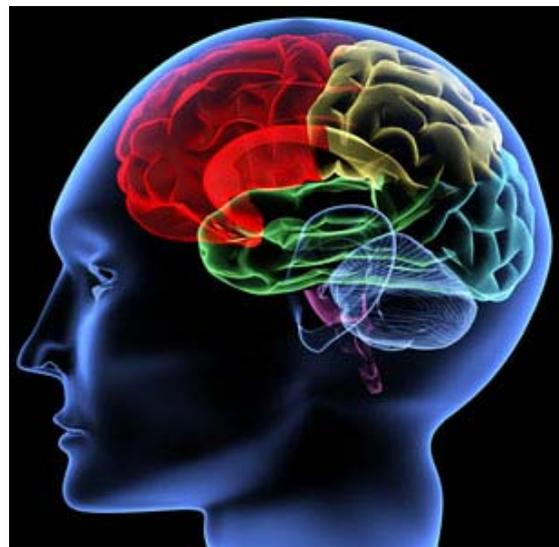


Clasificadores basados en Máquinas de Soporte Vertical para el Diagnóstico y Predicción de la Enfermedad de Alzheimer

G. Gavidia
E. Soudah
E. Oñate



Clasificadores basados en Máquinas de Soporte Vectorial para el Diagnóstico y Predicción de la Enfermedad de Alzheimer

G. Gavidia
E. Soudah
E. Oñate

CENTRO INTERNACIONAL DE MÉTODOS NUMÉRICOS EN INGENIERÍA
Edificio C1, Campus Norte UPC
Gran Capitán s/n
08034 Barcelona, España
www.cimne.com

Primera edición: Septiembre 2012

**CALSIFICADORES BASADOS EN MÁQUINAS DE SOPORTE VECTORIAL PARA EL DIAGNÓSTICO
Y PREDICCIÓN DE LA ENFERMEDAD DE ALZHEIMER**

Monografía CIMNE M131

© Los autores

ISBN: 978-84-940243-4-4

Depósito legal: B-27594-2012

En memoria de la Prof. Núria Durany Pich

La prof. Nuria, profesora de la facultad de Medicina de la Universitat de Barcelona, nació en Barcelona en 1961 y falleció en agosto del 2010.

Cursó la carrera de Bioquímica en Nueva Gales del Sur (Australia) y se interesó desde muy joven por las neurociencias, colaborando con el Banco de Cerebros en Würzburg, donde, con el profesor Riederer, elabora su tesis doctoral sobre cáncer cerebral.

Posteriormente vinieron importantes trabajos sobre la enfermedad de Alzheimer y esquizofrenia, los cuales impulsaron el desarrollo de proyectos de investigación en colaboración con CIMNE. Asumió importantes cargos directivos, como son el decanato de la facultad de Ciencias de la Salud y el vicerrectorado de Investigación; y dirigió proyectos internacionales de gran envergadura. Como fruto de su brillante e incesante carrera investigadora, en el 2008 ingresó en la Real Academia de Doctores.

En el recuerdo quedan su extraordinaria tenacidad, profesionalidad, amabilidad y sobre todo su gran ejemplo de entrega al trabajo.

Resumen

La enfermedad de Alzheimer (AD) es un desorden neurodegenerativo cuyo diagnóstico clínico es realizado después de excluir otros tipos de demencia y el diagnóstico clínico definitivo requiere además de la presencia de un alto déficit cognitivo, la confirmación histológica mediante autopsia de la presencia de proteínas como las neurofibras de tau (τ) y amyloid beta ($A\beta$) en los tejidos cerebrales. AD es una de las enfermedades de mayor impacto social en Europa y América. En el pasado 2005, en Europa fueron diagnosticados 3.600.000 pacientes afectados por AD (fuente: Frost & Sullivan) y en una reciente investigación impulsada por la asociación *Alzheimer Europe*, se estimó que 7.3 millones de personas padecen algún tipo de demencia. A pesar de que la ocurrencia de AD no es un suceso normal en la población mayor, el riesgo de desarrollar la enfermedad se incrementa en esa etapa.

A pesar que el diagnóstico definitivo es realizado post-muerte, un diagnóstico clínico de “probable” o “posible” AD puede ser realizado durante la vida de los pacientes, incluso en etapas muy tempranas de la enfermedad. En consecuencia, un tratamiento eficaz de AD durante las etapas iniciales tendría el mayor impacto debido a que la función cognitiva podría ser conservada en el nivel más alto posible.

Recientes estudios dirigidos por el *National Institute on Aging* (NIA) y *Alzheimer's Association* han redefinido las etapas establecidas para AD y los originales criterios establecidos por el estándar NINCDS-ADRDA (*The National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association*) para el diagnóstico de AD [Albert et al., 2011; McKhann et al., 2011; Sperling et al., 2011]. De este modo, tres etapas han sido establecidas: (1) AD-preclínico, (2) Deterioro Cognitivo Leve (MCI) ocasionado por AD, y (3) Demencia ocasionada por AD; cada una de las cuales implica criterios y características específicas. Asimismo, otro importante avance reportado por estos estudios constituye la incorporación de los biomarcadores en los criterios de diagnóstico orientados a investigación de AD y ensayos clínicos.

Existen considerables esfuerzos en identificar biomarcadores que proporcionen información adicional de AD y sirvan como sustitutos in vivo para la detección de los cambios patológicos en las diferentes etapas de la enfermedad. Al respecto, existen tres clases de biomarcadores comúnmente utilizados: (1) biomarcadores genéticos, (2) biomarcadores neuroquímicos, y (3) biomarcadores basados en imágenes médicas. La utilización de los biomarcadores basados en neuroimágenes se ha venido incrementando en las dos últimas décadas. Estos biomarcadores buscan diagnosticar y predecir las etapas tempranas de AD, describiendo las diferencias de las estructuras cerebrales y la relación entre anomalías de las zonas cerebrales y el desarrollo de la enfermedad. Son obtenidos a partir de las imágenes neurológicas de AD: (1) estructurales como las imágenes de Tomografía Computacional (CT) y las imágenes de Resonancia Magnética (RM); y (2) funcionales como las tomografías de emisión de fotón único (SPECT: Single Photon Emission Tomography) y la tomografía por emisión de positrones (PET: Positron-Emission Tomography).

Sin embargo, en el día a día del diagnóstico clínico de AD, la disposición de los biomarcadores morfométricos no resulta una tarea fácil y directa. El progreso del diagnóstico por imágenes médicas iniciado durante las pasadas décadas permite obtener una mayor cantidad de información funcional y anatómica del cerebro de cualquier individuo con una enorme precisión en tiempo y en espacio; sin embargo, es imposible analizar directamente esta gran cantidad de datos e imágenes, siendo necesario desarrollar nuevos modelos de cálculo que capturen y seleccionen los parámetros morfométricos de las zonas del cerebro relevantes para un eficiente diagnóstico de AD.

En este trabajo, se estableció una metodología de soporte al diagnóstico de AD, principalmente en sus etapas MCI ocasionado por AD y demencia ocasionada por AD. Para este fin fueron obtenidos casos clínicos de dos proyectos de investigación en demencia del tipo AD de reconocida trayectoria: la base de datos *Alzheimer's Disease Neuroimaging Initiative* (ADNI) (www.loni.ucla.edu/ADNI) y la base de datos *The Open Access Series of Imaging Studies* (OASIS) (<http://www.oasis-brains.org/>). Asimismo, fueron establecidas dos tareas principales: la selección de variables predictoras de AD y la construcción de modelos de clasificación basados en máquinas de soporte vectorial (SVM), entrenados a partir de las variables seleccionadas. Las variables predictoras seleccionadas estuvieron conformadas por biomarcadores morfométricos y características socio-demográficas y neuropsicológicas. Estas variables

deberían ser útiles para la discriminación de casos clínicos en tres estados: (1) Estado normal (generalmente personas mayores sanas); (2) Estado MCI ocasionado por AD; y (3) Etapa de demencia ocasionada por AD. Por otro lado, los modelos SVM estarían enfocados a dos tareas principales: (1) Diagnóstico de AD mediante la discriminación entre sujetos sanos y sujetos con AD; y (2) Predicción de AD, orientada a la discriminación de sujetos MCI con riesgo de convertirse a AD y sujetos MCI sin riesgo de conversión. Asimismo, estos modelos deberían garantizar resultados aceptables, respecto a la sensibilidad y especificidad de las tareas de clasificación.

Los resultados obtenidos en esta investigación son prometedores. Por un lado, el subconjunto de variables seleccionadas como relevantes para el diagnóstico de AD, tienen correlación con los resultados de investigaciones previas. Asimismo, en la etapa de testeo, los resultados demostraron que los modelos SVM son de gran utilidad para el soporte diagnóstico clínico de esta enfermedad, siendo capaces de discriminar sujetos con AD de sujetos sanos (diagnóstico) con una exactitud mayor al 99% y distinguir a los sujetos MCI con riesgo de conversión a AD de los sujetos MCI sin riesgo de conversión (predicción) con una exactitud superior al 94%.

Contenidos

Contenidos	v
Lista de figuras	xi
Indice de tablas	xv
1 Introducción y Alcances	1
1.1 Introducción	1
1.2 La enfermedad de Alzheimer (AD)	3
1.2.1 Fisiopatología de AD.	5
1.2.2 Progresión de $A\beta$ en el transcurso de AD.	7
1.3 Diagnóstico de AD.	7
1.3.1 Etapas de AD.	11
1.3.2 Diagnóstico en la etapa DAT.	13
1.3.3 Diagnóstico en la etapa MCI_{AD}	16
1.3.4 Diagnóstico en la etapa de AD-preclínico	17
1.3.5 Herramientas clínicas y neuropsicológicas para el diagnóstico clínico de AD	19
1.4 Biomarcadores de AD.	23
1.4.1 Biomarcadores de AD basados en neuroimagen.	25
1.5 Objetivos	27
2 Biomarcadores morfométricos obtenidos de MRI.	29
2.1 Introducción	29
2.2 Descripción.	30
2.3 Aplicación de los biomarcadores morfométricos en la predicción y diagnóstico de AD.	32
2.3.1 Estudios de MRIs en la etapa AD-preclínica.	33
2.3.2 Estudios de MRIs en la etapa MCI_{AD}	34

CONTENIDOS

2.3.3	Estudios de MRIs en la etapa de demencia AD.	34
2.4	Obtención de los biomarcadores morfométricos.	35
2.4.1	Herramientas automáticas.	36
2.4.2	Principales Rutinas de Procesamiento y Análisis de MRIs.	38
2.4.2.1	Eliminación de rasgos faciales (<i>defacing</i>).	38
2.4.2.2	Alineamiento de multisecciones de MRIs y corrección respecto a un atlas del cerebro	39
2.4.2.3	Eliminación del Bias.	40
2.4.2.4	Eliminación de las estructuras no cerebrales.	41
2.4.2.5	Cuantificación Morfométrica.	42
2.5	Enfoques de clasificación basados en biomarcadores morfométricos	42
2.5.1	Clasificación basada en <i>vóxel</i>	43
2.5.2	Clasificación basada en vértices (enfoque cortical).	44
2.5.3	Clasificación basada en ROIs.	45
2.5.4	Clasificación mixta.	48
2.5.5	Ventajas y desventajas de los enfoques de clasificación.	48
3	Modelos de Clasificación de AD basada en SVM.	51
3.1	Introducción.	51
3.2	Máquinas de soporte vectorial (SVM): Visión general	53
3.2.1	Clasificadores.	54
3.2.2	Función kernel	56
3.2.3	Definición de SVM.	56
3.3	Etapas de construcción de un clasificador SVM	59
3.3.1	Preparación de datos para entrenamiento y test.	59
3.3.2	Entrenamiento y Test.	60
3.3.2.1	Selección de la función kernel y búsqueda de los mejores parámetros.	60
3.3.2.2	Entrenamiento y test	61
3.3.3	Evaluación del desempeño de los clasificadores	61
3.3.3.1	Matriz de confusión	63
3.3.3.2	Mediciones basadas en la matriz de confusión	63
3.3.3.3	Factores que afectan el desempeño de un clasificador	67
3.4	Selección de variables	67
3.4.1	Ranking de variables	68
3.4.1.1	Método de Mínima Redundancia y Máxima Relevancia	69

