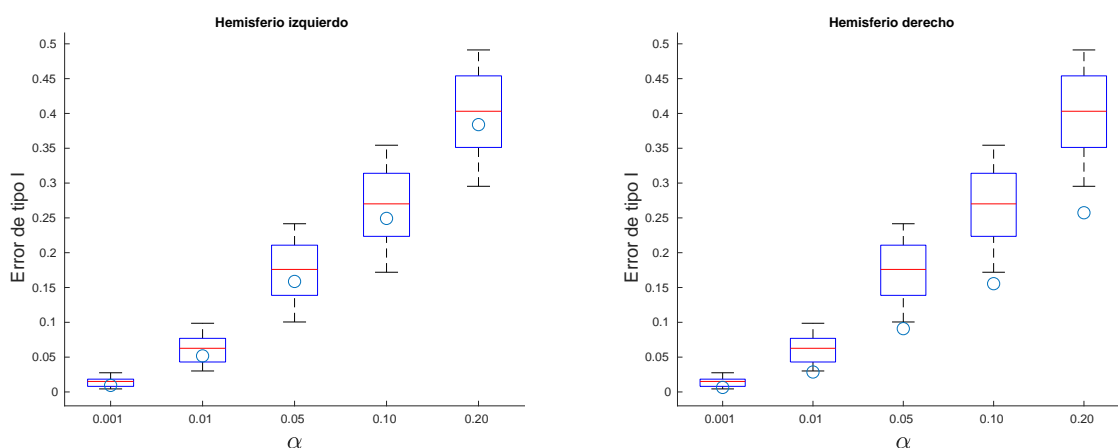


Figura 5.17: Error de tipo I de cada mascara. Se representa la media y las desviaciones típicas del error de tipo I de los 8 modelos. Los círculos representan el error del modelo final

5.4. Primera aproximación al modelo LME

Es ampliamente conocido la influencia de la edad como factor de riesgo para esta enfermedad. Gracias a estudios previos se conoce que existen otros factores de riesgo son el sexo femenino [32], nivel educativo [23], gen APOE-E4 [20], condición de fumador, obesidad y diabetes entre otros.

El modelo LME permite introducir estos marcadores como covariables al modelo. Estas variables clínicas permiten modificar la trayectoria de los marcadores utilizados y por tanto influir de un modo mucho más sutil pero general a la clasificación.

Por ello, seleccionar el conjunto de covariables idóneo es fundamental para la creación de un modelo preciso. En consecuencia, se ha decidido ensayar los mismos marcadores con diversos datos como covariables al modelo LME. Buscando determinar la mejor combinación de factores clínicos.

Este experimento ha sido la primera aproximación al problema, por lo que se ha decidido ensayar muchos modelos y muy diversos entre ellos, pudiendo así extraer la mayor información posible de los mismos. Como se necesitaba la creación de un número considerable de modelos se ha experimentando con bootstrap seleccionando y testeando con los mismos sujetos. Se ha realizado esto para agilizar el tiempo de computación aun sabiendo que se ha realizado double dopping. Por ello, los datos mostrados tienen valores superiores a los verdaderos y los resultados son solo orientativos para la creación del modelo. Aun así las conclusiones extraídas son completamente veraces en relación con la creación del modelo.

Marcadores de MRI estructural con todas la visitas

En la Figura 5.18 se pueden observar los diferentes diagramas con las diversas combinaciones de covariables. Los datos mostrados se han obtenido utilizando todas las visitas.

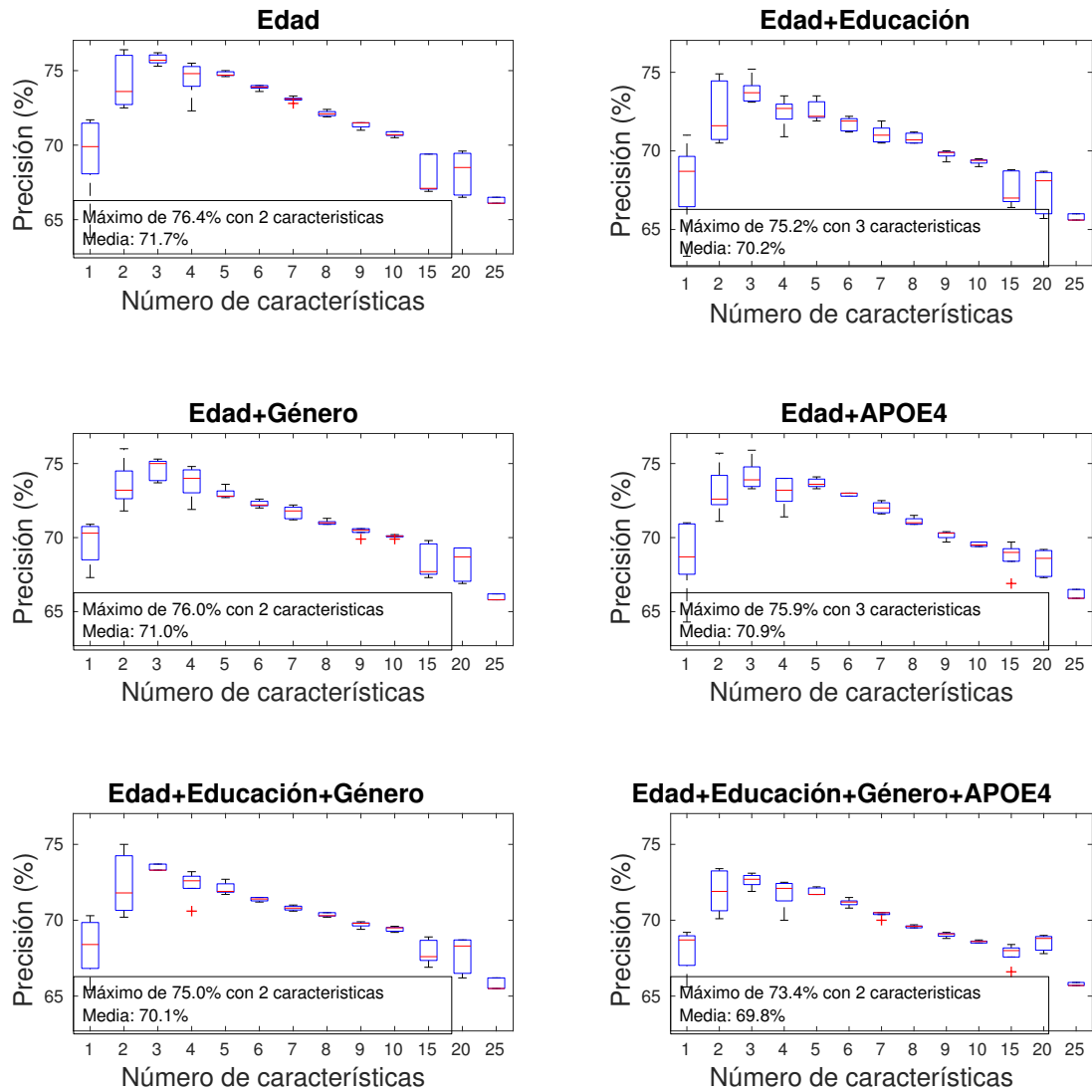


Figura 5.18: Diferentes combinaciones de las covariables con características de MRI en todas las visitas

Biomarcadores de MRI funcional junto con test psicológicos

Con datos longitudinales La Figura 5.19 muestra el resultado de la clasificación cuando se juntan los datos estructurales del cerebro y los valores cognitivos y funcionales.

Con datos del Baseline Aunque como se ha resaltado anteriormente, los test utilizados no son los que determinan el grupo clínico, si es posible que influyan en gran medida en el resultado. En consecuencia, se ha realizado el ensayo solo con datos del Baseline, debido a que en todos los casos los sujetos son MCI. Los resultados obtenidos son muy positivos debido a que reflejan unos valores muy favorables para tratarse tan solo de un punto en el tiempo.

De este primer experimento se han extraído las siguientes conclusiones :

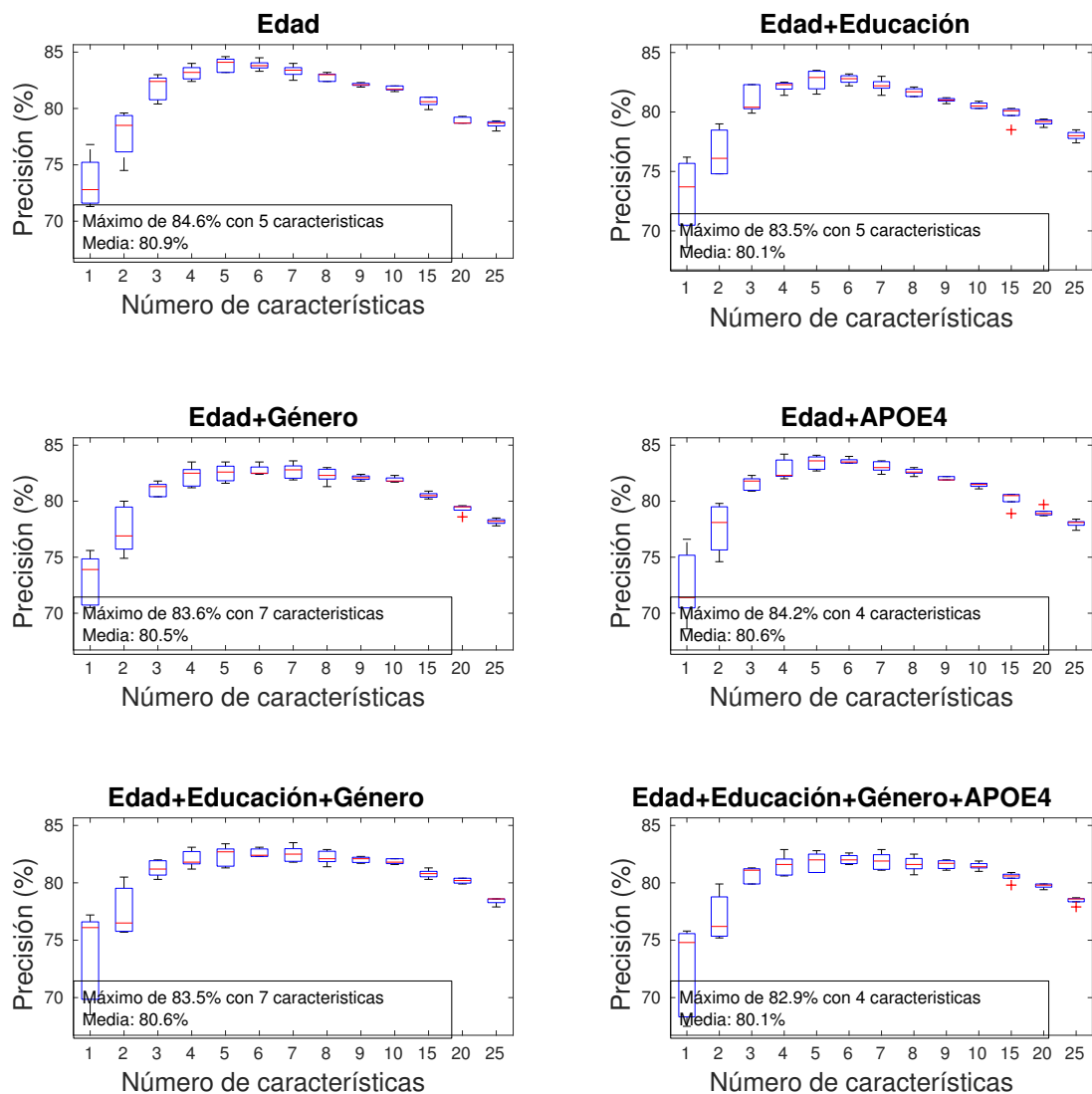


Figura 5.19: Diferentes combinaciones de las covariables con características de MRI junto con CAM en todas las visitas

- En referencia a las covariables del modelo.
 - En todas las ocasiones los mejores resultados se obtienen cuando la Edad es la covariable única.
 - Introducir el APOE4 como covariable al modelo no introduce ninguna mejora al modelo por lo que se descarta.
 - Las combinaciones de Edad+Educación, Edad+Género y Edad+Educación+Género pueden ser candidatas a general el modelo idóneo.
- En referente a los marcadores, se puede observar la gran mejora al introducir los marcadores cognitivos funcionales. Por lo que se introducirán como marcadores en el modelo.
- El número de características óptimo varía entre 3 y 5. Se debe subrayar que siempre se debe elegir el número mínimo de dimensiones que permite obtener