3.4.1.2	Método F-score	70
3.4.2	Métodos Wrapper	70
3.4.2.1	SVM como un método wrapper de selección de variables	71
3.5	Revisión de los estudios SVM aplicados en el diagnóstico de AD	71
4	Construcción de los modelos de diagnóstico y predicción de AD basados en SVM.	75
4.1	Introducción	75
4.2	Metodología propuesta	76
4.3	Obtención de datos de ensayos clínicos de AD	79
4.3.1	Fiabilidad de los datos de ensayo clínicos utilizados	80
4.3.2	Criterios de inclusión de los casos clínicos	81
4.3.3	Definición de la variable de clasificación clínica	82
4.3.4	Distribución de los datos	83
4.3.5	Variables morfométricas basadas en imágenes de Resonancia Magnética Estructural	84
4.3.6	Variables clínicas: demográficas y neuropsicológicas	89
4.3.7	Preprocesamiento de las variables	90
4.4	Selección de variables	91
4.4.1	Análisis inicial de las variables. Curvas ROC	91
4.4.2	Ranking de variables	92
4.4.3	Selección del subconjunto de variables	94
4.4.3.1	Configuración del clasificador SVM	95
4.4.3.2	Selección de variables morfométricas	96
4.4.3.3	Selección de variables clínicas	100
4.4.4	Discriminación entre los métodos de selección de variables.	101
5	Entrenamiento, Test y Evaluación de los modelos	107
5.1	Introducción	107
5.2	Entrenamiento de los modelos SVM de diagnóstico y predicción de AD.	109
5.2.1	Selección del kernel de modelo SVM	109
5.2.2	Búsqueda de los mejores parámetros de configuración de los modelos SVM.	109
5.2.3	Entrenamiento y validación	111
5.3	Test de los modelos SVM.	112
5.3.1	Casos de Estudio ADNI: Diagnóstico y predicción de AD	112

CONTENIDOS

5.3.2	Casos de estudio OASIS: Diagnóstico de AD	112
5.3.2.1	Obtención de los casos clínicos	112
5.3.2.2	Análisis de las variables de OASIS	115
5.3.2.3	Re-entrenamiento del modelo de diagnóstico	115
5.3.2.4	Clasificación	115
5.3.3	Desempeño de los modelos de diagnóstico y predicción de AD	116
6	Análisis de Resultados, Conclusiones y Trabajo Futuro	119
6.1	Introducción	119
6.2	Análisis de Resultados	120
6.2.1	Diagnóstico y predicción de AD. Problema de clasificación no-lineal . . .	120
6.2.2	Desempeño de los modelos de diagnóstico y predicción de AD	120
6.2.3	Influencia de las características sociodemográficas y neuropsicológicas en la enfermedad	122
6.2.4	Tiempo de procesamiento	124
6.3	Conclusiones	124
6.3.1	Máquinas de soporte vectorial (SVM) para el diagnóstico y predicción de AD	124
6.3.2	Variables predictoras seleccionadas	124
6.3.2.1	Correlación de las variables predictoras con otras investigaciones	124
6.3.2.2	Viabilidad de la obtención de las variables de entrada	124
6.3.2.3	Enfoque de variables predictoras singulares vs. enfoque de variables predictoras combinadas	125
6.3.2.4	Importancia de los biomarcadores morfométricos en el diagnóstico y predicción de AD	125
6.4	Limitaciones de este estudio	125
6.5	Trabajo futuro	126
6.5.1	Inclusión de nuevos biomarcadores de AD	126
6.5.2	Desarrollo e integración de un módulo de procesamiento y análisis de imágenes cerebrales.	126
	Apéndice A: Tablas y Gráficos	129
.1	Resultados de los métodos de selección de variables sobre las 318 variables morfológicas originales. Calificación final	129

CONTENIDOS

.2	Resultados de los métodos de selección de variables sobre las 6 variables clínicas originales. Calificación final	138
	Referencias	139

CONTENIDOS

Lista de figuras

1.1	Incremento/Decrecimiento de las principales causas de muerte en USA entre el periodo 2000-2008. Fuente: [Association, 2012]	4
1.2	Hipótesis “ <i>Amyloid cascade</i> ” de la fisiopatología de AD. 1.2(a) Diagrama original de la hipótesis propuesto por [Hardy and Selkoe, 2002]. 1.2(a) Representación gráfica propuesta por [Cummings, 2004].	6
1.3	Fase de progresión de $A\beta$ en el cerebro. Fase 1: $A\beta$ está muy restringido a las áreas neocorticales del cerebro. Fases 2 y 3, los depósitos del amiloide son localizados a lo largo del neocórtex, la allocorteza y en la corteza motora y las estructuras subcorticales. Fases 4 y 5, los depósitos del amiloide se observan a nivel global, incluyendo el tronco cerebral y en cerebelo en la etapa final de la enfermedad (fase 5). Clave: blanco, no hay depósitos de $A\beta$; rojo, nueva deposición de $A\beta$; gris, depósitos de $A\beta$ que no estaban presentes en la fase 1; negro, depósitos de $A\beta$ que ya estaban presentes en la fase 1. Fuente: [Ewers et al., 2011; Thal et al., 2002]	8
1.4	Modelo de la trayectoria clínica de AD. Fuente:[Sperling et al., 2011]	12
1.5	Ilustración del cambio del tejido cerebral y las células nerviosas en los estados: sano, MCI y AD severo. Fuente: <i>The American Health Assistance Foundation</i>	14
1.6	Modelo de las etapas de AD basado en biomarcadores. Fuentes:[Sperling et al., 2011] (adaptación),[Jack et al., 2010a] (original)	19
1.7	Esquema de clasificación de las subetapas del AD-preclínico. Fuente:[Sperling et al., 2011]	20
2.1	Modelo hipotético de los principales biomarcadores basados en neuroimágenes y su utilidad predictiva durante la etapas de AD. Fuente: [Ewers et al., 2011].	33

LISTA DE FIGURAS

2.2	Vistas ortogonales de tres sesiones de MRI (con cortos intervalos de tiempo) sin rasgos faciales. 2.2(a): sesión mpr-1. 2.2(a):sesión mpr-2. 2.2(a): sesión mpr-3. Fuente: Sesiones de MRIs del cerebro proporcionadas por el departamento de Biomedicina del <i>Centre Internacional de Mètodes Numèrics en Enginyeria</i> (CIMNE) (http://www.cimne.upc.edu) y procesadas con Freesurfer (http://surfer.nmr.mgh.harvard.edu/fswiki)	39
2.3	Vistas ortogonales de las MRIs resultantes después de aplicar las técnicas de alineamiento. Fuente: Sesiones de MRIs del cerebro proporcionadas por el departamento de Biomedicina del <i>Centre Internacional de Mètodes Numèrics en Enginyeria</i> (CIMNE) (http://www.cimne.upc.edu) y procesadas con Freesurfer (http://surfer.nmr.mgh.harvard.edu/fswiki)	40
2.4	Vistas ortogonales de la sesión de MRI corregistrada con respecto a un sistema de coordenadas del atlas Talairach. Fuente: Sesiones de MRIs del cerebro proporcionadas por el departamento de Biomedicina del <i>Centre Internacional de Mètodes Numèrics en Enginyeria</i> (CIMNE) (http://www.cimne.upc.edu) y procesadas con Freesurfer (http://surfer.nmr.mgh.harvard.edu/fswiki)	41
2.5	Vistas ortogonales de la imagen cerebral obtenida después de extraer las estructuras no-cerebrales.	41
2.6	Patrones anatómicos con máximo poder de discriminación entre sujetos sAD y sC localizados con modelo de clasificación SVM obtenido por [Ewers et al., 2011].	44
2.7	Representación tridimensional de los 34 ROIs (vista de un hemisferio) examinados en el estudio de [Desikan et al., 2009] y originalmente propuestos por el método de etiquetado de la corteza cerebral propuesto en [Desikan et al., 2006]	46
2.8	Ejemplo de los subcampos del hipocampo marcados manualmente: (1) Corte inicial; (2,3 y 4) Cortes intermedios; (5) Corte más posterior. Etiquetas: Roja, corteza entorrinal; Amarilla, subículo; Beige, CA1; azul, zona de transición CA1-2 ; Marrón, CA3&DG. Fuente:[Mueller et al., 2010].	47
3.1	Principales tópicos de ML. Fuente:[Wang and Summers, 2012].	53
3.2	Representación del aprendizaje de un clasificador.	55
3.3	Ejemplo de separación lineal de un problema de clasificación ideal.	56
3.4	Hipotética representación de un problema de clasificación binario: sC (puntos rojos) vs. sAD (puntos azules): (a) Diferentes posibilidades de separación de las clases; (b) Separación óptima obtenida con SVM.	62
3.5	Espacio ROC y la representación de los ejemplos de valores AUC.	66

LISTA DE FIGURAS

4.1	Principales etapas implicadas en la metodología de selección de variables y construcción de los modelos SVM para el diagnóstico y predicción de AD	78
4.2	Gráfica de las curvas ROC (AUC) obtenidas por las 324 variables originales: (a) 6 variables clínicas; (b) 318 variables morfométricas originales	93
4.3	Distribución no lineal de los grupos de entrenamiento y test para sC vs. sAD: (a) grupo sCsAD _{ADNI-train} . (b) Grupo sCsAD _{ADNI-test}	97
4.4	Gráfica de las curvas ROC (AUC) obtenidas por las 10 primeras variables morfométricas seleccionadas con F-score: (a) Resultados sobre el grupo de entrenamiento y validación sCsAD _{ADNI-train} . (b) Resultados sobre el grupo de test sCsAD _{ADNI-test}	99
4.5	Matrices de confusión obtenidas con los métodos de selección de variables. Variables morfométricas: (a) Método F-score, (b) Método SVM Wrapper, (c) Método F-score + SVM Wrapper. Variables clínicas: (d) Método F-score, (e) Método SVM Wrapper, (f) Método F-score + SVM Wrapper	104
4.6	Curvas ROC de las 29 variables seleccionadas. Valores AUC calculados sobre el grupo sCsAD _{ADNI-train}	105
5.1	Etapas implicadas de la construcción de los modelos SVM para el diagnóstico y predicción de AD	108
5.2	Distribución no lineal de los grupos de entrenamiento y test para el modelo de predicción: (a) grupo sMCI _{ADNI-train} . (b) Grupo sMCI _{ADNI-test}	110
5.3	Resultados del modelo de diagnóstico en la clasificación del grupo sCsAD _{ADNI-test} : (a) Matriz de confusión. (b) Curva ROC	113
5.4	Resultados del modelo de predicción en la clasificación del grupo sMCI _{ADNI-test} : (a) Matriz de confusión. (b) Curva ROC	114
5.5	Resultados del modelo de diagnóstico en la clasificación del grupo sCsAD _{OASIS-45} : (a) matriz de confusión. (b) curva ROC.	116
5.6	Resultados del modelo de diagnóstico en la clasificación del grupo sCsAD _{OASIS-7} : (a) matriz de confusión. (b) curva ROC.	117
6.1	Vista general de los principales procesos implicados en la obtención de biomarcadores morfométricos basados en MRIs	127

LISTA DE FIGURAS

Índice de tablas

1.1	Limitaciones de los criterios establecidos en el estándar NINCDS-ADRDA en 1984. Fuente: [DeKosky et al., 2011]	10
1.2	Etapas categóricas para la investigación en AD preclínico. Fuente: [Sperling et al., 2011]	19
3.1	Representación de una matriz de confusión	63
3.2	Clasificación AUC	67
3.3	Investigaciones en AD relacionadas con biomarcadores basados en imágenes y SVM	73
4.1	Clasificación de los participantes según sus características clínicas	82
4.2	Distribución de sujetos e instancias en los grupos de entrenamiento, validación y testeo de los modelos	84
4.3	Tipos de mediciones morfométricas consideradas en este estudio	85
4.4	Variables morfométricas basadas en MRIs	86
4.5	Variables morfométricas de la región no-cortical consideradas en este estudio	87
4.6	Variables clínicas: características demográficas y neuropsicológicas de los grupos $sCsAD_{ADNI-train}$, $sMCI_{ADNI-train}$, $sCsAD_{ADNI-test}$ y $sMCI_{ADNI-test}$	90
4.7	Lista de las 10 primeras variables morfométricas seleccionadas con el método F-score	98
4.8	Indicadores estadísticos de los seis modelos clasificadores auxiliares durante la etapa de test	102
4.9	Lista del subconjunto definitivo de 29 variables predictoras seleccionadas	103
5.1	Comparación de la variable “Educación” en los estudios ADNI y OASIS	115
5.2	Desempeño de los modelos SVM en el diagnóstico y predicción de AD.	118
6.1	Influencia de las características sociodemográficas y neuropsicológicas en AD. Resultados sobre la matriz de confusión de los grupos	123