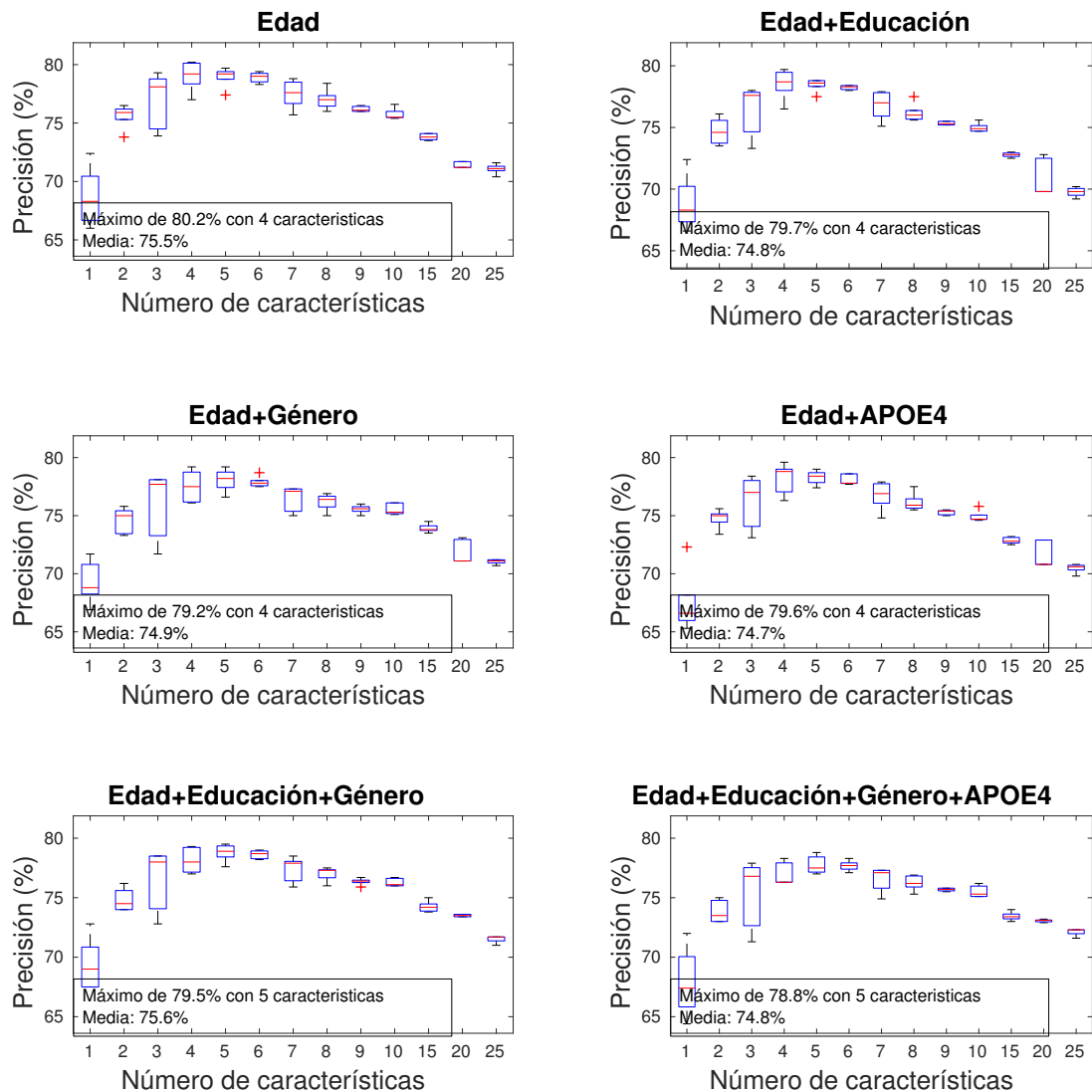


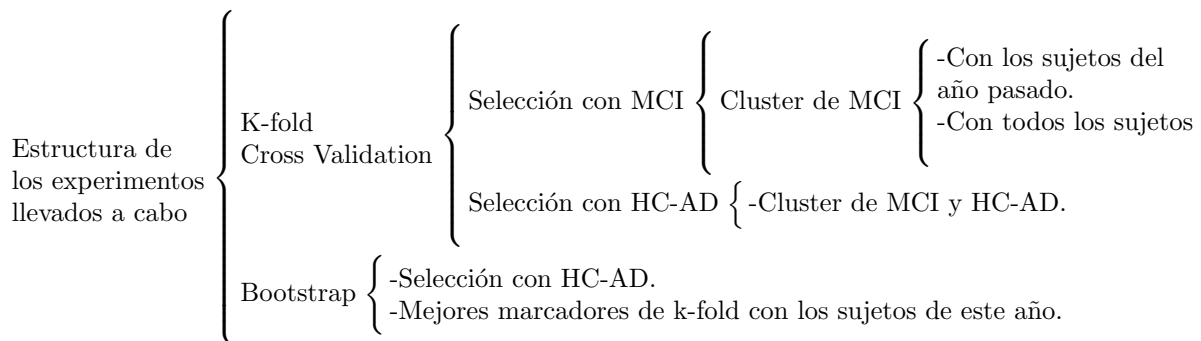
Figura 5.20: Diferentes combinaciones de las covariables con características de MRI junto con CAM en el Baseline

modelos adecuados debido a que los marcadores informan sobre la evolución de la enfermedad, en consecuencia se desea ser lo más preciso posible (uno de los motivos por los cuales se ha realizado el análisis masivo). La dificultad para extraer conclusiones aumenta en gran medida cuando la dimensión de las variables también lo hace. Como se ha comentado anteriormente la dimensión óptima se encontraba entre 3 y 5, en el siguiente experimento se dará una oportunidad a la dimensión 2 por el motivo previamente expuesto.

- Como es esperable los resultados mejoran al aumentar la información dada al modelo. Cuando se clasifica tan solo con la primera visita (Baseline) los resultados muestran una disminución en relación a cuando se tiene información de todas las visitas. Esto comprueba que el modelo creado es consistente con lo esperado.

Los siguientes pruebas realizadas buscan resultados reales por lo que no se ha rea-

lizado double dipping en ninguna de ellas. A continuación en el diagrama 5.4 se muestra la estructura de estos experimentos.



Nota: todas las gráficas mostradas a partir de ahora tendrán como eje y el accuracy del test. En ningún momento se mostrará el accuracy del training.

5.5. Doble K-fold Cross Validation

Siguiendo con la metodología de doble k-fold CV mostrada en el capítulo previo se han realizado los siguientes experimentos. En todos ellos, se ha variado el grupo de clínico para la selección y creación del modelo, es decir, los sujetos utilizados en el loop interior del doble k-fold CV. Mientras el modelo del loop exterior siempre es testeado con sujetos sMCI o pMCI.

5.5.1. Selección con MCI

La selección de marcadores se ha realizado con sujetos MCI.

Con clusters de MCI

Se conoce que los resultados dados por el doble k-fold creado son veraces al seleccionar las características con sujetos completamente independientes de los usados para comprobar el modelo. Aun así, con el objetivo de poder comprobar de una manera tremendamente purista los marcadores propuestos se van a utilizar una segunda estrategia. La cual comprobará los mejores marcadores mediante la metodología de bootstrap.

Datos de los sujetos del año pasado Se han utilizado solo una porción de los sujetos para así poder comprobar con la parte restante los marcadores propuestos.

La Figura 5.21 muestra el incremento que produce en el accuracy del modelo provenir al mismo de número mayor de visitas. A pesar de que se ha comentado en capítulos anteriores, es necesario subrayar que los modelos son creados con todas las visitas, el número de visitas expuesto en el eje x representa el número de visitas dadas para calcular los residuos y posteriormente clasificar, es decir, hace referencia a los sujetos del test.

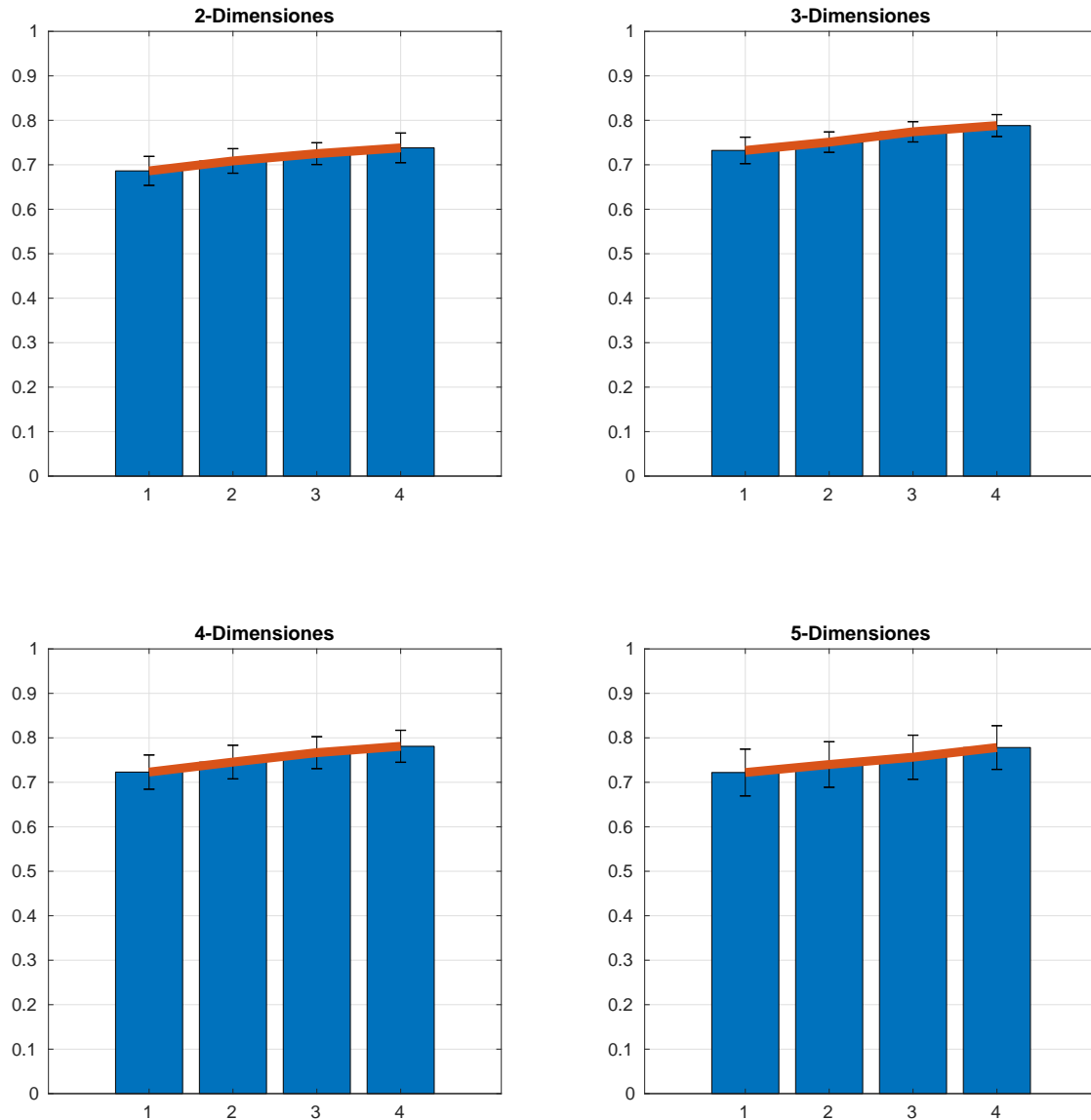


Figura 5.21: Comparativa de dimensiones en doble k-fold Cross Validation seleccionando con MCI. Los marcadores incluyen datos de MRI basados en atlas, CAM y clusters de MCI. En todas las subfiguras el eje x representa el número de marcadores mientras que el eje y representa la precisión en el test.

Como se puede observar en la Figura 5.22 existe un pequeño pico de accuracy en la dimensión 3. Se deberá contrastar si existe este fenómeno en los ensayos posteriores.