INDICE DE TABLAS

6.2	Influencia de las características sociodemográficas y neuropsicológicas en AD. Resultados sobre todos los casos positivos y negativos de los grupos	123
3	Calificación final de las 318 variables morfológicas originales. CTAvg :Cortical Thickness Average, VCP :Volume from Cortical Parcellation, VWM :Volume from White Matter, CTStd :Cortical Thickness Standard Deviation	129
4	Calificación final de las 6 variables clínicas originales.	138

Lista de Notaciones

Símbolos Romanos

- τ Tau. Es una proteína atípica depositada en el cerebro.
- $A\beta$ Amyloid beta. Es una proteína atípica depositada en el cerebro
- AD* Enfermedad de Alzheimer
- ANNs* *Artifical Neural Networks*/ Redes Neuronales Artificiales.
- CSF* Fluido cerebro espinal / líquido céfalo-raquídeo.
- MCI* Mild Cognitive Impairment/Deterioro Cognitivo Leve
- ML* *Machine Learning*/ Máquinas de aprendizaje automático.
- OASIS* *Open Access Series of Imaging Studies*. Base de datos de estudios transversales y longitudinales de AD
- sAD* Paciente con AD que tiene un índice de MMSE entre 20 y 26 (inclusive), un índice de CDR de 0.5 o 1.0, y reúne criterios del estándar (NINCDS/ADRDA) para probable caso de AD. Durante la construcción de modelos de clasificación corresponde a una instancia o muestra de datos.
- sC* Sujeto normal con valores de MMSE entre 24 y 30 (inclusive), un índice de CDR de 0, y calificado como no depresivos, no MCI y no demente. Durante la construcción de modelos de clasificación corresponde a una instancia o muestra de datos.
- sMCI* Paciente MCI que tiene un índice MMSE entre 24 y 30 (inclusive), reporta síntomas de problemas con la memoria, dificultades con la memoria objetiva, un CDR de 0.5, ausencia de niveles significativos de deterioro en otros dominios cognitivos, preservan las actividades esenciales de su vida diaria y poseen ausencia de demencia. Durante la construcción de modelos de clasificación corresponde a una instancia o muestra de datos.

LISTA DE NOTACIONES

- sMCIc* Instancia o muestra de datos de un sujeto diagnosticado con MCI convertido a AD
- sMCInc* Instancia o muestra de datos de un sujeto diagnosticado con MCI que no se ha convertido a AD. Sujeto MCI estable
- SVM* *Support Vector Machine* / Máquina de soporte vectorial.
- ACC *Accuracy* / Exactitud
- ADNI *The Alzheimer's Disease Neuroimaging Initiative*. Proyecto de investigación en demencia, principalmente de AD
- AUC *Area Under the ROC Curve* / Área bajo la curva ROC
- CDR *Clinical Dementia Rating* / Clasificación Clínica de la Demencia: CDR=0, no deterioro; CDR=0.5, muy leve; CDR=1, leve; CDR=2, moderado; y CDR=3, demencia severa
- DAT Dementia due to AD / demencia ocasionada por AD
- FNR False Negative Rate / Razón de falsos negativos
- FPR False Positive Rate / Razón de falsos positivos
- GM *Gray matter* / Materia o sustancia gris
- MCI_{AD} *MCI due to AD* / MCI ocasionado por AD
- MMSE *Mini-Mental State Examination* / Examen del Estado Mental Mínimo
- MRI *Magnetic Resonance Imaging* / Imágenes de Resonancia Magnética
- MTL *Medial Temporal Lobes* / lóbulos temporal medial
- NIA *The National Institute on Aging*
- NINCDS-ADRDA Estándar *The National Institute of Neurological and Communicative Disorders and Stroke* y el *Alzheimer's Disease and Related Disorders Association*
- PREC *Precision* / Precisión
- ROC *"Receiver Operating Characteristic"* / Receptor de Funcionamiento Característico
- ROI *Region of Interest* / Región de interés

LISTA DE NOTACIONES

- $sCsAD_{ADNI-test}$ Grupo de datos de ensayos clínicos obtenidos de ADNI, utilizado para la fase de test de los modelos de diagnóstico sC vs. sAD. Conformado por 1145 instancias: 591 instancias correspondientes a 204 sC y 554 instancias correspondientes a 246 sAD
- $sCsAD_{ADNI-train}$ Grupo de datos de ensayos clínicos obtenidos de ADNI, utilizado para la fase de entrenamiento y validación de los modelos de diagnóstico sC vs. sAD. Este grupo estuvo conformado por 1265 instancias: 631 instancias correspondientes a 216 sC y 634 instancias correspondientes a 287 sAD
- $sCsAD_{OASIS-45}$ Grupo de 45 casos clínicos sC(22) y sAD(23) seleccionados del estudio transversal de OASIS
- $sCsAD_{OASIS-7}$ Grupo de 7 casos clínicos seleccionados del estudio longitudinal de OASIS. Las MRIs de estos sujetos fueron procesadas para obtener las variables morfométricas
- $sMCI_{ADNI-test}$ Grupo de datos de ensayos clínicos obtenidos de ADNI, utilizado para la fase de test de los modelos de predicción sMCIc vs. sMCIc. Conformado por 814 instancias clasificadas inicialmente como sMCI, correspondientes a 297 pacientes
- $sMCI_{ADNI-train}$ Grupo de datos de ensayos clínicos obtenidos de ADNI, utilizado para la fase de entrenamiento y validación de los modelos de predicción sMCIc vs. sMCIc
- TNR *Specificity / True Negative Rate / Especificidad*
- TPR *Sensitivity / True Positive Rate / Sensibilidad*
- WM *White matter / Materia o sustancia blanca*

LISTA DE NOTACIONES

Capítulo 1

Introducción y Alcances

1.1 Introducción

La enfermedad de Alzheimer (*Alzheimer's Disease*, AD) es un desorden neurodegenerativo, cuyo diagnóstico es realizado después de excluir otros tipos de demencia. AD es una de las causas más comunes de demencia en el adulto mayor y es una de las enfermedades de mayor impacto social en Europa y América [Vellas et al., 2011]. En el 2005, en Europa fueron diagnosticados 3.600.000 pacientes afectados por AD (fuente: Frost & Sullivan) y en una reciente investigación impulsada por la asociación *Alzheimer Europe*, se estimó que 7.3 millones de personas padecen algún tipo de demencia. A pesar que la ocurrencia de AD no es un suceso normal en la población mayor, el riesgo de desarrollar la enfermedad se incrementa en esa etapa. En consecuencia, teniendo en cuenta el incremento anual de las personas mayores de 80 años en nuestra sociedad, esta enfermedad es uno de los retos más importantes a los que debemos enfrentarnos en este siglo.

El diagnóstico definitivo de AD requiere además de la presencia de un alto déficit cognitivo, la confirmación histológica mediante autopsia, a fin de verificar la presencia de proteínas como las neurofibras de tau (τ) y amyloid beta ($A\beta$, Abeta) en el fluido cerebro espinal (CSF) . Sin embargo, durante la vida de los pacientes es posible hacer un diagnóstico clínico, el cual sólo puede ser considerado como “probable” o “posible” [McKhann et al., 1984]. Al respecto, un diagnóstico preciso y temprano de la enfermedad es crucial en la perspectiva de futuros tratamientos. El diagnóstico preciso es requerido debido a la existencia de pacientes con otro tipo de demencia que comparten las características clínicas de los sujetos con AD [Knopman et al., 2001], los cuales podrían no responder bien ante un tratamiento específico de AD [Mueller et al., 2005b]. Por otro lado, el diagnóstico temprano se justifica por la evidencia creciente de que los mecanismos patogénicos moleculares de AD se activan varios años antes de que las neuronas comiencen a morir y se manifiesten los déficit cognitivos. Durante esta etapa, un tratamiento

1. INTRODUCCIÓN Y ALCANCES

eficaz de AD tendría el mayor impacto debido a que la función cognitiva podría ser conservada en el nivel más alto posible. En consecuencia, existe un gran interés en caracterizar los primeros signos clínicos del proceso degenerativo, el cual podría evolucionar a AD. Este interés ha dado lugar al desarrollo del concepto de *Deterioro Cognitivo Leve (Mild Cognitive Impairment, MCI)*. MCI representa el estado de transición entre el envejecimiento normal y la demencia del tipo AD [Morris et al., 2001].

En la práctica clínica, el diagnóstico de AD está basado principalmente en calificaciones neuropsicológicas, análisis clínicos y entrevistas con el paciente y familiares, las cuales son utilizadas para controlar la progresión de la enfermedad y la eficacia del tratamiento. Sin embargo, a pesar de que estas pruebas reflejan un aspecto importante de la progresión de AD (sobre todo el deterioro funcional), sus resultados en las últimas etapas de la enfermedad podrían no ser exactos [Mueller et al., 2005c].

Por estas limitaciones, principalmente aquéllas por las cuales la aplicación de las pruebas clínicas convencionales de AD pueden fallar, y la necesidad de controlar esta enfermedad a tiempo para mejorar la calidad de vida de los afectados, se ha producido un considerable esfuerzo en identificar adicionales biomarcadores capaces de “marcar” la presencia de la patología, incluso en etapas muy tempranas [Sperling et al., 2011]. Los biomarcadores de AD tienen como meta principal, brindar información adicional de esta enfermedad en las diferentes etapas de AD, así como del progreso de los tratamientos clínicos aplicados [Schneider et al., 2011; Vellas et al., 2011]. Sin embargo, la determinación de cuáles son los mejores biomarcadores que cumplan de manera segura y eficiente esta meta, constituye actualmente una tarea crítica.

En este trabajo, aplicamos una metodología para determinar las características más relevantes y discriminantes en el diagnóstico y predicción de AD, las cuales fueron utilizadas para construir modelos de clasificación basados en las máquinas de soporte vectorial (*Support Vector Machine, SVM*). Estas características, definidas en este trabajo como variables predictoras de los modelos SVM, contienen información de biomarcadores morfométricos y características demográficas y neuropsicológicas de los pacientes. Para la selección de las variables predictoras y el entrenamiento de los modelos, instancias de ensayos clínicos de AD fueron obtenidas del reconocido proyecto *The Alzheimer’s Disease Neuroimaging Initiative (ADNI)*. La selección fue realizada en varias etapas combinando métodos estadísticos como el método de “Mínima Redundancia y Máxima Relevancia” (*Minimum Redundancy Maximum Relevance, mRMR*) para la puntuación de las variables y los métodos de selección de variables basados en SVM. Finalmente, modelos SVM fueron entrenados considerando las variables seleccionadas. Estos modelos fueron entrenados para dos tareas principales: (1) diagnóstico de AD, es decir para la discriminación

entre sujetos mayores (sC) de sujetos con demencia del tipo AD (sAD); y (2) predicción de AD, para la discriminación de sujetos con MCI con riesgo de conversión a AD (sMCIc) de los sujetos con MCI sin riesgo de conversión (sMCIInc).

Los resultados obtenidos en esta investigación son prometedores. Por un lado, el subconjunto de variables seleccionadas tienen correlación con los resultados de investigaciones previas [Cui et al., 2011; Desikan et al., 2009; Heckemann et al., 2011]. Asimismo, en la etapa de test de los modelos SVM, utilizando los casos clínicos de ADNI y nuevos casos obtenidos de proyecto *The Open Access Series of Imaging Studies* (OASIS), los resultados de la clasificación confirman su eficiencia en el diagnóstico y predicción de AD. Nuestro modelo de diagnóstico fue capaz de discriminar sujetos sAD de los sujetos sC con una exactitud mayor al 99%. Asimismo, el modelo de diagnóstico fue capaz de discriminar los sujetos sMCIInc de los sMCIc con una exactitud superior al 94%.

1.2 La enfermedad de Alzheimer (AD)

La demencia es un desorden cerebral que afecta seriamente la habilidad de una persona para llevar a cabo sus actividades diarias. AD es la forma más común de demencia entre las personas mayores, afectando las partes del cerebro que controlan el pensamiento, la memoria y el lenguaje. AD es una enfermedad terminal y neurodegenerativa que progresa lentamente, para la cual actualmente no existe un tratamiento curativo. En España, afecta al 5-7% de las personas de más de 75 años, con aproximadamente 650.000 personas afectadas y más de 100.000 nuevos enfermos al año (Fuente: [Fundación Alzheimer de España](#), Octubre 2011). Para la estimación de 26 millones de personas con AD en todo el mundo se espera que se duplique en el 2030 y se cuadruplica en el 2050 [Brookmeyer et al., 2007; Jönsson et al., 2006]. En una reciente publicación de la Asociación de Alzheimer (Alzheimer's Association) [Association, 2012] se informó que en USA, AD es la sexta causa de muerte y la quinta causa de muerte en personas a partir de los 65 años; asimismo, mientras que la proporción de muertos debido a otras causas de muertes ha ido reduciéndose, la proporción de muertes debido a AD se ha incrementado considerablemente, ver figura 1.1. En la figura se observa que en USA, entre el 2000 y 2008, la proporción de muertes debido a ataques del corazón, accidentes cerebrovasculares, cáncer de próstata han decrecido en 13%, 20%, y 8%, respectivamente, mientras que la proporción de muertes debido a AD se ha incrementado en un 66%.