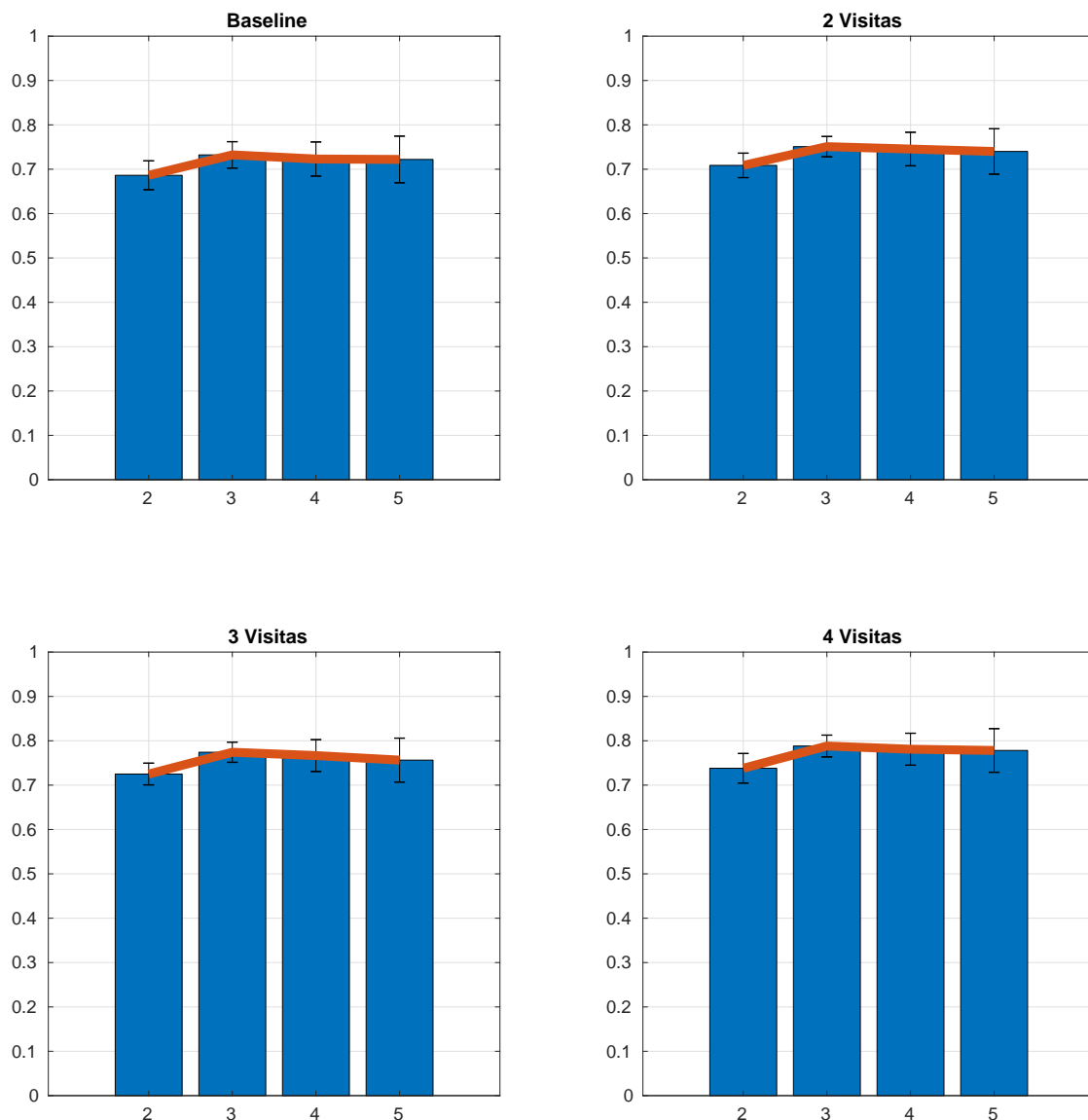


Figura 5.22: Comparativa de visitas en doble k-fold Cross Validation seleccionando con MCI. Los marcadores incluyen datos de MRI basados en atlas, CAM y clusters de MCI. En todas las subfiguras el eje x representa el número de marcadores mientras que el eje y representa la precisión en el test

La Tabla 5.5 muestra los marcadores que han sido seleccionados con mayor frecuencia para cada dimensión. Se debe destacar que se ha comprobado que estos los marcadores con mayor frecuencia (mayor veces seleccionados en el test, loop interior) son a su vez los que tiene mejores resultados en el test (bucle exterior). Por lo que el diseño y la implementación del doble k-fold una muy buena aproximación para encontrar mejores marcadores.

Datos de todos los sujetos Se continuo con los experimentos mediante la metodología de doble k-fold CV utilizando todos los sujetos disponibles, en la Tabla 5.6 se muestran las combinaciones con mayor frecuencia.

Tabla 5.5: Tabla con los mejores marcadores seleccionado con MCI, utilizando solo una porción del grupo de MCI. Los resultados se muestran con un intervalo de confianza de 95 %

SEN (%)	SPE (%)	ACC (%)	AUC	Marcadores
80.4 (79.4,81.5)	67.5 (66.3,68.7)	73.1 (72.4,73.9)	0.804 (0.797,0.811)	<i>ADAS13 + FAQ</i>
85.4 (84.5,86.2)	69.3 (68.2,70.3)	76.4 (75.8,77.1)	0.827 (0.820,0.834)	<i>ADAS13 + FAQ + rh_MCI.cl.1</i>
81.3 (79.5,83.2)	69.5 (67.4,71.6)	74.5 (73.1,75.9)	0.818 (0.805,0.832)	<i>PrecT_LH + ADAS13 + FAQ + lh_MCI.cl.2</i>
78.9 (77.3,80.4)	68.3 (66.5,70.1)	73.0 (71.9,74.1)	0.811 (0.799,0.823)	<i>PrecT_LH + ADAS13 + FAQ + ADASQ4 + lh_MCI.cl.2</i>

Tabla 5.6: Tabla con los mejores marcadores seleccionado con MCI con todos sujetos disponibles. Los resultados se muestran con un intervalo de confianza de 95 %

SEN (%)	SPE (%)	ACC (%)	AUC	Marcadores
82.4 (80.8,84.0)	59.2 (56.9,61.5)	70.8 (69.4,72.1)	0.786 (0.771,0.801)	<i>ADAS13 + FAQ</i>
84.6 (82.8,86.4)	64.1 (61.9,66.3)	74.2 (72.9,75.6)	0.804 (0.791,0.818)	<i>ADAS13 + FAQ + lh_MCI.cl.14</i>
84.0 (82.2,85.7)	64.6 (62.3,66.8)	74.3 (72.8,75.7)	0.821 (0.806,0.836)	<i>ADAS13 + FAQ + rh_MCI.cl.8 + lh_MCI.cl.14</i>
86.1 (83.4,88.9)	61.8 (58.4,65.3)	74.2 (71.9,76.5)	0.809 (0.790,0.828)	<i>CCingT_LH + ADAS13 + FAQ + rh_MCI.cl.8 + lh_MCI.cl.14</i>

Como se puede observar en la Figura 5.23, las sospechas eran ciertas en relación con la dimensión del número de características, dos características no aportan suficiente información. Además se puede observar como aumenta radicalmente el acuraccy con el aumento del número de visitas.

Con clusters de MCI y HC-AD

A pesar de que esta vez se incluían los marcadores de los clusters formados entre los grupos HC y AD, los marcadores seleccionados son los mismos que los seleccionados solo con los clusters de MCI. Esto tiene sentido puesto que el algoritmo mRMR selecciona las características más discriminantes entre los grupos. Se debe destacar que entre las 5 posiciones con mayor frecuencia para cada dimensión no existe ningún cluster del grupo HC-AD. En consecuencia, las gráficas son muy similares a las mostradas anteriormente por lo que se considera que no aportan valor. Aunque aparentemente este experimento no haya añadido nada se ha mostrado con el fin de informar sobre esta aproximación al problema no es muy fructífera.

5.5.2. Selección con HC y AD

Inspirado en la novedosa estrategia utilizada por Moradi [33] se ha decidido realizar la selección de marcadores con sujetos HC y AD. Esta estrategia se basa en que los sujetos sMCI deben tener muchas similitudes con respecto a los sujetos HC mientras que los sujetos AD deben tener similitudes con los AD.

A pesar de que los resultados mostrados en la Tabla 5.7 no tienen un gran poder de clasificación permiten extraer proposiciones de gran valor, las cuales son mostradas en el capítulo de conclusiones.

5.6. Bootstap

Como se ha resaltado en el capítulo de métodos, la estrategia de bootstrap ofrece una tremenda facilidad para realizar double dipping. Por ello, la selección de marcadores se ha realizado con los siguientes datos:

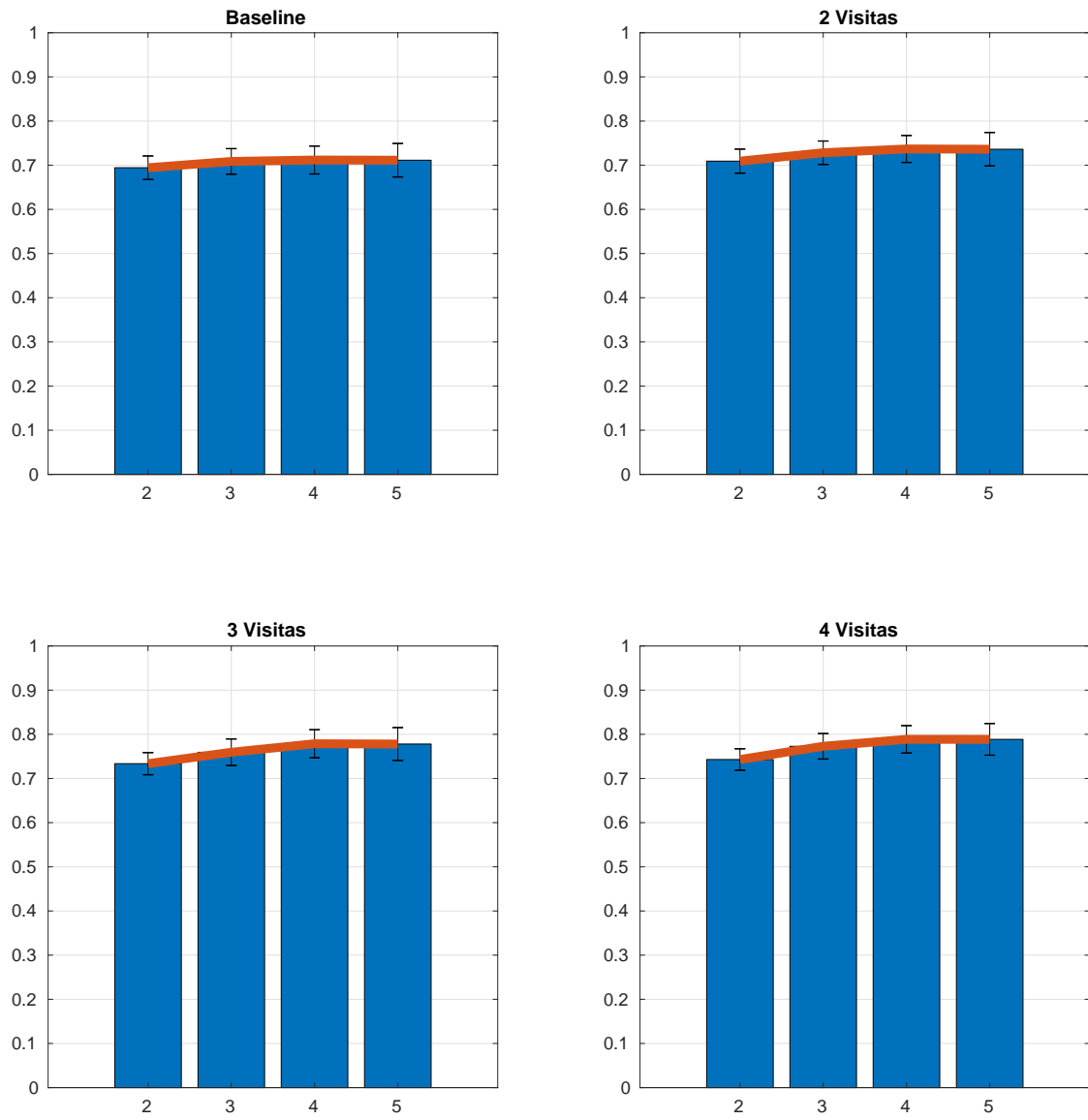


Figura 5.23: Comparativa de visitas en doble k-fold Cross Validation seleccionando con todos los sujetos. En todas las subfiguras el eje x representa el número de marcadores mientras que el eje y representa la precisión en el test

Tabla 5.7: Tabla de resultados de k-fold seleccionando con HC y AD. Los resultados se muestran con un intervalo de confianza de 95 %

SEN (%)	SPE (%)	ACC (%)	AUC	Marcadores
82.2 (81.3,83.0)	55.6 (54.7,56.6)	67.3 (66.7,68.0)	0.730 (0.723,0.738)	<i>RAVLT_learning + FAQ</i>
80.9 (80.0,81.8)	61.4 (60.4,62.3)	70.1 (69.5,70.7)	0.816 (0.810,0.822)	<i>FAQ + MMSE + ADASQ4</i>
80.0 (79.1,80.9)	61.7 (60.8,62.6)	69.9 (69.2,70.5)	0.810 (0.804,0.816)	<i>RAVLT_learning + FAQ + MMSE + ADASQ4</i>
81.2 (80.3,82.1)	62.7 (61.8,63.6)	70.9 (70.3,71.5)	0.816 (0.810,0.822)	<i>CCingT_LH + RAVLT_learning + FAQ + MMSE + ADASQ4</i>

5.6.1. Selección con HC y AD

Continuando con la estrategia de seleccionar con sujetos HC y AD, se ha implementado en bootstrap obteniendo los siguientes resultados.

Tabla 5.8: Tabla de resultados de bootstrap seleccionando con HC y AD. Los resultados se muestran con un intervalo de confianza de 95 %

SEN (%)	SPE (%)	ACC (%)	AUC	Marcadores
61.2 (59.8,62.6)	62.0 (60.6,63.4)	61.7 (60.4,63.0)	0.669 (0.655,0.683)	<i>FAQ</i>
66.2 (64.7,67.6)	62.0 (60.7,63.4)	63.9 (62.5,65.2)	0.692 (0.677,0.706)	<i>NCauV_RH + FAQ</i>
74.5 (72.9,76.1)	56.0 (54.7,57.3)	64.2 (62.8,65.6)	0.735 (0.719,0.750)	<i>NHV_RH + FAQ + ADASQ4</i>
72.4 (70.8,74.0)	55.6 (54.3,56.9)	63.1 (61.7,64.4)	0.723 (0.708,0.739)	<i>RAVLT_learning + FAQ + MMSE + ADASQ4</i>
74.0 (72.4,75.7)	58.0 (56.7,59.4)	65.2 (63.8,66.6)	0.736 (0.720,0.751)	<i>CCingT_LH + RAVLT_immediate + FAQ + MMSE + ADASQ4</i>

5.6.2. Mejores marcadores de k-fold

Se ha implementado la metodología de bootstrap con las mejores combinaciones obtenidas en el doble k-fold, utilizando como test la parte restante de la población. Esto se ha realizado a pesar de saber que los datos del doble k-fold son veraces.

La Tabla 5.9 muestra los resultados de los mejores marcadores marcadores de k-fold. Otros marcadores previamente seleccionados como adecuados por el k-fold han mostrado tener mayor potencial para esta población, la Tabla 5.10 muestra estos marcadores.

Tabla 5.9: Tabla de resultados de bootstrap utilizando las mejores características del k-fold anterior, por ello se utiliza la otra fracción de los sujetos. Los resultados se muestran con un intervalo de confianza de 95 %

SEN (%)	SPE (%)	ACC (%)	AUC	Marcadores
79.9 (79.0,80.9)	59.0 (57.8,60.2)	70.1 (69.4,70.8)	0.777 (0.769,0.785)	<i>FAQ + ADASQ4</i>
85.0 (84.2,85.9)	54.3 (53.1,55.5)	70.5 (69.7,71.2)	0.807 (0.799,0.814)	<i>ADAS13 + FAQ + rh_MCI_cl.1</i>
83.9 (83.0,84.8)	54.5 (53.3,55.7)	70.0 (69.2,70.7)	0.788 (0.780,0.796)	<i>PrecT_LH + ADAS13 + FAQ + lh_MCI_cl.2</i>
88.2 (87.4,88.9)	55.1 (53.9,56.3)	72.4 (71.7,73.1)	0.815 (0.808,0.823)	<i>PrecT_LH + ADAS13 + FAQ + MMSE + lh_MCI_cl.2</i>

5.7. DEMO del Software

La estructura de la del software utilizado se ha dividido en dos grandes demos. Lógicamente, ambas demos están estrechamente ligadas con el flow-chart mostrando

Tabla 5.10: Tabla de resultados reseñables de bootstrap utilizando características del k-fold anterior, por ello se utiliza la otra fracción de los sujetos. Los resultados se muestran con un intervalo de confianza de 95 %

SEN (%)	SPE (%)	ACC (%)	AUC	Marcadores
87.3 (86.5,88.1)	63.5 (62.3,64.7)	75.9 (75.2,76.6)	0.843 (0.836,0.850)	<i>ADAS13 + FAQ + rh_MCI.cl.4 + lh_MCI.cl.14</i>
91.0 (90.3,91.7)	61.2 (60.1,62.4)	76.8 (76.1,77.5)	0.839 (0.832,0.846)	<i>IstCingT.LH + MTT.LH + FAQ + MMSE + ADASQ4</i>
80.9 (80.0,81.8)	62.7 (61.5,63.8)	72.3 (71.6,73.0)	0.803 (0.795,0.810)	<i>IstCingT.LH + FAQ + ADASQ4 + rh_MCI.cl.1 + lh_MCI.cl.9</i>

anteriormente.

El enlace para descargar esta demo es:

<https://bitbucket.org/JaimeSV/tfg-alzheimer-early-diagnosis/src/master/>

5.7.1. DEMO extract features

Esta demo, contiene el procesamiento de imágenes y la extracción de características corticales y subcorticales. Estos scripts han sido modificados para que tan solo sea necesario extraer el archivo descargado de ADNI en la carpeta con dicho nombre y ejecutar los scripts en orden, tal y como se ha descrito en el capítulo de metodología.

La estructura de la demo es la siguiente:

$$\begin{array}{l}
 \left. \begin{array}{l}
 - \text{ADNI} \\
 - \text{Base} \\
 - \text{Cross} \\
 - \text{Error} \\
 - \text{Long}
 \end{array} \right\} \\
 \text{DEMO_extract_features} \left\{ \begin{array}{l}
 \text{scripts} \left\{ \begin{array}{l}
 -FS_recon_all \\
 -FS_recon_all_base \\
 -FS_recon_all_long \\
 -getAsegAparcADNI \\
 \text{getAsegAparc} \left\{ \begin{array}{l}
 -parseXML \\
 -readXML \\
 -sortData
 \end{array} \right. \\
 \text{Sorted} \left\{ \begin{array}{l}
 -nii \\
 -xml
 \end{array} \right. \\
 -README.txt
 \end{array} \right.
 \end{array}
 \end{array}
 \end{array}$$

5.7.2. DEMO machine learning

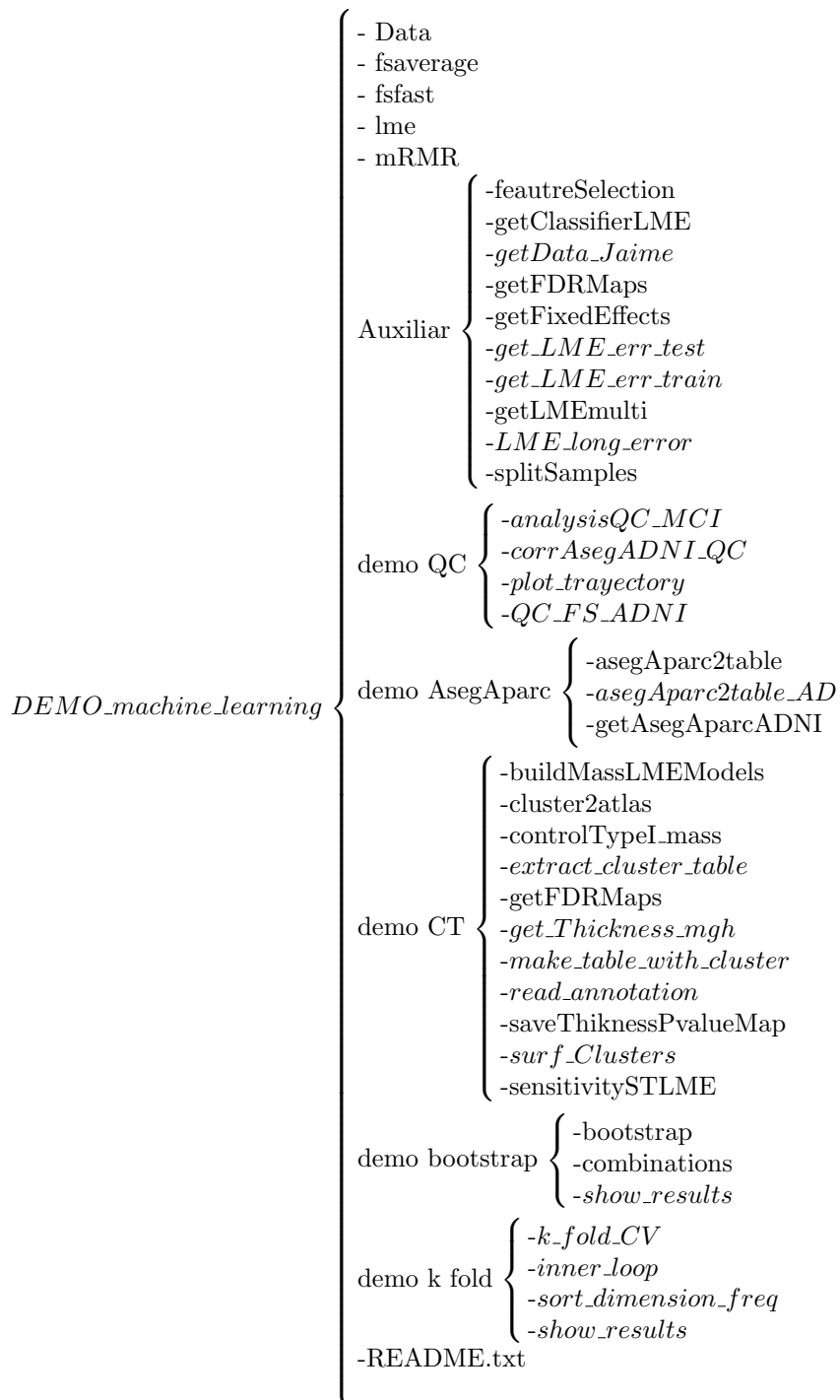
Esta demo contiene el control de calidad junto con la selección de características y la clasificación. Ambas se han implementado de manera automática mediante las estrategias expuestas en los algoritmos.

En esta demo se puede encontrar grupos diferenciados de carpetas

- Las carpetas demo contienen los scripts necesarios para poder llevar a cabo dicha actividad. El script principal, el que se debe ejecutar, tiene el mismo

nombre que la carpeta.

- Las carpetas que NO estan nombradas como demo son carpetas auxiliares. Estas carpetas suelen venir incluidas al descargar free-surfer pero se han incluido para mejorar la portabilidad de todo el sistema. Se recomienda leer el README de cada demo, puesto que indican que carpetas auxiliares deben ser incluidas para ejecutar cada una de las diversas demos creadas.



Capítulo 6

Discusiones y conclusiones

En primer lugar, se debe destacar el predominio de los marcadores de evaluaciones funcionales cognitivas complejas, siendo estos marcadores especialmente valiosos en bajas dimensiones. Esto refleja que las evaluaciones clínicas aportan gran cantidad de información por sí mismas, mientras que la atrofia en las regiones anatómicas se da de manera conjunta en diferentes regiones simultáneamente. Por lo que es necesario aumentar la dimensión para poder mostrar estas complejas relaciones.

El test **FAQ** es el marcador más recurrido, puesto que, como se hace referencia en el capítulo de Materiales, este test permite medir las capacidades del paciente para realizar las actividades de su vida cotidiana. La Figura 6.1 el modelo de la trayectoria de este marcador.

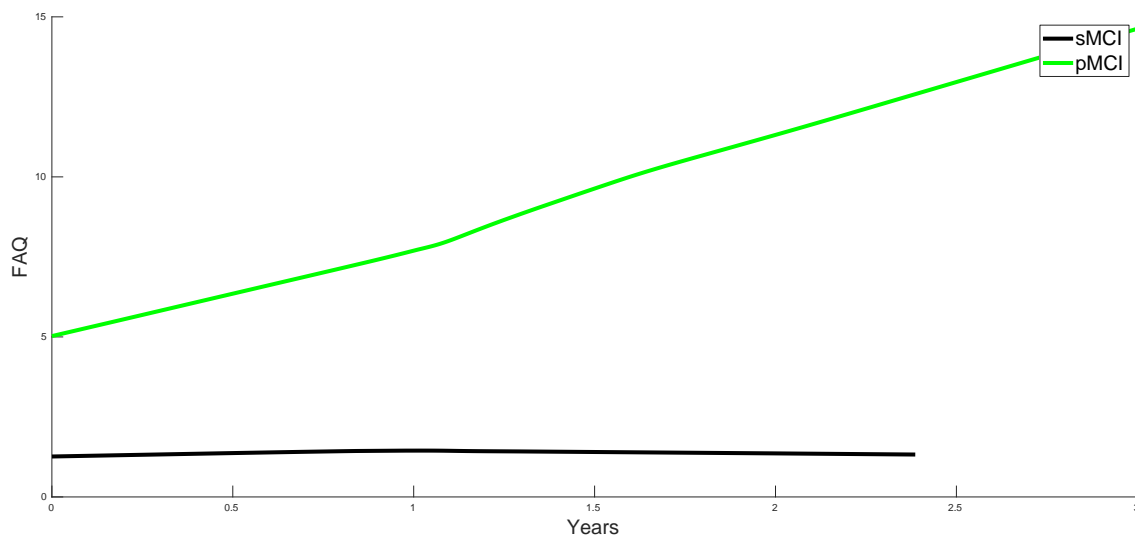


Figura 6.1: Trayectoria del marcador FAQ

El modelo de la trayectoria del grupo pMCI tiene una pendiente positiva, lo cual implica mayor incapacidad para realizar sus actividades de la vida diaria, véase el test FAQ en la Figura 3.4. Mientras que los sMCI se mantienen constantes.