Esta enfermedad fue observada por primera vez por el médico alemán Alois Alzheimer en 1906 [Stelzmann et al., 1995]. Alzheimer observó en el tejido cerebral de una mujer que había

1. INTRODUCCIÓN Y ALCANCES

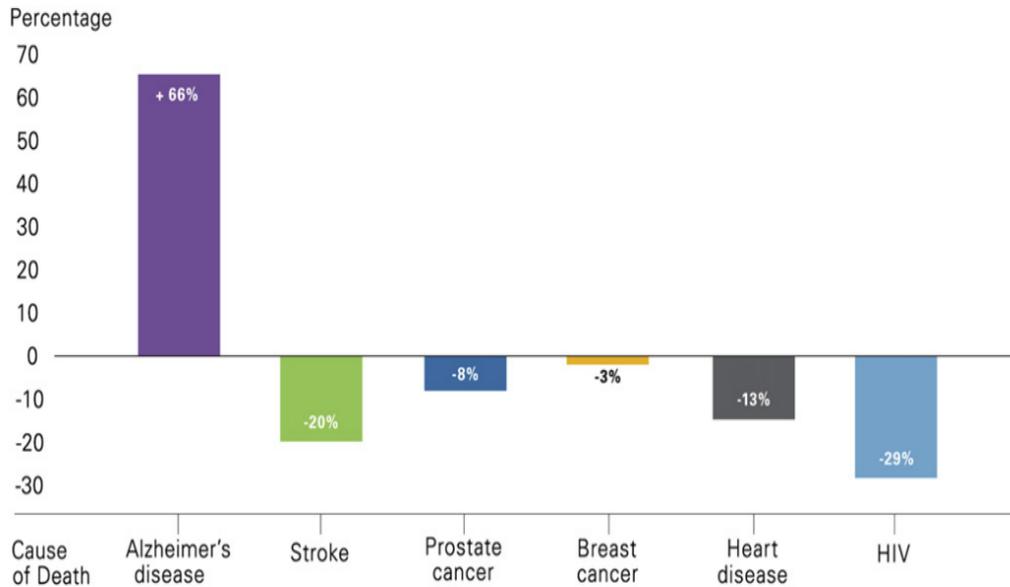


Figura 1.1: Incremento/Decrecimiento de las principales causas de muerte en USA entre el periodo 2000-2008. Fuente: [Association, 2012]

muerto de una extraña enfermedad mental, una disminución de las neuronas en el cortex cerebral junto con un aglutinamiento anormal de proteínas y una masa confusa de fibras en citoplasma de las neuronas [Maurer et al., 1997]. Actualmente ese aglutinamiento es conocido como “placas de amiloide” ($A\beta$) y la masa confusa son llamadas neurofibras de tau (τ)¹. Se decidió llamar a la nueva patología “Enfermedad de Alzheimer”² en honor a su “descubridor” y el término fue utilizado por primera vez por Kraepelin en la octava edición del “Manual de Psiquiatría” en 1910. Es así como la patología que en esas épocas era definida como una enfermedad rara, actualmente es considerada como una de las causas más frecuentes de demencia.

Los síntomas clínicos requeridos para el diagnóstico de AD son desarrollados sólo después de la pérdida de células importantes en las regiones cerebrales más vulnerables, asimismo, estos síntomas aparecen después de varios años e incluso décadas de iniciarse la enfermedad [Price et al., 2001; Sperling et al., 2011]. El diagnóstico clínico definitivo de AD requiere además de la presencia de un alto déficit cognitivo, la confirmación histológica (mediante autopsia) de la presencia de las proteínas τ y $A\beta$ en los tejidos cerebrales. Durante este diagnóstico son analizadas la densidad y distribución de las placas seniles y los ovillos neurofibrilares (*neurofibrillary*

¹Las neurofibras de τ fueron descritas por primera vez por Alois Alzheimer

²Esta enfermedad también es denominada mal de Alzheimer, o demencia senil de tipo Alzheimer o simplemente Alzheimer

tangles, NFTs) en el cerebro. [Morris et al., 1991], para las cuales, su componente principal representa el depósito anormal de las dos proteínas neuronales $A\beta$ y τ , respectivamente.

1.2.1 Fisiopatología de AD.

Existe un creciente consenso de investigaciones que afirman que la producción y acumulación de $A\beta$ (beta-amyloid) es crucial en la patogénesis de AD. Estas afirmaciones han dado lugar a la denominada hipótesis “*amyloid cascade*” [Hardy and Selkoe, 2002], la cual constituye actualmente la base de las opciones emergentes para el tratamiento de AD.

Resultados de autopsias y estudios basados en biomarcadores han reforzado esta hipótesis al proporcionar nueva información sobre el rol de la deposición de este amiloide en el cerebro, a través del tiempo [Cummings, 2011]. Asimismo, la formación de los ovillos de neurofibras, la oxidación y peroxidación lipídica, la inflamación y la activación de la muerte celular en “cascada”, han sido identificadas como consecuencias secundarias de la generación y deposición de $A\beta$ [Cummings, 2004]. En la figura 1.2 se muestra una representación de los procesos implicados en esta hipótesis basada en el rol de $A\beta$. Figura 1.2(a) presenta la secuencia (en cascada) de los procesos patogénicos implicados en la hipótesis propuesta originalmente por [Hardy and Selkoe, 2002]. Figura 1.2(b) [Cummings, 2004] nos presenta con mayor detalle estos procesos, en donde se observa que la secuencia de eventos patogénicos es iniciada con la generación de $A\beta$ a partir de la proteína amiloide precursora (amyloid precursor protein, APP) y progresa, a través de múltiples etapas, hasta la muerte celular.

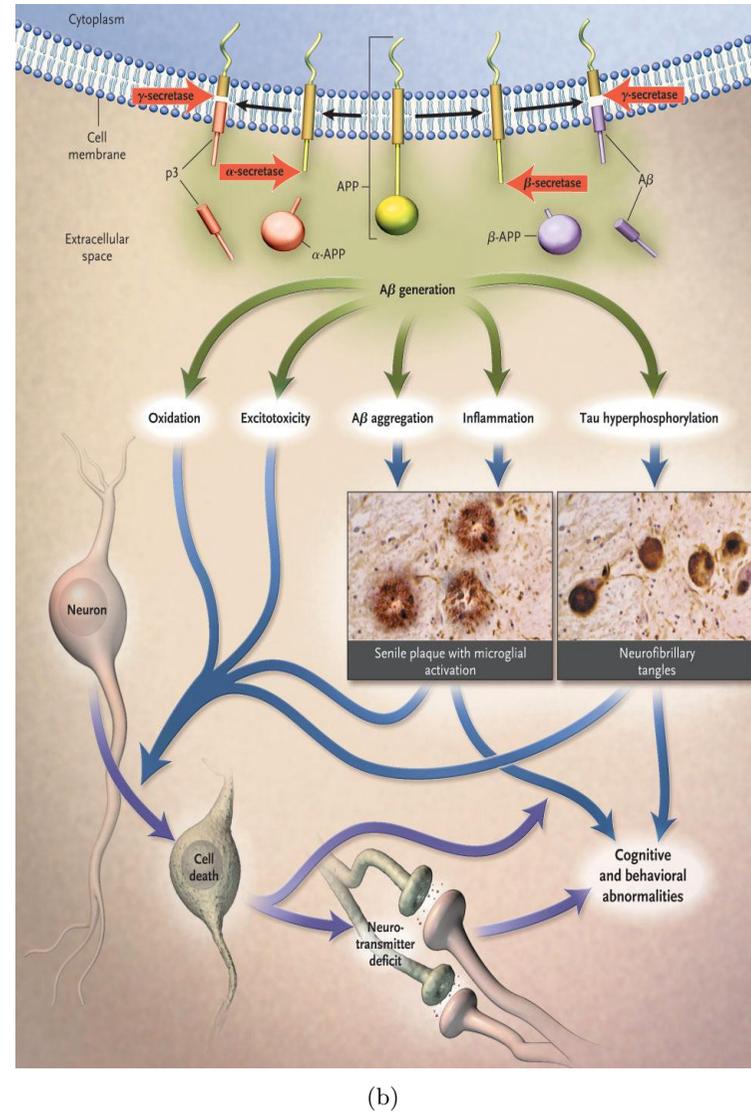
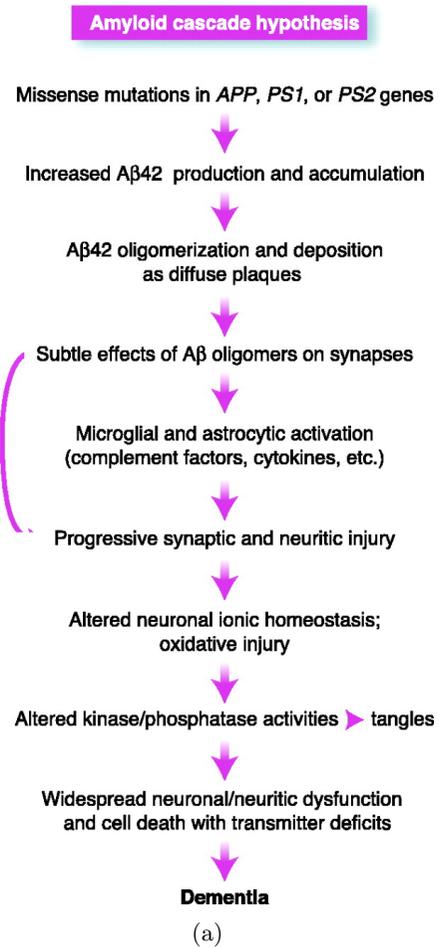


Figura 1.2: Hipótesis “*Amyloid cascade*” de la fisiopatología de AD. 1.2(a) Diagrama original de la hipótesis propuesto por [Hardy and Selkoe, 2002]. 1.2(a) Representación gráfica propuesta por [Cummings, 2004].

1.2.2 Progresión de $A\beta$ en el transcurso de AD.

Durante las etapas preclínicas de AD, el patrón de $A\beta$ en el cerebro presenta una distribución inicial difusa en la zona del neocórtex y progresa hacia la allocorteza temporal, incluyendo el hipocampo y las estructuras cerebrales subcorticales, durante el curso de la enfermedad. Durante las etapas clínicas del deterioro cognitivo, los depósitos de $A\beta$ son incrementados tanto en espacio y cantidad en todo el cerebro. Asimismo, en la etapa de demencia a causa de AD, los depósitos de $A\beta$ son más frecuentes en las regiones del cerebro neocortical, el hipocampo, la amígdala y las regiones subcorticales del cerebro (por ejemplo: el cuerpo estriado y los ganglios basales), sin embargo, estos depósitos son débilmente distribuidas en el cerebelo y en algunas zonas del tronco cerebral [Ewers et al., 2011; Thal et al., 2002].

En la figura 1.3 se presenta una ilustración presentada en [Thal et al., 2002] sobre la progresión de $A\beta$ en las tres diferentes etapas de AD. Obsérvese que en la figura, las etapas presentadas hacen referencia a la deposición de $A\beta$ en el cerebro y no necesariamente corresponden con la gravedad clínica del proceso. En la fase 1, el amiloide $A\beta$ está muy restringido a las áreas neocorticales del cerebro. En las fases 2 y 3, cuando los signos clínicos del deterioro cognitivo aún tienen que aparecer, los depósitos del amiloide son localizados a lo largo del neocórtex, la allocorteza (incluyendo el lóbulo temporal medio), así como en la corteza motora y las estructuras subcorticales. En las etapas clínicas de AD, fases 4 y 5, los depósitos del amiloide se observan a nivel global, incluyendo el tronco cerebral y el cerebelo en la etapa final de la enfermedad (fase 5)

1.3 Diagnóstico de AD.

Comúnmente, los médicos utilizan el término AD para referirse a un suceso clínico que típicamente se presenta con una característica de desorden amnésico progresivo seguido de la aparición de otros cambios cognitivos, conductuales y neuropsiquiátricos, los cuales afectan las actividades diarias de los afectados. Sin embargo, en la mayoría de los casos, los médicos hacen el diagnóstico de AD con mayor o menor grado de confianza, en base a su comprensión de la enfermedad y la evidencia existente para el diagnóstico [Cummings, 2004].