La demencia se define como un deterioro de la capacidad mental, suficientemente grave como para interferir en la vida diaria, razón por la cual este test es el marcador más recurrido.

Otro marcador cognitivo utilizado frecuentemente es el ADAS. Esta evaluación esta especialmente diseñada para la enfermedad de Alzheimer y es ampliamente utilizada para monitorizar la evolución de la enfermedad.

En dimensiones mayores (3-5) solo las evaluaciones funcionales no dan suficiente información, por lo que es necesario la utilización de datos anatómicos. Esta inclusión provoca un incremento en el poder del clasificador como se puede observar en las Figuras 5.22 y 5.23. Entre estos marcadores provenientes del MRI estructural suelen ser escogidos los clusters en lugar de las regiones definidas mediante atlas. Esto se debe a que los clusters son zonas donde se da el mayor deterioro. Los clusters seleccionados con mayor frecuencia vienen dado en la siguiente Tabla 6.1 y se muestran en la Figura 6.2.

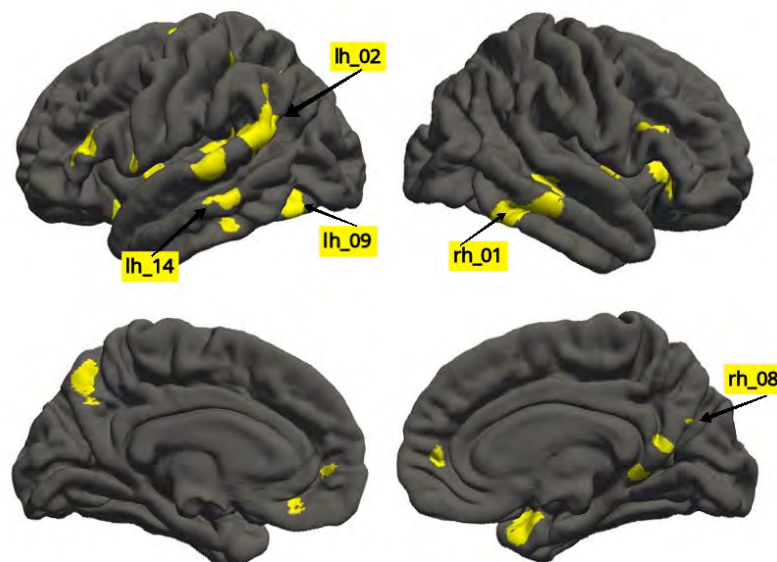


Figura 6.2: Figura donde se muestran los cluster más significativos

Tabla 6.1: Tabla donde se muestra la relación entre los cluster más significativos y las diferentes estructuras del atlas de Desikan-Killiany

Cluster	Estructuras
lh_MCI.cl.2	supramarginal, bankssts, inferiorparietal, superiortemporal
lh_MCI.cl.9	lateraloccipital, fusiform
lh_MCI.cl.14	middletemporal
rh_MCI.cl.1	middletemporal, inferiortemporal
rh_MCI.cl.8	precuneus

Como muestra la tabla, casi todos los cluster están formados por la unión de varias regiones, esto refleja que debido a dificultad intrínseca de esta enfermedad

era necesario realizar el complejo análisis masivo llevado a cabo.

Tras la multitud de experimentos realizados es necesario destacar la influencia de las siguientes estructuras:

- Lóbulo Parietal Inferior
- Lóbulo Temporal Superior
- Lóbulo Temporal Medial
- Precúneo
- Cótex del Cíngulo Anterior

En segundo lugar, en todos los casos del experimento “Primera aproximación”, los mejores resultados se encuentran cuando la única covariable al modelo LME era la edad. Cabe destacar la mala actuación obtenida al introducir al modelo el gen APOE-4, posiblemente este detrimento venga causado por la binariedad del dato lo cual provoca una gran influencia en el modelado de las trayectorias. Por otro lado, los resultados son adecuados al utilizar como covariables la edad o la educación, lo cual refleja la influencia de estos factores de riesgo en la enfermedad. Tras la duradera investigación necesaria sobre esta enfermedad neurodegenerativa llevada a cabo para poder realizar este proyecto, me veo en la obligación de informar de la tremenda influencia que puede tener no los años de educación, sino el ejercicio mental. Este factor de riesgo no va a ser determinante en padecer o no la enfermedad, va a tener gran influencia en la evolución y calidad de la vida de dicha enfermedad.

A pesar de que la literatura haya demostrado la importancia del género, la educación o el APOE-4 como factores de riesgo. La **edad** es un factor de riesgo tan relevante para esta enfermedad que como se muestra en el experimento, la introducción de cualquier otra covariable solo empeora el resultado.

En tercer lugar, se debe resaltar la siguiente observación, ¿cómo es posible que con los mismos marcadores den resultados dispares en el doble k-fold CV, solo porque en un experimento se haya utilizado una subgrupo de los sujetos y en el otro la totalidad de los mismos? Esta reflexión hace referencia a las Tablas 5.5 y 5.6; donde la misma combinación de marcadores, el ADAS13 y el FAQ, según la población o aumenta la sensibilidad disminuyendo la especificidad o al revés, siendo menor la precisión del grupo completo. Esto junto con la variedad de marcadores presentes en la Tabla 5.10 denota que este subgrupo es mucho más heterogéneo que el otro y por lo tanto cuando se unen ambos los resultados empeoran. Disponer de un grupo heterogéneo es muy beneficioso puesto que permite evaluar la actuación del clasificador en situaciones adversas.

En cuarto lugar, debido a su importancia se quiere resaltar de nuevo el gran incremento dado al aumentar el número de visitas, puesto que al disponer de mayor información el algoritmo es capaz de predecir con mayor precisión la trayectoria que va a seguir el sujeto. Es necesario subrayar la complejidad del algoritmo creado puesto que solo con datos en el baseline se realiza una predicción basada en la trayectoria. Es decir, solo con un punto de la misma.

Entre la multitud de combinaciones de marcadores obtenidos se ha determinado como óptimo la combinación de **ADAS13, FAQ y rh MCI cl 1**. Se ha decidido esta combinación de marcadores por motivos como: muy alta precisión, del **76.4 %** con un intervalo de confianza del 95 % entre los valores 75.8 % y 77.1 %, baja dimensión, relación adecuada entre sensibilidad y especificidad; y resultados adecuados para un grupo muy heterogéneo como es el subgrupo con el que se comprobaron los resultados de nuevo en el bootstrap.

Tabla 6.2: Tabla de resultados con los marcadores ADAS13, FAQ y rh MCI cl 1 y su evolución con el número de visitas. Los resultados se muestran con un intervalo de confianza de 95 %

Número de visitas	SEN (%)	SPE (%)	ACC (%)	AUC
1	85.4 (84.5,86.2)	69.3 (68.2,70.3)	76.4 (75.8,77.1)	0.827 (0.820,0.834)
2	78.6 (77.6,79.6)	74.1 (73.2,75.1)	76.1 (75.4,76.8)	0.843 (0.836,0.849)
3	76.5 (75.4,77.5)	79.1 (78.2,79.9)	77.8 (77.1,78.5)	0.871 (0.865,0.877)
4	73.4 (72.3,74.4)	81.5 (80.6,82.3)	77.8 (77.1,78.4)	0.883 (0.878,0.889)

Es necesario destacar que existen otras combinaciones con mejores actuaciones con 4 visitas, llegan a obtener precisiones de 82.1 %.

Por último, se cabe precisar que cuando la selección en el doble k-fold CV se realizaba con sujetos HC y AD se notaba un claro predominio de las evaluaciones clínicas de las funciones cognitivas, tanto simples como complejas. El principal motivo de esta selección puede ser que a pesar de existir cambios anatómicos, las mayores diferencias entre sujetos HC y AD se dan en las evaluaciones clínicas.

Aun así este proyecto de investigación ha demostrado que la utilización exclusiva de evaluaciones clínicas no es la solución más óptima, puesto que la información dada suele ser redundante. Por lo que es estrictamente necesario la inclusión de marcadores basados en la anatomía para así añadir más información sobre el desarrollo de la enfermedad.

6.1. Contraste con otros estudios

Utilizando *Sequential Forward Feature Selection (SFS)*, los marcadores más utilizados para la clasificación entre HC, AD y MCI son el volumen del hipocampo, el volumen del ventrículo, el espesor del lóbulo parietal, el espesor del lóbulo occipital y el espesor del cingulate [46]. Como se ha podido observar los datos obtenidos en este proyecto están en línea con los de la literatura.

Otro artículo muy reciente (Septiembre 2017) realizado por el Departamento de neurología y centro de Alzheimer de Amsterdam [50] demuestra la importancia del lóbulo temporal. Este resultado es muy similar al mostrado en este experimento puesto que se ha incluido en la mejor combinación de marcadores al cluster *rh_MCI_cl_1*, formado por el lóbulo temporal medial y el lóbulo temporal superior. En contraste,

en su estudio ajustaron el modelo LME utilizando como covariables la edad, el sexo, la educación. Los experimentos llevados a cabo en este estudio demuestran que la covariable únicamente de la edad es la más poderosa y tantas covariable empeoran el resultado.

La siguiente Tabla 6.3 resume los resultados de los artículo más novedosos en la clasificación de sMCI y pMCI:

Tabla 6.3: Tabla comparativa entre los estudios más reseñables en la literatura y este proyecto

	Follow-up	Número de sujetos sMCI/pMCI	Información	AUC	ACC (%)	SEN (%)	SPE (%)	Número de características
Gavidia [18]	36		MRI		70.30	64.26	75.14	
Beheshti [2]	36	65/71	MRI	0.751	75	76.92	73.23	
Eskilsen [15]	36	227/161	MRI	0.763	71.9	69.6	73.6	6
Platero	36	101/81	MRI	0.796	76.0	77.7	76.0	3
Ferreira [16]	36	160/122	CF	0.870	81.56	85	79	7
Korolev [25]	36	120/139	MRI,CF	0.87	79.9	83.4	76.4	10
Moradi [33]	36	164/100	MRI,CF	0.9	82	87	74	
Gavidia [18]	36		MRI,CF		76.72	70.77	81.62	
Platero	36		MRI,CF	0.885	81.2	75.4	85.8	3
Proyecto Actual	36		MRI,CF	0.827	76.4	85.5	67.5	3

A pesar de que la clasificación en del grupo MCI, en general, es difícil, debido a ser un grupo muy heterogéneo entre HC y AD [46], los resultados obtenidos son muy satisfactorios puesto que no discrepan en gran medida de los obtenido por los mejores investigadores de la actualidad en este campo.

Aunando todas las conclusiones previamente mencionadas, se destacar que en este proyecto se ha propuesto una metodología robusta capaz de modelizar, seleccionar y analizar aquellos marcadores cuyas trayectorias longitudinales son más significativas. Permitiendo así, predecir el futuro diagnóstico dentro de tres años del grupo de población con mayor riesgo de padecer la enfermedad, los sujetos con deterioro cognitivo leve.

6.2. Desarrollos futuros

Se ha demostrado que las imágenes estructurales tienen mayor dificultad para diferenciar MCI tempranos que otros biomarcadores como CSF o PET [21]. En consecuencia, la combinación de estas imágenes funcionales con imágenes estructurales puede ser de gran utilidad para la diferenciación en concreto entre HC y MCI.

Como se ha expresado anteriormente, existen 4 dominios diferentes dentro de los sujetos MCI. Por ello, sería muy aconsejable diferenciar entre estos grupos y estudiarlos por separado. Pudiendo obtener marcadores específicos para cada dominio afectado (como este proyecto) y marcadores para el grupo completo de MCI. Se podrían extraer conclusiones muy reseñables al contrastar los marcadores obtenidos para cada uno de estos grupos.

Una mejora potencialmente satisfactoria debido a la metodología utilizada en este proyecto sería modificar el sector de características actual por uno longitudinal.

Es decir, introducir tanto el ratio de atrofia anual como el valor en el baseline de los marcadores a la entrada del algoritmo encargado de selección de características, en este caso mRMR. Otra posible mejora se podría realizar si los datos se extraen de diferentes atlas asemejando al trabajo realizado por Liu [28]. Otro adelanto sería introducir un método basados en la posición de los vértices de algunas estructuras concretas, pudiendo así obtener marcadores utilizados por la literatura tales como el espesor o la profundidad máxima del surco de regiones como el hipocampo.

Por otro lado, se podría modificar el paradigma mostrado solo realizando una pequeña variación el flow char seguido. Cambiando el actual clasificador utilizado LDA por uno de tipo SVM, preferiblemente con un kernel gaussiano, se obtendrían otras dimensiones óptimas, posiblemente más altas. Por lo tanto, este clasificar reflejaría mejor las complejas relaciones entre estructuras cerebrales.

Por el contrario, se desaconseja utilizar clasificadores basados en redes neuronales puesto que a pesar de que los resultados obtenidos con estos clasificares pueden ser muy adecuados, la alta complejidad de relaciones enmascaran cual es el grupo ideal de marcadores. Ocultando por ello, información muy relevante sobre la enfermedad y el desarrollo de la misma.

Apéndice A

Conceptos teóricos básicos

A.1. Imágenes de resonancia magnética

Las imágenes de resonancia magnética (MRI) son imágenes anatómicas tridimensionales detalladas. Su finalidad es la detección de enfermedades junto con la monitorización de tratamientos. Esta tecnología se basa en la detección del cambio en la dirección del eje de rotación de los protones que se encuentran en los átomos de hidrógeno.

A.1.1. Glosario básico

En primer lugar, las MRI son generadas mediante cortes (slices) bidimensionales y se representan mediante una escala de grises. Por ello, se deben definir varios conceptos usados en la descripción imágenes de resonancia magnética:

- Hiperintenso. Tonos blancos, brillantes.
- Isointenso. Tonos similares a los tejidos circundantes.
- Hipointenso. Tonos oscuros, negros.

A continuación, se va a definir las diferentes estructuras que existen en el cerebro humano.

- Materia blanca. Compuesto por fibras nerviosas o axones de las neuronas. Se encuentran en las estructuras centrales del cerebro. Su función es comunicar a la materia gris con las otras partes del cuerpo.
- Materia gris. Compuesto por cuerpos neuronales y neuropil¹. Su función implica el control muscular, la percepción sensorial, la memoria, las emociones y la toma de decisiones.
- Ventrículos. Conjunto de cuatro cuerpos callosos interconectados entre si. Su función es la producción del líquido cefalorraquídeo. Este líquido es el medio por el cual el cerebro elimina residuos de su actividad.

Se puede observar la estructura del cerebro en la Figura A.1 .

¹) Red de finas conexiones multineuronales entre axones, sus colaterales y las espinas dendríticas.

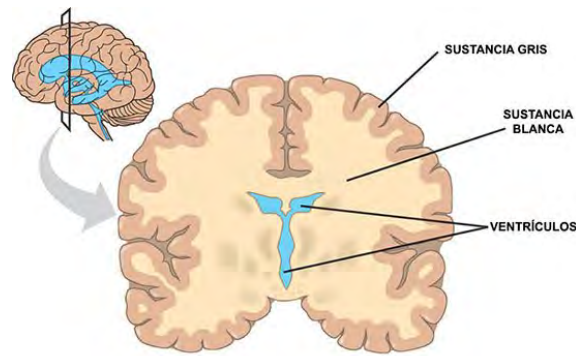


Figura A.1: Estructuras del cerebro humano

A.1.2. Funcionamiento

Como se ha descrito anteriormente, las técnicas de MRI se basan en los protones de nuestro cuerpo, los cuales giran en multitud de direcciones. Dentro del escáner se encuentran unas bobinas que crean un campo magnético de 0,1 a 7 Teslas, dependiendo de la precisión del escáner. Este campo magnético provoca que todos estos protones se alineen (en paralelo o antiparalelo) con las líneas de dicho campo.

A continuación, se emiten un pulso de radiofrecuencia. Es necesario destacar que solo los protones que tienen la misma frecuencia de precesión² son afectados por el pulso. Este aporte extra de energía provoca una rotación de los protones. Este fenómeno es llamado **resonancia**.

Cuando estos pulsos finalizan los protones vuelven a su situación de equilibrio, es decir, alineados con el campo magnético. Este movimiento provoca un desprendimiento de energía que es captada por los sensores del escáner. Este punto de equilibrio también es llamado de relajación.

En resumen, el tiempo que tardan los protones para realinearse con el campo magnético, así como la cantidad de energía liberada, cambian dependiendo del entorno y la naturaleza química de las moléculas (propiedades magnéticas). Mientras más rápido se realineen los protones, más brillante será la imagen. En algunas ocasiones es necesario utilizar un elemento contraste (normalmente el gadolinio) para aumentar la velocidad a la cual los protones se realinean con el campo magnético.

Pulso de radiofrecuencia En cuanto a las formas de onda de los pulsos, las más comunes son la Gaussiana y la seno cardinar o también llamada sinc ($\text{sinc}(x) = \sin(x)/x$). Cabe destacar que la transformada de Fourier del pulso Gaussiano conduce a una forma gaussiana, mientras que la transformada de pulso sinc lleva a una forma casi rectangular (siendo un filtro paso banda ideal), lo que permite una mejor resolución.

² Movimiento circular rotatorio, con una frecuencia específica para cada elemento químico. Viene determinada por la ecuación de Larmor

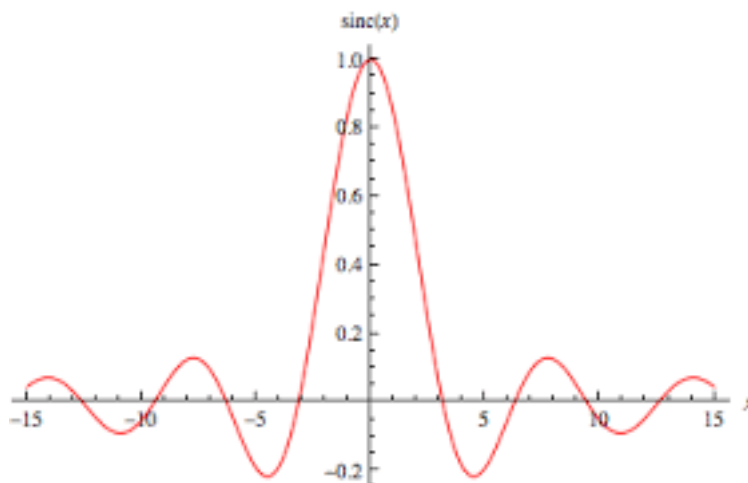


Figura A.2: Forma de onda del seno cardinal

A.1.3. Tipos de imágenes

Según el tiempo de exposición Según los ejes estudiados para llegar a al punto de equilibrio podemos destacar dos tipos de tiempos y en consecuencia dos tipos de imágenes:

- **Longitudinal o T1.** Tiempo necesario para recuperar el 63% de la magnetización original, en otras palabras, para alinearse con el campo magnético externo. Los valores son fuertemente dependientes del campo magnético ya que estos imponen la frecuencia de trabajo. Los valores aumentan con el valor del campo magnético. Este tipo de imágenes se corresponden adecuadamente con la **anatomía**. El T1 es más largo en los líquidos que en los sólidos y es más corto en los tejidos grasos. Esto provoca que los líquidos como el agua, se presente en un T1 de color negro. Para potenciar el T1, tanto el tiempo de repetición de la secuencia, como el tiempo eco (recoger la señal que nos envían los protones) tienen que ser muy cortos.
- **Transversal o T2.** Tiempo necesario para que los protones se desfasen un 37% de lo que estaban. No es muy sensible a los cambios del campo magnético. Para potenciar el T2, es necesario que tanto el tiempo de secuencia como el tiempo eco sea largó. La característica principal es ser muy sensibles a los **cambios patológicos**.

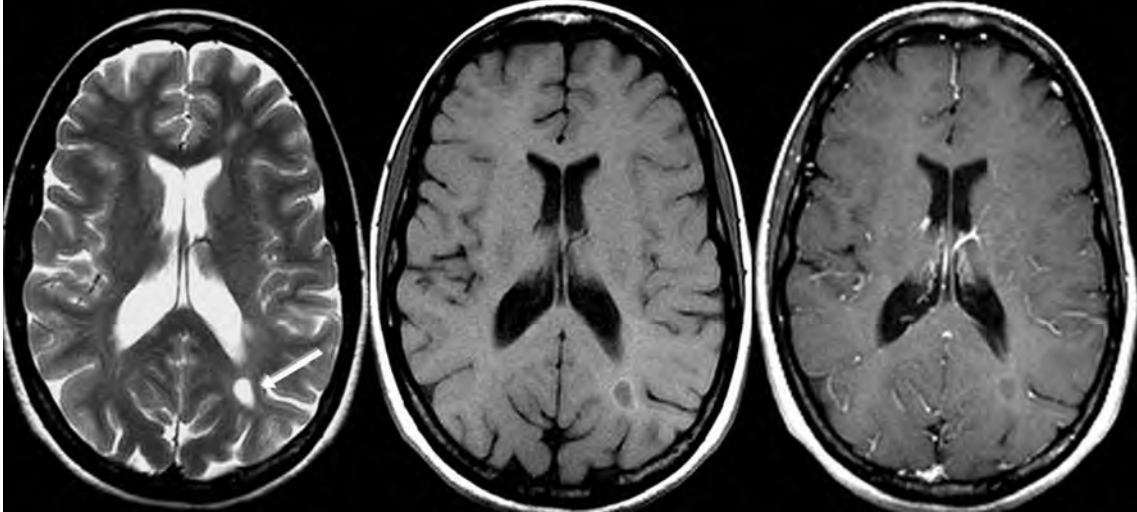
Tabla A.1: Tabla comparativa de los tiempos de repetición y eco

	T1	T2
TR (ms)	450-650	2.800-4.000
TE (ms)	<25	80-120

En la Figura A.3 se puede observar claramente la diferencias. En las imagen T1, la sustancia blanca es hipertensa mientras que el liquido cefalorraquídeo (LCR) hipointenso. Por otro lado, mediante la secuencia T2, la sustancia blanca es hipointesa

mientras el líquido cefalorraquídeo hiperintenso.

Figura A.3: Resonancia magnética cerebral. Secuencias ponderadas en T2 (izquierda), T1 (centro) y T1 con contraste (derecha) en el plano transversal



A continuación, en la tabla comparativa A.2 se muestra a modo de resumen la diferente gamas de colores según el elemento y la ponderación utilizada.

Tabla A.2: Tabla comparativa entre las diferentes apariencias de diversos elementos según la ponderación utilizada

Elemento	Secuencia T1	Secuencia T2
Agua	hiperintenso	hipointenso
Grasa	hipointenso	hiperintenso
Aire	hipointenso	hipointenso
Sustancia gris	gris-hipointenso	gris-hiperintenso
Sustancia blanca	gris-hiperintenso	gris-hipointenso
Tejido colágeno	hipointenso	hipointenso
Hueso cortical	hipointenso	hipointenso
Calcificaciones	hipointenso	hipointenso
Sangre en periodo remoto (hemosiderina)	hipointenso	hipointenso

Según el tipo de corte En la Figura A.1.3 se puede observar el cerebro de un paciente desde los diferentes puntos de vista normalizados.

A- Axial o transversal. Vista desde el inferior

B- Sagital. Vista desde la izquierda

C- Coronal. Vista desde el anterior

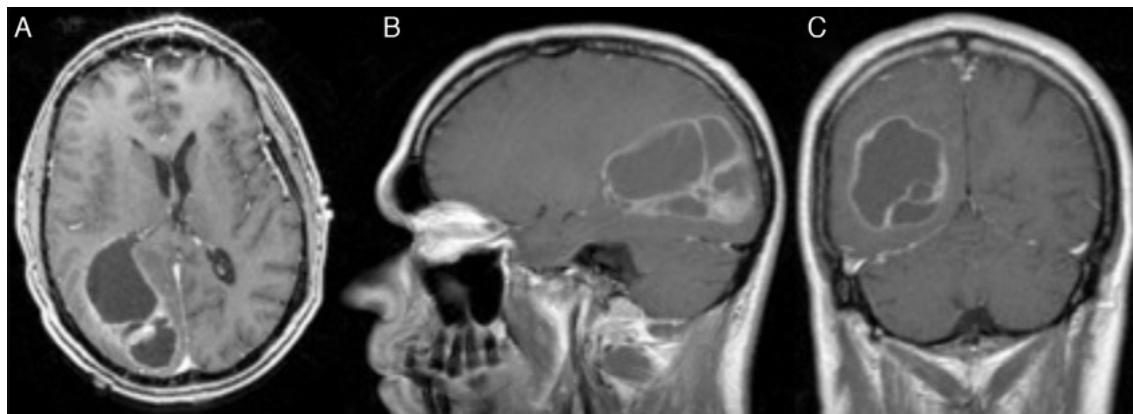


Figura A.4: Tipos de cortes

A.1.4. Ventajas

La principal ventaja de las MRI es que se trata de una técnica de exploración no invasiva debido a que no implica exposición a la radiación ionizante. Adicionalmente, las MRI del cerebro y otras estructuras craneales son más claras y detalladas que las que se obtienen mediante otros métodos de diagnóstico por imágenes.