Los criterios establecidos por el estándar *The National Institute of Neurological and Communicative Disorders and Stroke*¹ y el *Alzheimer's Disease and Related Disorders Association*²

¹Actualmente conocido como el *National Institute of Neurological Disorders and Stroke*.

²Actualmente conocido como *Alzheimer's Association*.

1. INTRODUCCIÓN Y ALCANCES

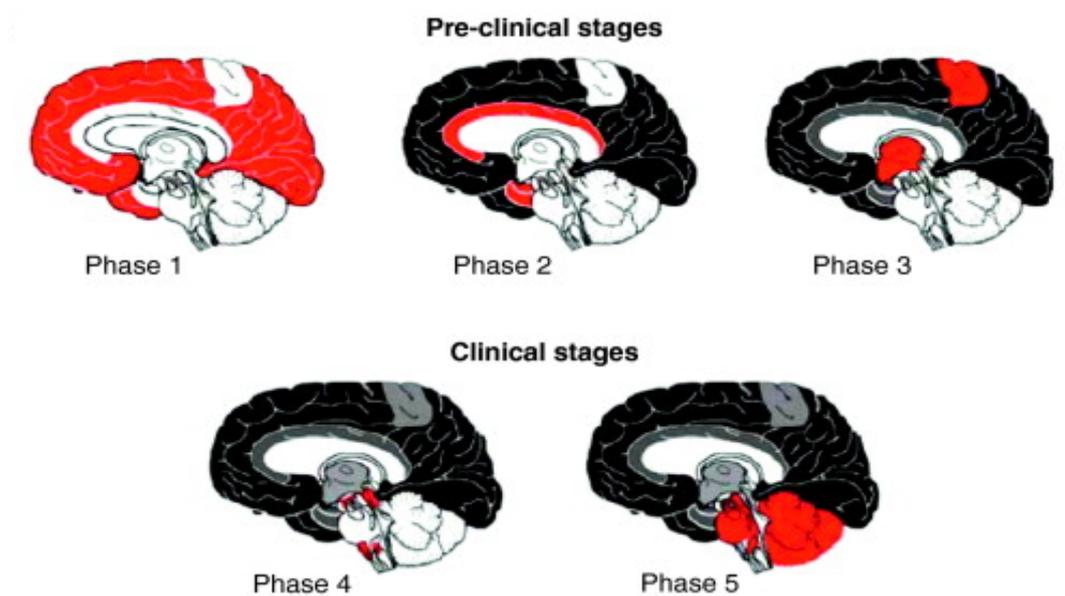


Figura 1.3: Fase de progresión de $A\beta$ en el cerebro. Fase 1: $A\beta$ está muy restringido a las áreas neocorticales del cerebro. Fases 2 y 3, los depósitos del amiloide son localizados a lo largo del neocórtex, la allocorteza y en la corteza motora y las estructuras subcorticales. Fases 4 y 5, los depósitos del amiloide se observan a nivel global, incluyendo el tronco cerebral y en cerebelo en la etapa final de la enfermedad (fase 5). Clave: blanco, no hay depósitos de $A\beta$; rojo, nueva deposición de $A\beta$; gris, depósitos de $A\beta$ que no estaban presentes en la fase 1; negro, depósitos de $A\beta$ que ya estaban presentes en la fase 1. Fuente: [Ewers et al., 2011; Thal et al., 2002]

(NINCDS-ADRDA) [McKhann et al., 1984] han constituido por más de 25 años, la base para el diagnóstico de AD.

Para hacer frente a la incertidumbre en el diagnóstico de esta enfermedad, el estándar NINCDS-ADRDA estableció la siguiente clasificación de los posibles diagnósticos de demencia ocasionada por AD: (1) “definitivo” para referirse al diagnóstico clínico con confirmación histológica post-muerte; (2) “probable”, en referencia al diagnóstico basado por la presencia de los síntomas clínicos típicos, sin confirmación histológica; y, (3) “posible”, basado en la presencia de características clínicas atípicas, pero sin aparente diagnóstico alternativo, ni confirmación histológica.

Para el diagnóstico clínico de posible o probable AD, se estableció que la presencia del deterioro cognitivo y la sospecha del síndrome de demencia debía ser confirmado a través de pruebas clínicas, tales como historial médico, exámenes clínicos, pruebas neuropsicológicas y mediciones de laboratorio; mientras que el diagnóstico definitivo de AD sólo podría ser hecho a partir de confirmación histológica (examinación microscópica del tejido cerebral) a través de autopsia o biopsia.

Adicionalmente, se resaltó que el diagnóstico “probable” o “posible” de demencia del tipo AD, sólo deberá efectuarse cuando el límite de severidad de demencia es alcanzado y cuando otras causas de demencia han sido excluidas a través de evaluaciones biológicas o por neuroimágenes.

Sin embargo, a pesar que el estándar ha demostrado una razonable especificidad¹ en el diagnóstico, en los últimos años se han hecho evidentes sus limitaciones, principalmente en su baja sensibilidad² en el diagnóstico durante etapas muy tempranas de la enfermedad [DeKosky et al., 2011], principalmente desde que el estado MCI empezó a ser referenciado. Asimismo, estas limitaciones se deben a que el estándar NINCDS-ADRDA no había sido revisado para incorporar los descubrimientos en la epidemiología y genética de esta enfermedad, así como los nuevos estudios clinopatológicos y el descubrimiento de potenciales biomarcadores de AD.

En la tabla 1.1, extraída del estudio realizado por [DeKosky et al., 2011], se presentan algunas de las limitaciones del estándar.

Revisión de los antiguos criterios de diagnóstico

Las deficiencias encontradas en los originales criterios de diagnóstico de AD, establecidos desde hace más de 25 años, así como los recientes avances en el uso de biomarcadores de AD, han

¹En el contexto del diagnóstico de AD, especificidad (*specificity*) se refiere a la capacidad de detectar la ausencia de AD en sujetos sanos.

²Sensibilidad (*sensitivity*) en este contexto se refiere a la capacidad de diagnosticar AD en pacientes realmente afectados por esta enfermedad.

1. INTRODUCCIÓN Y ALCANCES

Tabla 1.1: Limitaciones de los criterios establecidos en el estándar NINCDS-ADRDA en 1984. Fuente: [DeKosky et al., 2011]

Incertidumbre de la clasificación en el límite de - Demencia con enfermedades cerebrovasculares - Demencia con enfermedad de Parkinson, alucinaciones, etc. (por ejemplo, demencia con el síndrome de cuerpos de <i>Lewy</i>)
Problemas menores con - Afasias progresivas - Variación en el comportamiento por demencia frontotemporal
Casos muy leves que no cumplen los criterios de diagnóstico
No existen biomarcadores específicos de la enfermedad

propiciado el replanteamiento de estos criterios.

Iniciativa del 2007. En el 2007, Dubois y compañeros [Dubois et al., 2007], basados en los últimos avances en biomarcadores de AD, propusieron nuevos criterios de investigación, los cuales han actualizado el diagnóstico de AD en torno a dos factores: como un patrón específico de cambios cognitivos y como la evidencia estructural/biológica de la patología de Alzheimer. Asimismo, los autores propusieron un “léxico común” como punto de referencia para los clínicos e investigadores, el cual consiste en considerar a AD exclusivamente como una entidad clínica y sintomática que abarca ambas etapas de predemencia y demencia; y reformula las fases implicadas en AD. Para mayor detalle, ver [Dubois et al., 2007, 2010].

Iniciativa del 2009. Por otro lado, en junio del 2009, en una reunión realizada para revisar y actualizar los criterios de diagnóstico de AD implantados desde 1984, *The National Institute on Aging* (NIA) y *Alzheimer’s Association* conformaron tres grupos integrados por más de 40 expertos en el campo, a fin de revisar los actuales criterios. Un primer grupo fue asignado para revisar los criterios del estándar NINCDS-ADRDA para todos los casos de demencia y para el estado de demencia del tipo AD. El segundo grupo fue asignado para revisar la caracterización de la etapa MCI_{AD} y definir criterios para su diagnóstico. Asimismo, un tercer grupo fue constituido para revisar la caracterización de la fase AD-preclínico, centrándose en la revisión de los factores en individuos normales que podrían desarrollar disfunción cognitiva por causas de MCI y AD. Como resultado, en el año 2011 han surgido tres importantes publicaciones sobre: (1) Recomendaciones para el estado DAT (demencia ocasionada por AD) [McKhann et al., 2011]; (2) Recomendaciones sobre los criterios para el diagnóstico del estado MCI_{AD} [Albert et al., 2011]; y, (3) Recomendaciones para los criterios de diagnóstico de la etapa AD-preclínico [Sperling et al., 2011]. De este modo, los antiguos criterios establecidos para el diagnóstico de

AD en 1984 han sido reformulados para incluir los últimos avances obtenidos en los últimos 25 años.

A pesar de las diferencias presentes en las dos iniciativas, existe una formulación común en ambas. Ambos grupos de trabajo enfatizan en la necesidad de “validación” y “estandarización” de los nuevos criterios, antes de que éstos puedan ser transferidos dentro de un preciso y detallado diagnóstico clínico. Sin embargo, en cualquiera de los casos, ambos trabajos constituyen uno de los principales avances para transformar el concepto de AD desde ser considerado anteriormente como un síndrome clínico hacia ser considerado como un proceso patogénico progresivo [Zetterberg, 2011].

Las formulaciones más importantes de las reformulaciones establecidas por la NIA y *Alzheimer’s Association* son comentados en las siguientes secciones.

1.3.1 Etapas de AD.

A medida que las células nerviosas (neuronas) mueren y diferentes zonas del cerebro se atrofian, los pacientes con AD son caracterizados en su forma típica por una pérdida progresiva de la memoria y de otras capacidades mentales. Sin embargo, estudios han revelado que a pesar de que puedan pasar años para que estos primeros síntomas se observen en la personas afectadas, en las etapas iniciales de la enfermedad empiezan a producirse cambios fisiopatológicos que afectan a las neuronas hasta producirles la muerte.

Nota 1 *En este trabajo presentaremos los nuevos conceptos de esta enfermedad establecidos en el 2011 por el National Institute on Aging (NIA) y Alzheimer’s Association [Albert et al., 2011; McKhann et al., 2011; Sperling et al., 2011], los cuales redefinen las etapas de AD y establecen la reformulación de los antiguos criterios para el diagnóstico en cada una de éstas.*

Los nuevos criterios establecidos por el NIA y *Alzheimer’s Association* han definido las tres etapas de AD como: (1) AD-preclínico, (2) Deterioro cognitivo leve ocasionado por AD (MCI_{AD}), y (3) Demencia ocasionado por AD (*Dementia due to AD* o *Dementia of the Alzheimer’s Type*, DAT). Antes de esto, las etapas estaban definidas como: (1) “leve/temprano” (mild/early) AD, (2) “moderado/medio” (moderate/mid) AD, (3) “severo/tardío” (severe/late) AD. De este modo, los actuales criterios proponen que DAT inicia antes de la etapa “leve/temprano” y que cuando los cambios tempranos en el cerebro son identificados, un individuo diagnosticado con los nuevos criterios debería ser considerado en la etapa preclínica de AD o en la etapa MCI_{AD} [Association, 2012].

1. INTRODUCCIÓN Y ALCANCES

En la figura 1.4 se presenta un hipotético modelo de la trayectoria de transición clínica de AD¹, según la redefinición de las etapas. Obsérvese que el estado MCI_{AD} es precedido por el estado AD-preclínico y a su vez precede al estado de DAT [Sperling et al., 2011].

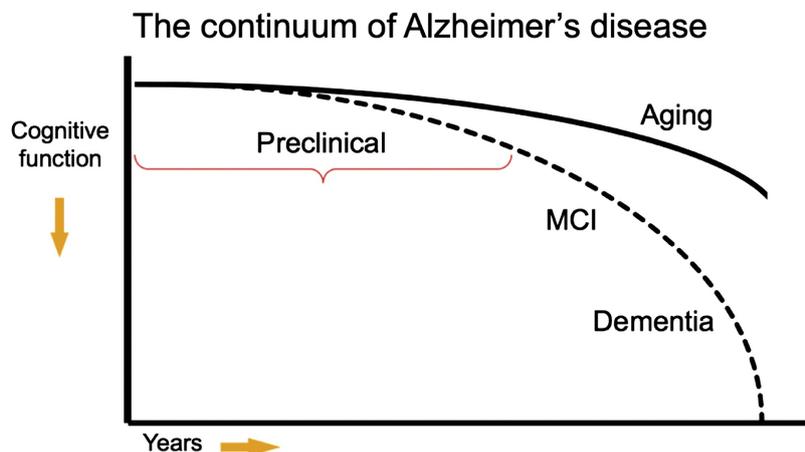


Figura 1.4: Modelo de la trayectoria clínica de AD. Fuente:[Sperling et al., 2011]

AD-preclínico.

En esta etapa las personas tienen cambios medibles en el cerebro, fluido cerebroespinal (CSF) y sangre que indican los primeros síntomas de la enfermedad pero aún no han desarrollado los síntomas como pérdida de memoria. Esta etapa preclínica o pre-sintomática se basa en las actuales afirmaciones que AD empieza a generar cambios en el cerebro varios años atrás (posiblemente entre 10 y 20 años) antes de la aparición de los primeros síntomas.

MCI ocasionado por AD (MCI_{AD}).

El interés en caracterizar los primeros signos clínicos del proceso degenerativo de AD, el cual podría evolucionar a DAT, ha dado lugar al desarrollo del concepto de “Deterioro Cognitivo Leve” (MCI), el cual representa el estado de transición entre el AD-preclínico y demencia del tipo AD [Morris et al., 2001].

Las personas con MCI tienen leve, pero medibles, cambios en las habilidades cognitivas (relacionadas con el pensamiento), sin embargo las habilidades individuales no se ven afectadas para llevar a cabo las actividades diarias [Association, 2012]. Asimismo, el rendimiento cognitivo

¹Este modelo también es conocido como “*The continuum of Alzheimer's disease*”.

se mantiene relativamente estable en los individuos sanos a medida que envejecen, pero disminuye rápidamente en las personas con MCI.

Dependiendo de qué dominios cognitivos hayan sido afectados, se pueden distinguir diferentes subtipos de MCI. El subtipo más relevante en relación a AD es el llamado “MCI ocasionado por AD” (MCI_{AD})¹. Evidencia significativa sugiere que en una persona con MCI_{AD} , el cerebro ya está cargado con marcadores de la enfermedad, incluyendo la pérdida irreversible de las neuronas [Price et al., 2001]. Sin embargo, es importante diferenciar que los sujetos en el estado MCI_{AD} no son dementes, pero tienen importantes deficiencias en uno o más dominios cognitivos y un mayor riesgo de demencia [Albert et al., 2011]. Asimismo, en este subtipo suele producirse un tipo de cambio anual hacia el estado DAT (demencia ocasionada por AD) de aproximadamente 12% y es caracterizado por la presencia de problemas en la memoria subjetiva y un deterioro leve de la memoria objetiva.

Estudios histológicos han encontrado que los cambios neuropatológicos entre las etapa MCI_{AD} y DAT son consistentes con la densidad y distribución de las neurofibras de τ y $A\beta$. Sin embargo, a pesar que el concepto de MCI es muy útil para identificar el grupo de sujetos con un alto riesgo de conversión a AD, es importante considerar que también incluye un número no despreciable de sujetos, en quienes la enfermedad nunca progresa hacia AD o en los que se desarrolla otro tipo de demencia [Morris et al., 2001].

Nota 2 *En la actualidad, los factores implicados en la progresión (o no progresión) del estado MCI_{AD} hacia el estado DAT no han sido definidos claramente por lo que se hace necesario el establecimiento de medidas adicionales para distinguir de manera fiable entre estas dos condiciones de MCI.*

Demencia ocasionada por AD (DAT).

Esta etapa es caracterizada por síntomas de deterioro de la memoria, cognición y la conducta; los cuales afectan las habilidades de las personas para la realización de sus actividades diarias.

En la figura 1.5 se ilustra los cambios producidos en la zona cortical, los ventrículos y el hipocampo durante las etapas MCI_{AD} y DAT (en su representación más severa). La figura también muestra las características de las células nerviosas en un sujeto sano y un sujeto DAT.

1.3.2 Diagnóstico en la etapa DAT.

Como resultado de los estudios del primer grupo liderado por Guy McKhann, se han redefinido los criterios clínicos establecidos en el estándar NINCDS-ADRDA para el diagnóstico del estado

¹Este estado también es llamado MCI amnésico

1. INTRODUCCIÓN Y ALCANCES

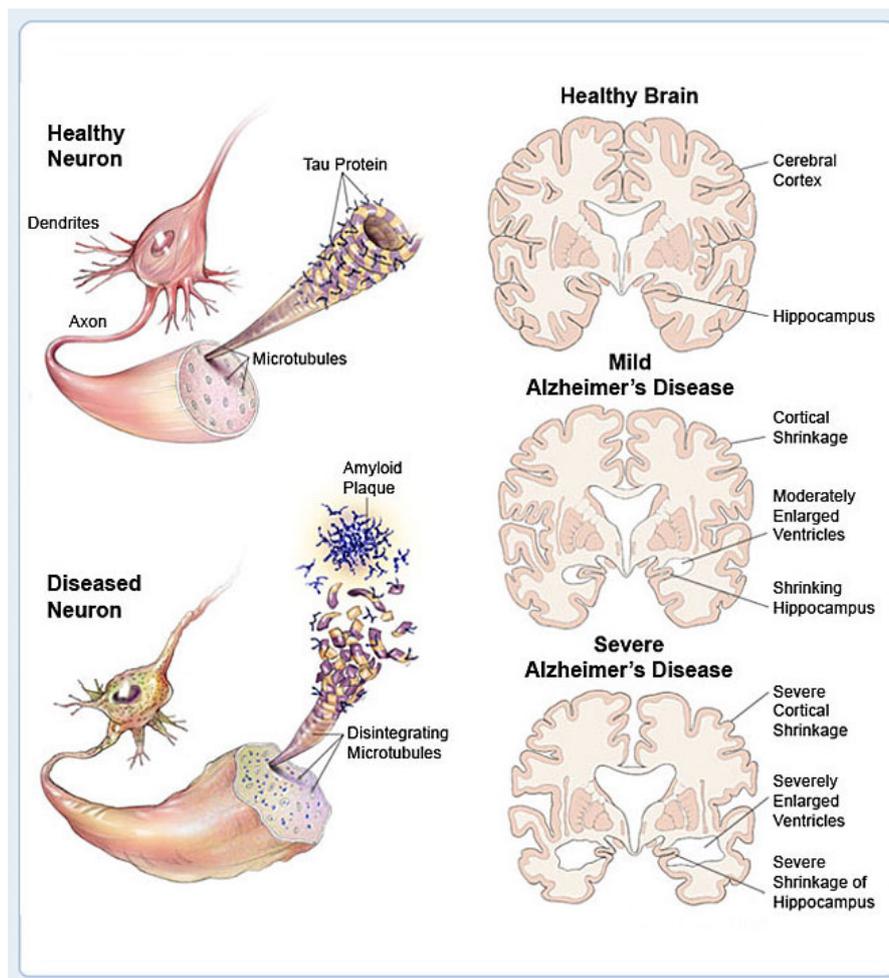


Figura 1.5: Ilustración del cambio del tejido cerebral y las células nerviosas en los estados: sano, MCI y AD severo. Fuente: *The American Health Assistance Foundation*.

DAT. Mayor detalle sobre los criterios originales pueden ser encontrados en [McKhann et al., 1984]; asimismo, los nuevos conceptos establecidos para el diagnóstico de esta etapa pueden ser encontrados con mayor detalle en [McKhann et al., 2011].

Algunos de los conceptos más importantes de la investigación de [McKhann et al., 2011] son resumidos a continuación.

Redefinición de la terminología de clasificación. Ha sido propuesta una terminología para la clasificación de los individuos con demencia ocasionada por AD, redefiniendo los antiguos criterios establecidos en el estándar NINCDS-ADRDA de 1984:

- **Probable demencia ocasionada por AD.** En esta terminología los criterios antiguos del estándar han sido conservados de manera general, ver [McKhann et al., 1984].
- **Posible demencia ocasionada por AD.** En este caso, los criterios del original estándar han sido redefinidos, por lo que se sugiere la re-evaluación de los pacientes.
- **Probable o posible demencia ocasionada por AD, con evidencia de procesos fisiopatológicos de AD.** Esta terminología es actualmente entendida para propósitos de investigación y es sustentada por los últimos avances encontrados en los biomarcadores de AD. En personas que cumplen los criterios clínicos para ser diagnosticados con probable o posible demencia AD, la evidencia de biomarcadores podría incrementar la certeza que la base del síndrome de demencia clínica presentado por estos individuos, es el proceso fisiopatológico de AD.

Uso de los biomarcadores para propósito de investigación de AD. Resaltan que los criterios clínicos continúan siendo la “piedra angular” del diagnóstico en la práctica clínica; sin embargo, Guy McKhann y compañeros sugieren que el uso de biomarcadores puede resultar útil para el diagnóstico basado en el proceso fisiopatológico de esta enfermedad, en tres circunstancias: (1) estudios de investigación, (2) ensayos clínicos; y como herramientas clínicas opcionales cuando estén disponibles y se considere apropiado por el médico.

El descubrimiento de biomarcadores de AD constituyen uno de los principales avances en la investigación sobre AD, llegando al punto de propiciar el re-planteamiento de los antiguos criterios de diagnóstico del estándar NINCDS-ADRDA de 1984. En la sección 1.4 hacemos una breve introducción a estos biomarcadores.

1. INTRODUCCIÓN Y ALCANCES

1.3.3 Diagnóstico en la etapa MCI_{AD} .

La investigación dirigida por Marilyn Albert, PhD. [Albert et al., 2011] estuvo orientada al desarrollo de criterios para la fase de pre-demencia sintomática de AD, denominada como “MCI ocasionado por AD” (“*MCI due to AD*”) (MCI_{AD})¹.

Nota 3 *Es importante resaltar que en los estudios de Albert et al. [2011], MCI_{AD} es descrita como una etiqueta de diagnóstico a ser utilizada en vista de la evidencia clínica y la evidencia de biomarcadores. Sin embargo, los autores señalan que este diagnóstico no implica que la patología de AD sea necesariamente la causa de los síntomas clínicos, o que necesariamente los sujetos en el estado MCI_{AD} progresen a un estado de demencia a causa de AD. Dado estas aclaraciones, el estado MCI_{AD} debe ser asumido como un concepto de investigación que aún no ha sido plenamente validado. Asimismo, los autores siguen confirmando que la demarcación clara entre un estado cognitivo normal y MCI_{AD} , así como, entre el estado MCI_{AD} y el estado DAT, continúa siendo una tarea difícil, por lo que es muy importante el juicio de los clínicos para hacer estas distinciones.*

Como resultado, se han establecido dos tipos de criterios de diagnóstico en esta etapa:

1. **Criterios clínicos para el diagnóstico de MCI_{AD} .** Estas recomendaciones incluyen los principales criterios a ser utilizados por los médicos sin acceso a técnicas avanzadas de imágenes médicas ni mediciones sobre el líquido cefalorraquídeo (CSF). Básicamente, estos criterios están orientados a:
 - Los cambios en la cognición, en comparación con previos niveles cognitivo de la persona;
 - El deterioro en uno o más dominios cognitivos de la persona, los cuales podrían ser: memoria, función ejecutiva, atención, lenguaje y habilidades visio-espaciales;
 - Preservación de independencia en habilidades funcionales; y,
 - No-demencia.
2. **Criterios para la investigación del estado MCI_{AD} .** Estos criterios han incorporado el uso de biomarcadores para el diagnóstico del estado MCI ocasionado por AD (basándose en imágenes médicas y mediciones del CSF), limitando el uso de éstos a entornos de investigación clínica, incluyendo los ensayos clínicos.

¹La etapa equivalente a MCI_{AD} según el nuevo léxico propuesto por Dubois y compañeros [Dubois et al., 2007] sería “prodromal AD”.

Nota 4 *Los nuevos criterios recomiendan la aplicación de herramientas de test basadas en biomarcadores en personas diagnosticadas con MCI para determinar si han sufrido cambios cerebrales que pudiesen ponerles en riesgo de desarrollar AD u otro tipo de demencia.*

1.3.4 Diagnóstico en la etapa de AD-preclínico

Se considera que el proceso fisiopatológico de AD, abreviado como AD-P (*AD-pathophysiological process*) [Sperling et al., 2011], es iniciado muchos años atrás, antes que la persona sea diagnosticada clínicamente con DAT. El inicio de este proceso está marcado dentro de la etapa “pre-sintomático” o “preclínica” de AD, ver figura 1.4.

En el estudio del grupo de trabajo dirigido por Reisa Sperling, MD [Sperling et al., 2011] se establecieron un conjunto de reformuladas recomendaciones para determinar los factores que predicen de mejor manera el riesgo de progresión de un estado cognitivo normal hacia los estados MCI_{AD} y DAT, siendo incorporados los actuales avances en la investigación de biomarcadores de AD. Asimismo, han sido redefinidos los conceptos, hasta la fecha utilizados, de las etapas tempranas de AD.