A.2. Registro de imágenes

El registro de imágenes consiste en eliminar las diferencias geométricas entre dos imágenes. Esta disparidad puede ser causada porque las imágenes hayan sido tomadas bajo diversas posiciones y/o condiciones. El objetivo del registro de imágenes es que cada pixel de ambas imágenes correspondan al mismo punto del objeto real. En consecuencia, se podrá comparar ambas imágenes con facilidad. Cabe destacar que la resolución es vital para un buen registro de imágenes.

A.2.1. Transformaciones geométricas

Las diferencias geométricas previamente comentadas son eliminadas mediante transformaciones geométricas. Las transformaciones geométricas modifican la relación espacial entre los píxeles. Se basa en dos principios, transformación espacial y interpolación de los niveles de grises.

Se debe subrayar que una imagen digital es una matriz, por ello las operaciones matriciales son fundamentales en el procesamiento de las imágenes.

A.2.2. Transformaciones espaciales

Existen dos tipos, mapeo directo e inverso. El primero es más costoso computacionalmente, por lo que se suele realizar el mapeo inverso. En este procedimiento se interpola el valor de los píxel de la imagen de entrada más cercanos para así determinar el valor del píxel de salida.

A.2.3. Transformaciones rígidas

Son un tipo de transformación geométrica donde se mantienen todas las distancias entre los diferentes puntos de cada imagen. En consecuencia, se mantiene la rectitud de las líneas, la planitud de las superficies y los ángulos formados entre las rectas. Entre las transformaciones rígidas se destaca la traslación y la rotación.

$$x' = Rx + t$$

Traslación La traslación viene determinada por el vector t . Sus componentes vienen dados respecto al sistema de coordenadas original (x,y,z) .

Rotación La rotación viene determinada por la *matriz ortogonal* R ³. Las matrices de rotación incluyen rotaciones apropiadas, que producen transformaciones rígidas, e inapropiadas, que rotan y reflejan el objeto. Añadiendo la restricción de $\det(R) = 1$ eliminamos las rotaciones inapropiadas. Existen multitud de maneras de rotar: ángulos de Euler, cuaternios, matrices ortogonales etc.

Ejemplo de rotación y translación rígida en 2D

Si nos encontramos en el plano x,y y deseamos realizar es una rotación de un ángulo θ y una traslación al punto (t_1, t_2) , esta transformación viene dada por la siguiente matriz.⁴

$$M = \begin{pmatrix} \cos\theta & -\sin\theta & t_1 \\ \sin\theta & \cos\theta & t_2 \\ 0 & 0 & 1 \end{pmatrix}$$

A.2.4. Transformaciones no rígidas

Transformaciones afines

Vienen dadas por:

$$x' = Ax + t$$

donde no existen restricciones para los elementos de A . En las transformaciones afines, se conserva la rectitud de las líneas y la planitud de las superficies pero permite la variación de los que forman las líneas. Este tipo de transformación es utilizada para *corregir las inclinaciones* producidas en la adquisición de imágenes.

³Una matriz ortogonal implica que $R^{-1} = R^t$

⁴Se debe destacar que nos encontramos ante la misma matriz que representa el movimiento de rotación en la robótica

$$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Un caso particular de transformación afin es el escalado, donde la matriz A es una matriz diagonal y sus elementos representan el escalado en los diversos ejes. Si estos escalares son el mismo se esta ante un escalado isotrópico o transformación de similitud.

Ejemplo de transformación afin en 2D

Las coordenadas de un punto (x',y') se expresada mediante un combinación lineal de las coordenadas del punto original (x,y) :

$$x' = a_{11}x + a_{12}y + t_1$$

$$y' = a_{21}x + a_{22}y + t_2$$

Este sistema se puede representar de forma matricial del siguiente manera:

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & t_1 \\ a_{21} & a_{22} & t_2 \\ 0 & 0 & 1 \end{pmatrix} * \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

Estos coeficientes de la matriz de transformación (M) tienen un sentido físico:

- a_{11} . Representa el escalado en el eje x
- a_{12} . Representa la inclinación o presión en dirección horizontal
- a_{21} . Representa la inclinación o presión en dirección vertical
- a_{22} . Representa el escalado en el eje y
- t_1 . Representa el desplazamiento horizontal
- t_2 . Representa el desplazamiento vertical

Transformaciones proyectivas

Se conserva la rectitud de las líneas pero se puede modificar el paralelismo de las mismas.

$$M = \begin{pmatrix} a_{11} & a_{12} & a_{13} & t_1 \\ a_{21} & a_{22} & a_{23} & t_2 \\ a_{31} & a_{32} & a_{33} & t_3 \\ p_1 & p_2 & p_3 & \alpha \end{pmatrix}$$

donde p es el vector que especifica el eje de proyección y α un escalar.

Se puede observar en la Figura A.5 que tan solo se conserva la rectitud de las líneas.

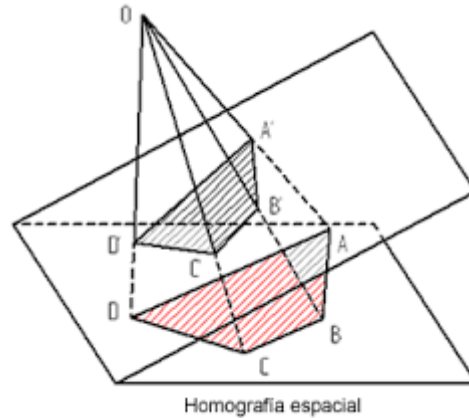


Figura A.5: Ejemplo de una transformación proyectiva

A.3. Contraste de hipótesis

El contrastes de hipótesis se basa en datos muestrales. Se comienza con la formulación de la **hipótesis nula H_0** la cual será mantenida a menos que la muestra aporte suficiente evidencia contraria (la H_0 es aceptada antes de llevar a cabo el test). Por ello la muestra debe contrastar la hipótesis. La hipótesis nula siempre forma un conjunto cerrado.

Si la hipótesis nula no es válida alguna debe de ser correcta. Esta **hipótesis alternativa H_1** debe ser opuesta a la nula, por lo que forma un conjunto abierto. Una vez definidas ambas hipótesis se debe tomar una decisión, rechazar o no rechazar H_0 . La decisión se basa en el valor de la “distancia” (valor de un estadístico de contraste) entre los valores muestrales que disponemos y aquellos que tienen una alta probabilidad si se cumple la hipótesis nula. Se debe destacar que existe una posibilidad de llegar a una conclusión equivocada al solo disponer de una muestra aleatoria, por lo que no podemos afirmar con certeza que la hipótesis nula sea correcta o no. En consecuencia se puede cometer dos tipos de errores

- Error de tipo I. Rechaza una hipótesis nula correcta. La probabilidad de cometer un error de tipo I es α y se denomina nivel de significación.

$$\alpha = P(\text{Rechazar } H_0 \mid H_0 \text{ correcta})$$

- Error de tipo II. No rechazar la hipótesis nula incorrecta.

$$\beta = P(\text{Norechazar } H_0 \mid H_1 \text{ incorrecta})$$

- Potencia. Probabilidad de rechazar una hipótesis nula cuando es incorrecta.

$$\text{Potencia} = P(\text{Rechazar } H_0 \mid H_1 \text{ incorrecta})$$

La siguiente tabla muestra de una manera más visual los tipos de errores.

Como se podía intuir la probabilidad del error de tipo I es inversamente proporcional a la probabilidad de error de tipo II.

En la Figura A.6 auna todos los conceptos vistos y ayuda a entender de una manera muy visual los valores α y β nombrados anteriormente.

Tabla A.3: Tabla comparativa donde se muestran los tipos de errores

Decisión Situación actual	Ho correcta	Ho incorrecta
No rechazar Ho	Sin error ($1 - \alpha$)	Error de tipo II β
Rechazar Ho	Error de tipo I α	Sin error ($(1 - \alpha) \equiv Potencia$)

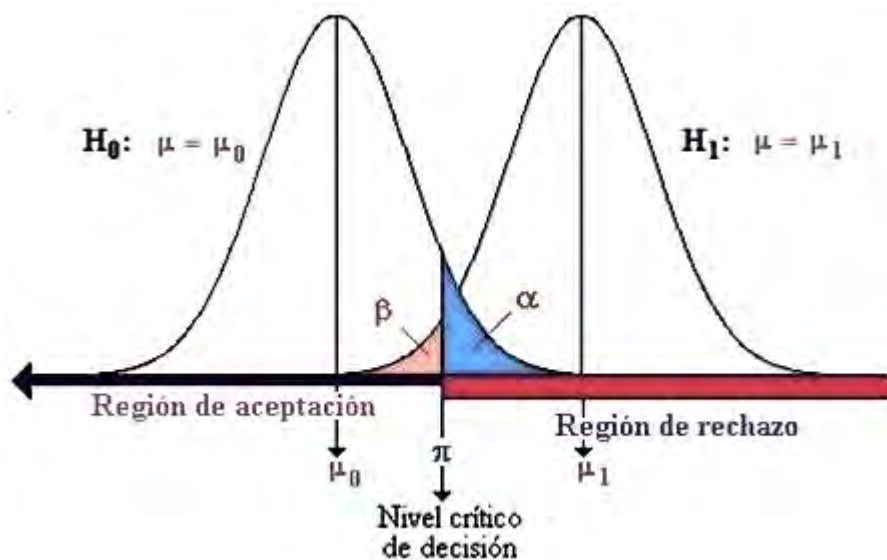


Figura A.6: Contraste entre dos distribuciones normales

Estadístico de contraste Informa si es probable o improbable que se observen los datos muestrales suponiendo la hipótesis nula como cierta. Se debe rechazar la H_0 si el valor estadístico de contraste está dentro de la región de rechazo (la cual viene definida por el nivel de significación).

P-valor Es el menor valor de α para que se pueda rechazar la H_0 .

- Si el p-valor es $< \alpha$, rechazamos H_0 (al proporcionar evidencias en contra)
- Si el p-valor es $\geq \alpha$, no rechazamos H_0 (al proporcionar evidencias a favor)

Apéndice B

Acrónimos

Campo semántico	Acrónimo	Origen	Traducción
Grupos clínicos	AD	Alzheimer's Disease	Enfermedad de Alzheimer
	HC	Healthy Control	Sujeto control
	MCI	Mild cognitive impairment	Sujeto con deterioro cognitivo leve
	sMCI	stable MCI	Sujeto MCI que no transforma a AD
	pMCI	progressive MCI	Sujeto MCI que transforma a AD
Marcadores	CF	Clinical risk factors	Factores de riesgo clínicos
	CAM	Clinical assessments markers	Evaluaciones clínicas cognitivas y funcionales
	MRI	Magnetic resonance imaging	Imagen de resonancia magnética
	APOE-4	Apolipoprotein E-e4	Apolipoproteína E-e4
	CSF	Cerebrospinal fluid proteins	Proteínas del fluido cerebrospinal
	PET	Positron emission tomography	Emisión de positrones
Anatomía cerebral	WM	White matter	Materia blanca
	GM	Grey matter	Materia gris
	ICV	Intra-craneal volumen	Volumen intracraneal
Machine learning	SVM	Support vector machine	Máquinas de soporte vectorial
	LDA	Linear discriminant analysis	Análisis discriminante lineal
	CV	Cross validation	Validación cruzada

Tabla B.1: Tabla de acrónimos

Bibliografía

- [1] Niels Andreasen, Magnus Sjögren, and Kaj Blennow. Csf markers for alzheimer's disease: total tau, phospho-tau and $a\beta_{42}$. *The world journal of biological psychiatry*, 4(4):147–155, 2003.
- [2] Iman Beheshti, Hasan Demirel, Hiroshi Matsuda, Alzheimer's Disease Neuroimaging Initiative, et al. Classification of alzheimer's disease and prediction of mild cognitive impairment-to-alzheimer's conversion from structural magnetic resource imaging using feature ranking and a genetic algorithm. *Computers in biology and medicine*, 83:109–119, 2017.
- [3] Yoav Benjamini and Yosef Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society. Series B (Methodological)*, pages 289–300, 1995.
- [4] Jorge L Bernal-Rusiel, Douglas N Greve, Martin Reuter, Bruce Fischl, Mert R Sabuncu, Alzheimer's Disease Neuroimaging Initiative, et al. Statistical analysis of longitudinal neuroimage data with linear mixed effects models. *Neuroimage*, 66:249–260, 2013.
- [5] Jorge L Bernal-Rusiel, Martin Reuter, Douglas N Greve, Bruce Fischl, Mert R Sabuncu, Alzheimer's Disease Neuroimaging Initiative, et al. Spatiotemporal linear mixed effects modeling for the mass-univariate analysis of longitudinal neuroimage data. *Neuroimage*, 81:358–370, 2013.
- [6] Heiko Braak and Eva Braak. Neuropathological staging of alzheimer-related changes. *Acta neuropathologica*, 82(4):239–259, 1991.
- [7] M. Carmen Tobar Carlos Platero, Lin Lin. Longitudinal neuroimaging hippocampal markers for diagnosing alzheimer's disease.
- [8] Andrea Chincarini, Paolo Bosco, Piero Calvini, Gianluca Gemme, Mario Esposito, Chiara Olivieri, Luca Rei, Sandro Squarcia, Guido Rodriguez, Roberto Bellotti, et al. Local mri analysis approach in the diagnosis of early and prodromal alzheimer's disease. *Neuroimage*, 58(2):469–480, 2011.
- [9] Youngsang Cho, Joon-Kyung Seong, Yong Jeong, Sung Yong Shin, Alzheimer's Disease Neuroimaging Initiative, et al. Individual subject classification for alzheimer's disease based on incremental learning using a spatial frequency representation of cortical thickness data. *Neuroimage*, 59(3):2217–2230, 2012.
- [10] Gordana Derado, F DuBois Bowman, Lijun Zhang, and Alzheimer's Disease Neuroimaging Initiative. Predicting brain activity using a bayesian spatial model. *Statistical methods in medical research*, 22(4):382–397, 2013.

- [11] Rahul S Desikan, Howard J Cabral, Christopher P Hess, William P Dillon, Christine M Glastonbury, Michael W Weiner, Nicholas J Schmansky, Douglas N Greve, David H Salat, Randy L Buckner, et al. Automated mri measures identify individuals with mild cognitive impairment and alzheimer's disease. *Brain*, 132(8):2048–2057, 2009.
- [12] BC Dickerson, TR Stoub, RC Shah, RA Sperling, RJ Killiany, MS Albert, BT Hyman, Deborah Blacker, et al. Alzheimer-signature mri biomarker predicts ad dementia in cognitively normal adults. *Neurology*, 76(16):1395–1402, 2011.
- [13] Chris Ding and Hanchuan Peng. Minimum redundancy feature selection from microarray gene expression data. *Journal of bioinformatics and computational biology*, 3(02):185–205, 2005.
- [14] Bradley Efron. Bootstrap methods: another look at the jackknife. In *Breakthroughs in statistics*, pages 569–593. Springer, 1992.
- [15] Simon F Eskildsen, Pierrick Coupé, Vladimir S Fonov, Jens C Pruessner, and D Louis Collins. Structural imaging biomarkers of alzheimer's disease: predicting disease progression. *Neurobiology of aging*, 36:S23–S31, 2015.
- [16] Francisco L Ferreira, Sandra Cardoso, Dina Silva, Manuela Guerreiro, Alexandre de Mendonça, and Sara C Madeira. Improving prognostic prediction from mild cognitive impairment to alzheimer's disease using genetic algorithms. In *International Conference on Practical Applications of Computational Biology & Bioinformatics*, pages 180–188. Springer, 2017.
- [17] Garrett M Fitzmaurice, Nan M Laird, and James H Ware. *Applied longitudinal analysis*, volume 998. John Wiley & Sons, 2012.
- [18] Giovana Gavidia-Bovadilla, Samir Kanaan-Izquierdo, María Mataró-Serrat, Alexandre Perera-Lluna, Alzheimer's Disease Neuroimaging Initiative, et al. Early prediction of alzheimer's disease using null longitudinal model-based classifiers. *PloS one*, 12(1):e0168011, 2017.
- [19] Emilie Gerardin, Gaël Chételat, Marie Chupin, Rémi Cuingnet, Béatrice Desgranges, Ho-Sung Kim, Marc Niethammer, Bruno Dubois, Stéphane Lehéricy, Line Garnero, et al. Multidimensional classification of hippocampal shape features discriminates alzheimer's disease and mild cognitive impairment from normal aging. *Neuroimage*, 47(4):1476–1486, 2009.
- [20] Edna Grünblatt, Sonja Zehetmayer, Jasmin Bartl, Christiane Löffler, Ildiko Wicht, Michael K Rainer, Susanne Jungwirth, Peter Bauer, Walter Danielczyk, Karl-Heinz Tragl, et al. Genetic risk factors and markers for alzheimer's disease and/or depression in the vita study. *Journal of psychiatric research*, 43(3):298–308, 2009.
- [21] Clifford R Jack Jr, David S Knopman, William J Jagust, Ronald C Petersen, Michael W Weiner, Paul S Aisen, Leslie M Shaw, Prashanthi Vemuri, Heather J Wiste, Stephen D Weigand, et al. Tracking pathophysiological processes in alzheimer's disease: an updated hypothetical model of dynamic biomarkers. *The Lancet Neurology*, 12(2):207–216, 2013.

- [22] Sanjeev M Kamat, Anjali S Kamat, and George T Grossberg. Dementia risk prediction: are we there yet? *Clinics in geriatric medicine*, 26(1):113–123, 2010.
- [23] Robert Katzman. Education and the prevalence of dementia and alzheimer’s disease. *Neurology*, 1993.
- [24] David S Knopman, Steven T DeKosky, JL Cummings, H Chui, J Corey-Bloom, N Relkin, GW Small, B Miller, and JC Stevens. Practice parameter: Diagnosis of dementia (an evidence-based review) report of the quality standards subcommittee of the american academy of neurology. *Neurology*, 56(9):1143–1153, 2001.
- [25] Igor O Korolev, Laura L Symonds, Andrea C Bozoki, Alzheimer’s Disease Neuroimaging Initiative, et al. Predicting progression from mild cognitive impairment to alzheimer’s dementia using clinical, mri, and plasma biomarkers via probabilistic pattern classification. *PloS one*, 11(2):e0138866, 2016.
- [26] Nikolaus Kriegeskorte, W Kyle Simmons, Patrick SF Bellgowan, and Chris I Baker. Circular analysis in systems neuroscience: the dangers of double dipping. *Nature neuroscience*, 12(5):535, 2009.
- [27] Muwei Li, Yuanyuan Qin, Fei Gao, Wenzhen Zhu, and Xiaohai He. Discriminative analysis of multivariate features from structural mri and diffusion tensor images. *Magnetic resonance imaging*, 32(8):1043–1051, 2014.
- [28] Mingxia Liu, Daoqiang Zhang, Dinggang Shen, and Alzheimer’s Disease Neuroimaging Initiative. View-centralized multi-atlas classification for alzheimer’s disease diagnosis. *Human brain mapping*, 36(5):1847–1865, 2015.
- [29] Xin Liu, Duygu Tosun, Michael W Weiner, Norbert Schuff, Alzheimer’s Disease Neuroimaging Initiative, et al. Locally linear embedding (lle) for mri based alzheimer’s disease classification. *Neuroimage*, 83:148–157, 2013.
- [30] A Lobo, LJ Launer, L Fratiglioni, K Andersen, A Di Carlo, MMB Breteler, JRM Copeland, JF Dartigues, C Jagger, J Martinez-Lage, et al. Prevalence of dementia and major subtypes in europe: a collaborative study of population-based cohorts. *Neurology*, 54(5):S4, 2000.
- [31] I McDowell. Alzheimer’s disease: insights from epidemiology. *Aging Clinical and Experimental Research*, 13(3):143–162, 2001.
- [32] Michelle M Mielke, Prashanthi Vemuri, and Walter A Rocca. Clinical epidemiology of alzheimer’s disease: assessing sex and gender differences. *Clinical epidemiology*, 6:37, 2014.
- [33] Elaheh Moradi, Antonietta Pepe, Christian Gaser, Heikki Huttunen, Jussi Tohka, Alzheimer’s Disease Neuroimaging Initiative, et al. Machine learning framework for early mri-based alzheimer’s conversion prediction in mci subjects. *Neuroimage*, 104:398–412, 2015.
- [34] John C Morris. The clinical dementia rating (cdr): current version and scoring rules. *Neurology*, 1993.

- [35] Richard J O'Brien and Philip C Wong. Amyloid precursor protein processing and alzheimer's disease. *Annual review of neuroscience*, 34:185–204, 2011.
- [36] Pedro Paulo de Magalhães Oliveira Jr, Ricardo Nitrini, Geraldo Busatto, Carlos Buchpiguel, João Ricardo Sato, and Edson Amaro Jr. Use of svm methods with surface-based cortical and volumetric subcortical measurements to detect alzheimer's disease. *Journal of Alzheimer's Disease*, 19(4):1263–1272, 2010.
- [37] Hyunjin Park, Jin-ju Yang, Jongbum Seo, Jong-min Lee, et al. Dimensionality reduced cortical features and their use in predicting longitudinal changes in alzheimer's disease. *Neuroscience letters*, 550:17–22, 2013.
- [38] RI Pfeffer, TT Kurosaki, CH Harrah Jr, JM Chance, and S Filos. Measurement of functional activities in older adults in the community. *Journal of gerontology*, 37(3):323–329, 1982.
- [39] Claudia Plant, Stefan J Teipel, Annahita Oswald, Christian Böhm, Thomas Meindl, Janaina Mourao-Miranda, Arun W Bokde, Harald Hampel, and Michael Ewers. Automated detection of brain atrophy patterns based on mri for the prediction of alzheimer's disease. *Neuroimage*, 50(1):162–174, 2010.
- [40] Martin Prince, Renata Bryce, Emiliano Albanese, Anders Wimo, Wagner Ribeiro, and Cleusa P Ferri. The global prevalence of dementia: a systematic review and metaanalysis. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, 9(1):63–75, 2013.
- [41] Martin Reuter, H Diana Rosas, and Bruce Fischl. Highly accurate inverse consistent registration: a robust approach. *Neuroimage*, 53(4):1181–1196, 2010.
- [42] Martin Reuter, Nicholas J Schmansky, H Diana Rosas, and Bruce Fischl. Within-subject template estimation for unbiased longitudinal image analysis. *Neuroimage*, 61(4):1402–1418, 2012.
- [43] André Rey. L'examen clinique en psychologie. 1958.
- [44] Wilma G Rosen, Richard C Mohs, and Kenneth L Davis. A new rating scale for alzheimer's disease. *The American journal of psychiatry*, 1984.
- [45] Alberto Serrano-Pozo, Matthew P Frosch, Eliezer Masliah, and Bradley T Hyman. Neuropathological alterations in alzheimer disease. *Cold Spring Harbor perspectives in medicine*, 1(1):a006189, 2011.
- [46] Lauge Sørensen, Christian Igel, Akshay Pai, Ioana Balas, Cecilie Anker, Martin Lillholm, Mads Nielsen, Alzheimer's Disease Neuroimaging Initiative, et al. Differential diagnosis of mild cognitive impairment and alzheimer's disease using structural mri cortical thickness, hippocampal shape, hippocampal texture, and volumetry. *NeuroImage: Clinical*, 13:470–482, 2017.
- [47] Xiaoying Tang, Dominic Holland, Anders M Dale, Laurent Younes, and Michael I Miller. Baseline shape diffeomorphometry patterns of subcortical and ventricular structures in predicting conversion of mild cognitive impairment to alzheimer's disease. *Journal of Alzheimer's Disease*, 44(2):599–611, 2015.

- [48] Wesley K Thompson, Joachim Hallmayer, Ruth O'Hara, and Alzheimer's Disease Neuroimaging Initiative. Design considerations for characterizing psychiatric trajectories across the lifespan: application to effects of apoe- ϵ 4 on cerebral cortical thickness in alzheimer's disease. *American Journal of Psychiatry*, 168(9):894–903, 2011.
- [49] Tom N Tombaugh and Nancy J McIntyre. The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society*, 40(9):922–935, 1992.
- [50] Sander CJ Verfaillie, Rosalinde E Slot, Betty M Tijms, Femke Bouwman, Marije R Benedictus, Jozefien M Overbeek, Teddy Koene, Hugo Vrenken, Philip Scheltens, Frederik Barkhof, et al. Thinner cortex in patients with subjective cognitive decline is associated with steeper decline of memory. *Neurobiology of aging*, 61:238–244, 2018.
- [51] Chong-Yaw Wee, Pew-Thian Yap, Dinggang Shen, and Alzheimer's Disease Neuroimaging Initiative. Prediction of alzheimer's disease and mild cognitive impairment using cortical morphological patterns. *Human brain mapping*, 34(12):3411–3425, 2013.
- [52] Eric Westman, Carlos Aguilar, J-Sebastian Muehlboeck, and Andrew Simmons. Regional magnetic resonance imaging measures for multivariate analysis in alzheimer's disease and mild cognitive impairment. *Brain topography*, 26(1):9–23, 2013.
- [53] Anders Wimo, Linus Jönsson, John Bond, Martin Prince, and Bengt Winblad. The worldwide economic impact of dementia 2010. *Alzheimer's & dementia: the journal of the Alzheimer's Association*, 9(1):1–11, 2013.