Importantes recomendaciones han sido establecidas a fin de diagnosticar ese grupo de sujetos “No-dementes, no MCI” incluidas en la etapa preclínica de AD. Sin embargo, estas recomendaciones han sido propuestas para uso exclusivo en el campo de la investigación, sin tener ninguna implicación clínica (por el momento). Algunas de estas recomendaciones más resaltantes son comentadas a continuación.

Redefinición de los principales sucesos en etapas tempranas de AD. Se ha establecido que la detección de los cambios patológicos deberían estar basados en la acumulación de la proteína $A\beta$; sin embargo este estudio ha resaltado que esta característica es “necesaria” pero no “suficiente” para producir las manifestaciones clínicas de AD.

Por otro lado, el estudio ha reconocido que algunos individuos pueden evidenciar todas las características neuropatológicas de AD durante la autopsia, pero nunca haber presentado demencia durante su vida. Asimismo, existen pacientes que a pesar de manifestar los síntomas clínicos, han vivido por mucho tiempo.

El estudio, justifica que estas excepciones, pueden haber sido propiciadas por los factores “reserva cerebral” y “reserva cognitiva”. “Reserva cerebral” se refiere a la capacidad del cerebro para soportar daños patológicos, tal vez debido a una mayor densidad sináptica o un mayor número de neuronas sanas, de tal manera que el sustrato neuronal sigue siendo suficiente para apoyar el funcionamiento normal del individuo. Por otro lado, la “reserva cognitiva” se refiere

1. INTRODUCCIÓN Y ALCANCES

a la capacidad de alternar las redes cerebrales para hacer frente a los efectos de la patología. Adicionalmente, se ha establecido la posibilidad que factores socio-demográficos como el alto nivel de educación o un alto nivel socio-económico contribuyan a ambas formas de reservas.

Inclusión de un modelo de biomarcadores para la etapa preclínica de AD

El modelo de biomarcadores propuesto inicialmente por [Jack et al., 2010a] ha sido modificado para referenciar la fase preclínica y el uso de biomarcadores, ver figura 1.6. El modelo refleja la trayectoria de cambios de un estado normal hacia anormal (eje Y) de los biomarcadores de AD (aceptados por el nuevo estudio) como una función de forma sigmoideal del estado de la enfermedad (eje X). Los biomarcadores incluidos en este modelo son: (1) Biomarcadores de la acumulación de $A\beta$; (2) biomarcadores de disfunción sináptica; y, (3) Biomarcadores de imágenes de resonancia magnética (MRI) centrados en la pérdida neuronal.

Obsérvese que en el modelo, ninguno de los biomarcadores es estático, es decir, las tasas de cambio de éstos varían de manera no-lineal con el tiempo (curva sigmoideal). Asimismo, tres características relacionadas con los biomarcadores son claramente diferenciadas :

- Los biomarcadores basados en la acumulación de $A\beta$ en el cerebro se vuelven anormales en la etapa AD preclínica y una sustancial carga de $A\beta$ es acumulada antes de la aparición de los síntomas clínicos¹. Asimismo, se puede observar que la acumulación de $A\beta$ en el cerebro es necesario pero no suficiente para producir los síntomas clínicos de las etapas MCI_{AD} y DAT , lo cual constituye uno de los nuevos conceptos establecidos por estos investigadores.
- Los biomarcadores de disfunción sináptica basados en *fluorodeoxyglucose 18F* (FDG) y la Resonancia Magnética Funcional (fMRI), pueden reflejar anomalías en etapas muy tempranas, particularmente sobre los portadores del gen $\epsilon 4$ de la apolipoproteína E (APOE). Estas anomalías funcionales pueden presentarse incluso antes de la detección de la acumulación $A\beta$ en el cerebro.
- Los biomarcadores de MRIs basados en el análisis de la pérdida neuronal, se vuelven anormales un poco más tarde. Asimismo, estos biomarcadores conservan una estrecha relación con el rendimiento cognitivo a través de la etapas clínicas MCI_{AD} y DAT .

¹La fase de latencia entre la acumulación de $A\beta$ y los síntomas clínicos aún no se ha cuantificado, pero algunos estudios sugieren que podría abarcar más de una década.

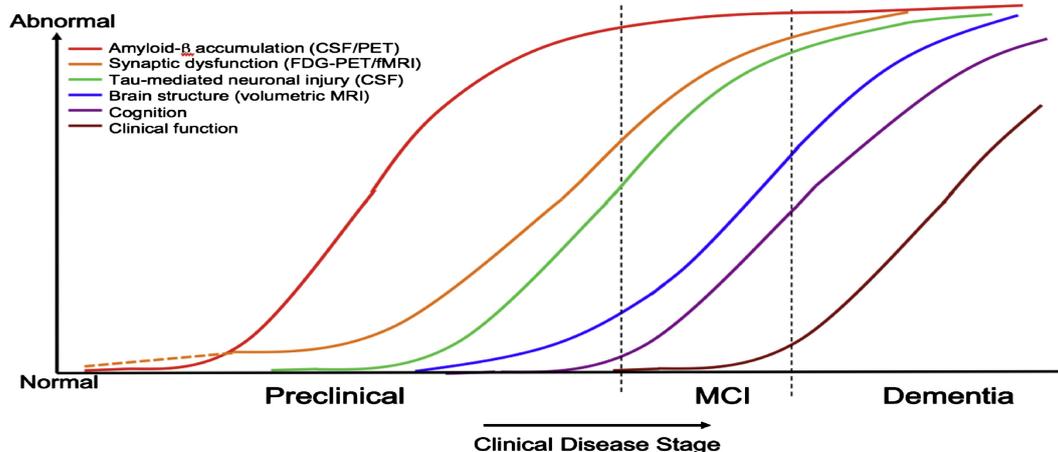


Figura 1.6: Modelo de las etapas de AD basado en biomarcadores. Fuentes:[[Sperling et al., 2011](#)] (adaptación),[[Jack et al., 2010a](#)] (original)

Nueva estructura de clasificación de la etapa AD-preclínico. Finalmente, se ha propuesto un nuevo esquema de clasificación basada en tres sub-etapas de la etapa AD-preclínico, las cuales sólo deberían ser aplicadas para propósitos de investigaciones relacionadas con AD y no como criterios de diagnóstico para propósitos clínicos. Las sub-etapas y su caracterización a partir de los biomarcadores incorporados son presentados en la tabla 1.2. Asimismo, el nuevo esquema de clasificación generado a partir de estas sub-etapas es presentada en la figura 1.7. Obsérvese que el esquema se basa en la hipótesis de que la acumulación de $A\beta$ en el cerebro (etapa amiloidosis cerebral) es actualmente la etapa más temprana de AD a ser medida y ocurre antes que alguna otra sintomatología cerebral, ver sección 1.2.1.

Tabla 1.2: Etapas categóricas para la investigación en AD preclínico. Fuente: [[Sperling et al., 2011](#)]

Etapas	Descripción	$A\beta$ (PET o CSF)	Biomarcadores de daño neuronal (τ , FDG, MRI estructural)	Evidencia de cambio cognitivo sutil
Etapa 1	Amiloidosis asintomática cerebral	Positivo	Negativo	Negativo
Etapa 2	Amiloidosis asintomática + neurodegeneración decreciente ("downstream")	Positivo	Positivo	Negativo
Etapa 3	Amiloidosis asintomática +daño neuronal + declinamiento sutil cognitivo/conductual	Positivo	Positivo	Positivo

1.3.5 Herramientas clínicas y neuropsicológicas para el diagnóstico clínico de AD

Como ha sido descrito en las secciones anteriores, la detección de AD en las etapas iniciales es especialmente importante para orientar óptimamente las terapias de la enfermedad, antes que el

1. INTRODUCCIÓN Y ALCANCES

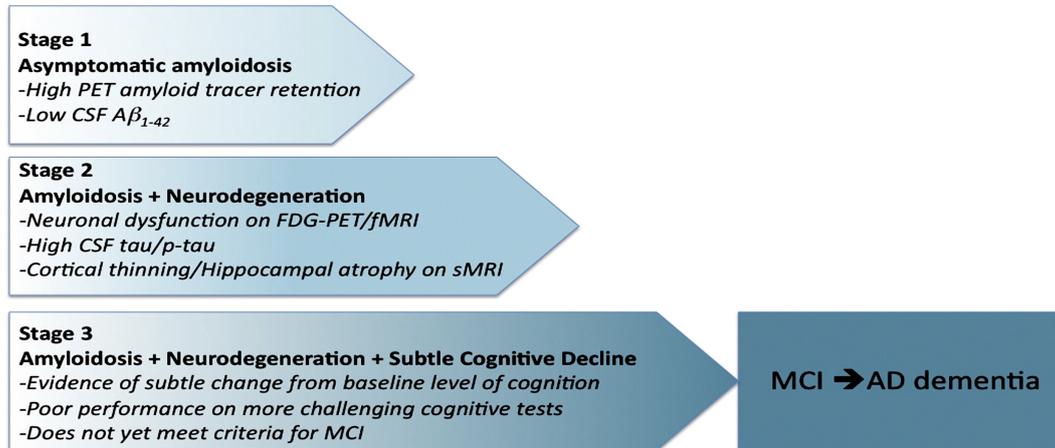


Figura 1.7: Esquema de clasificación de las subetapas del AD-preclínico. Fuente:[Sperling et al., 2011]

daño neuropatológico se haya producido y antes de la manifestación de la demencia. Debido a que los estudios indican que los individuos diagnosticados con MCI podrían presentar síntomas tempranos de AD [Morris et al., 2001; Petersen et al., 2010] y además se ha demostrado que la ocurrencia de cambios neuropatológicos entre la etapas MCI_{AD} y DAT son consistentes con la densidad y distribución de las neurofibras de τ y $A\beta$ [Hampel et al., 2004; JL and JC., 1999; Sperling et al., 2011; Tremblay et al., 2007], el diagnóstico temprano es crucial para establecer un tratamiento eficaz que conserve la función cognitiva en el nivel más alto posible.

Generalmente, el diagnóstico en las etapas clínicas MCI_{AD} y DAT y en menor escala la etapa preclínica de AD, incluyen varios tipos de evaluaciones basadas en exámenes clínicos y calificaciones neuropsicológicas. Estas evaluaciones son comúnmente utilizadas para controlar la progresión de la enfermedad o la eficacia de un tratamiento. En su forma más común, estas evaluaciones incluyen la aplicación de exámenes o pruebas, tales como:

Historial médico. Consiste en una entrevista o un cuestionario para identificar los problemas médicos anteriores, las dificultades en las actividades diarias y los medicamentos (recetas, vitaminas, suplementos y medicamentos de venta libre), entre otras cosas. El médico es informado de cualquier antecedente familiar de Alzheimer u otros problemas médicos relacionados. Asimismo, el médico dialoga con un miembro cercano de la familia para complementar la información, lo cual es importante para obtener una imagen completa de la historia médica de la persona.

Examen físico. Incluyen evaluaciones del oído, corazón y los pulmones, así como la temperatura, la presión arterial y las pulsaciones. El médico también puede preguntar sobre la dieta y

la nutrición , el uso de alcohol y los productos del tabaco.

Pruebas estándar de laboratorio. Incluyen pruebas de sangre y orina diseñadas para descartar otras condiciones posibles. En estas pruebas se miden la biometría hemática, la tiroides y la función del hígado; y los niveles de glucosa y otros indicadores basados en la sangre de la enfermedad. Un examen de la depresión también es realizado. En algunos casos, una pequeña muestra de fluido espinal puede ser recogida para la prueba.

Análisis de las imágenes del cerebro. Las imágenes médicas estructurales y funcionales del cerebro han sido utilizadas para excluir otros tipos de enfermedades que se puedan confundir con AD, mejorar la exactitud del diagnóstico clínico de AD y monitorear la progresión de la enfermedad y el efecto de los tratamiento ¹. Por un lado, las imágenes estructurales, tales como las imágenes basadas en Resonancia Magnética (MRI: *Magnetic Resonance Imaging*) y las imágenes de Tomografía Computarizada (CT: *Computed Tomography*), son utilizadas para descartar o confirmar que la razón de los síntomas se deba a la presencia de tumores cerebrales o coágulos de sangre en el cerebro. Asimismo, las imágenes funcionales como la Tomografía por Emisión de Positrones (PET: *Positron-Emission Tomography*) y la Tomografía de Emisión de Fotón Único (SPECT: *Single Photon Emission Tomography*), son utilizadas para observar el trabajo y la actividad de ciertas partes del cerebro. Sin embargo, con los últimos avances en ambas modalidades de neuroimagen, las investigaciones se han orientado a detectar marcadores biológicos o biomarcadores, los cuales, basados en el análisis de ciertas zonas del cerebro afectadas por AD, ayuden a la detección temprana y cambios asociados con MCI_{AD} y DAT, así como a predecir el grado de progresión de la enfermedad [[Hampel et al., 2011](#)].

Pruebas neuropsicológicas para el diagnóstico y monitoreo de AD. El objetivo de estas pruebas es caracterizar de manera más completa los tipos de síntomas cognitivos, a fin de caracterizar de mejor manera el deterioro cognitivo y obtener pistas sobre la causa subyacente de la enfermedad. Actualmente constituyen las herramientas más utilizadas por los especialistas para el diagnóstico clínico de un “probable” o “posible ” AD. En estas pruebas, una variedad de herramientas son aplicadas para evaluar: la memoria, la resolución de problemas, atención, visión, coordinación motora y el pensamiento abstracto, tales como la realización de cálculos sencillos por los pacientes. A continuación son listadas las pruebas neuropsicológicas más comunes:

¹Actualmente, se está tratando de determinar otras técnicas de imágenes cerebrales capaces de identificar los signos reveladores del AD temprano con la suficiente certeza para ser utilizados como herramientas de diagnóstico.

1. INTRODUCCIÓN Y ALCANCES

- **Examen del Estado Mental Mínimo (MMSE).** El Examen del Estado Mental Mínimo (MMSE: Mini-Mental State Examination)[[Folstein et al., 1975](#)] es uno de los test cognitivos más utilizados y consiste en una serie de preguntas que ponen a prueba una variedad de habilidades mentales comunes. Es el test cognitivo abreviado de mayor validez y difusión internacional. Su puntaje máximo es de 30 y, originariamente, el valor límite o de corte fue de 24; sin embargo, los resultados deben ser interpretados de acuerdo a diferentes criterios tales como la edad, el nivel educacional y otras variables personales.
- **Escala de Deterioro Global (GDS).** La escala GDS (*Global Deterioration Scale*), también llamada “Escala de Reisberg” (*Reisberg Scale*) [[Reisberg et al., 1982](#)] suele ser utilizada por algunos médicos para medir la progresión de AD. Esta escala está dividida en 7 etapas de habilidades: Etapa 1: Ausencia de alteración cognitiva (*no cognitive decline*); etapa 2: Disminución cognitiva muy leve (*very mild cognitive decline*); etapa 3: Defecto cognitivo leve (*mild cognitive decline*); etapa 4: Defecto cognitivo moderado (*moderate cognitive decline*); etapa 5: Defecto cognitivo moderado-grave (*moderately severe cognitive decline*); etapa 6: Defecto cognitivo grave (*severe cognitive decline*); y, etapa 7: Defecto cognitivo muy grave (*very severe cognitive decline*). La escala define cada etapa en términos operacionales y en base a un deterioro supuestamente homogéneo. Sin embargo, dado que la secuencia de aparición de los síntomas es a menudo variable, se ha argumentado que se podrían producir errores durante la inclusión de un paciente en alguna de las etapas. Si embargo, se trata de una de las escalas más completas, simples y útiles para la estimación de la severidad de la demencia.
- **Clasificación Clínica de la Demencia (CDR: Clinical Dementia Rating).** El factor CDR fue desarrollado por el *Memory and Aging Project at Washington University School of Medicine* en 1979 y desde entonces ha sido utilizado en investigaciones y diagnósticos clínicos para caracterizar el nivel cognitivo y desempeño funcional de pacientes con riesgo de padecer AD u otro desorden de demencia [[Morris, 1993](#)]. En el diagnóstico clínico, el CDR califica el grado de deterioro cognitivo, utilizando un algoritmo estándar y sus posibles valores pueden ser: CDR=0, no deterioro; CDR=0.5, muy leve; CDR=1, leve; CDR=2, moderado; y CDR=3, demencia severa.
- **Escala de Evaluación para la Enfermedad de Alzheimer (ADAS).** El ADAS (Alzheimer’s disease assessment scales) es un test breve que evalúa la disfunción cognitiva y conductual. El ADAS cognitivo incluye la evaluación de la memoria, lenguaje, orientación, praxias y visuconstrucción. Es altamente efectivo para discriminar pacientes

con AD de sujetos controles, y es considerablemente más sensible que el MMSE para las demencias leves. El efecto del nivel educacional no es estadísticamente significativo y las diferencias según la edad son de pequeña magnitud.

El diagnóstico a través de las pruebas clínicas y neuropsicológicas en la etapa MCI_{AD} o incluso en una de las etapas preclínicas de AD sin la presencia de síntomas, sigue siendo un continuo desafío para los investigadores en neurociencias. Sin embargo, aunque estas pruebas, reflejan un aspecto importante de la progresión de la enfermedad (sobre todo el deterioro funcional), presentan algunas desventajas muy marcadas como su baja fiabilidad test-retest y su incapacidad para distinguir entre los tratamientos que modifican la evolución de AD en el tiempo y los tratamientos que son puramente sintomáticos. En consecuencia, existen considerables esfuerzos para identificar nuevas herramientas de diagnóstico que podrían proporcionar información adicional de esta enfermedad a lo largo de todas las etapas de AD.

En los últimos años, estas nuevas herramientas han sido constituidas por diversos biomarcadores destinados al diagnóstico y predicción de AD. Las últimas investigaciones sobre esta enfermedad se han dedicado a proponer biomarcadores que sirvan como sustitutos en vivo para la detección de los cambios patológicos y, por tanto, sirvan para identificar con exactitud los individuos que se encuentren en las primeras etapas de AD [Morris et al., 2001; Mueller et al., 2005c]. En la sección 1.4 se introduce el concepto de biomarcadores de AD y se describen con mayor detalle los principales tipos.

1.4 Biomarcadores de AD.

De acuerdo con el Instituto Nacional de Salud (National Institute of Health, NIH), un biomarcador es definido como : “..A characteristic that is objectively measured and evaluated as an indicator of normal biologic processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention” [Ho et al., 2010].

Los marcadores biológicos o biomarcadores reflejan una variedad de características de la enfermedad y tienen un potencial uso como mediciones de pruebas, criterios de selección de sujetos para pruebas o como marcadores de los procesos de una enfermedad.

Clasificación. Existen diversas clasificaciones asignadas a los biomarcadores, una de ellas está basada en las etapas de la enfermedad bajo estudio:

1. Biomarcadores antecedentes, utilizados para identificar el riesgo de desarrollar una enfermedad;

1. INTRODUCCIÓN Y ALCANCES

2. Biomarcadores de diagnóstico, útiles en la identificación de la enfermedad; y,
3. Biomarcadores de pronóstico o predicción, útiles para predecir el futuro curso de la enfermedad, incluyendo la respuesta a un tratamiento.

Nota 5 *El desarrollo de biomarcadores de AD no sólo sirve de soporte a los clínicos para el reconocimiento de los síntomas de la enfermedad en sus tres etapas AD-preclínico, MCI_{AD} y DAT; además pueden ayudar a identificar la enfermedad antes de que la demencia u otros síntomas aparezcan. Ambas funcionalidades corresponden la base de los biomarcadores de diagnóstico y predicción de AD y su identificación constituyen uno de los principales objetivos asumidos en este trabajo.*

Utilización de los biomarcadores de AD. En [Fjell and Walhovd, 2011; Hampel et al., 2011] han sido descritas las diferentes formas de utilización de los biomarcadores en investigaciones y ensayos clínicos de AD, destacando las siguientes:

1. En ensayos clínicos, en combinación con pruebas clínicas y cognitivas, como criterios de exclusión/inclusión de pacientes;
2. Como “covariantes de referencia” o “predictores” en la futura conversión del estado MCI_{AD} hacia DAT; es decir, las mediciones iniciales de estos biomarcadores son incorporados como covariantes de referencia en los modelos lineales de estudios de AD, aumentando su poder estadístico para detectar los efectos de los tratamientos aplicados o marcando directrices en la determinación de la muestra necesaria para los estudios;
3. En la determinación y evaluación de los resultados (“*outcome measures*”) de estudios de intervención y de esta manera detectar modificaciones en los efectos de tratamientos, ya sea en las etapas MCI_{AD} o DAT, así como en etapas preclínicas. En estas etapas suele ser difícil demostrar los efectos de los tratamientos utilizando únicamente mediciones clínicas y neuropsicológicas;
4. Detección de efectos adversos como inflamación, reacciones inmunológicas, etc.
5. Mejorar nuestro entendimiento de la patogénesis de AD.

Principales tipos de biomarcadores de AD. Respecto a los biomarcadores utilizados en el diagnóstico y predicción de AD, encontramos tres clases de biomarcadores comúnmente utilizados:

-
1. Biomarcadores genéticos, los cuales son estudiados para definir las mutaciones de genes asociados al aumento de la producción de $A\beta$ [Bergem et al., 1997; Bertram and Tanzi, 2008; Davinelli et al., 2011; Rocchi et al., 2003; Tanzi and Bertram, 2005] y determinar la etapa de la enfermedad.
 2. Biomarcadores neuroquímicos, basados en el estudio de las concentraciones de las proteínas tau y $A\beta$ en el serum, el plasma y en el CSF [Davinelli et al., 2011; Hampel et al., 2004, 2008, 2010b; Irizarry, 2004; Mattsson et al., 2011; Teunissen et al., 2002; Zetterberg and Blennow, 2008; Zetterberg et al., 2010].
 3. Biomarcadores basados en imágenes médicas, los cuales analizan el daño neuronal a partir de dos principales modalidades de neuroimagen: (i) estructurales, como las imágenes de CT y las MRIs; y, (ii) funcionales, como las imágenes SPECT y PET. Las imágenes estructurales son utilizadas para obtener biomarcadores que evalúen los cambios del cerebro relacionados a AD. Las imágenes funcionales son utilizadas para obtener biomarcadores que describan la reducción de la función neuronal en etapas tempranas de la enfermedad, lo cual puede ser ocasionado por el deterioro de la funcionalidad neuronal o por la reducción de la densidad neuronal.

De los tipos biomarcadores mencionados, los biomarcadores basados en el análisis de CSF y las mediciones de atrofia cerebral derivados de las MRIs, constituyen los primeros candidatos de biomarcadores para el diagnóstico temprano y monitoreo de la progresión de AD [Albert et al., 2011; Dubois et al., 2007; Fjell and Walhovd, 2011; McKhann et al., 2011; Sperling et al., 2011]. Asimismo, en los últimos años, las investigaciones se han orientado a combinar los biomarcadores basados en neuroimagen con biomarcadores neuroquímicos y genéticos, e incluso con información clínica y neuropsicológica a fin de compensar las deficiencias de cada tipo de biomarcador y garantizar un diagnóstico temprano de AD más preciso y completo [Mattsson et al., 2012; Tosun et al., 2010; Vos et al., 2012].

1.4.1 Biomarcadores de AD basados en neuroimagen.

Como fue descrito en la sección 1.2.1, las características neuropatológicas básicas de AD están asociadas a los cambios en el cerebro como la acumulación de la proteína amiloide $A\beta$ (base de la hipótesis “*Amyloid Cascade*”) y el desarrollo de ovillos neurofibrilares, ambos asociados a la degeneración y los síntomas clínicos de demencia.

La detección de estos cambios cerebrales, incluso en etapas muy tempranas de la enfermedad, ha sido posible a través de la utilización de diferentes modalidades de imágenes médicas.

1. INTRODUCCIÓN Y ALCANCES

A continuación presentamos un resumen de las principales técnicas y métodos de imágenes neuronales comúnmente utilizadas para detectar “*in vivo*” los cambios del cerebro asociados a la neurodegeneración y el deterioro cognitivo en los sujetos afectados por AD.

PET. Esta modalidad de imagen es utilizada en combinación con agentes de amiloide “*in vivo*” basados en imágenes, los cuales “marcan” los depósitos fibrilares de $A\beta$ del cerebro. Esta combinación ha sido una valiosa técnica de visualización y cuantificación de la deposición de $A\beta$ en vida de los sujetos.

FDG-PET. Este método consiste en combinar las modalidades de imagen PET con el isótopo 18F-fluorodesoxiglucosa (FDG) (18F-fluorodeoxyglucose) a fin de detectar cambios en el metabolismo de la glucosa y el flujo sanguíneo en el cerebro. FDG es un análogo a la glucosa, muy utilizada en el diagnóstico por imagen, principalmente en combinación con las imágenes PET. Cuando el flúor de la molécula FDG es convertido en flúor-18, el cual es un isótopo radioactivo emisor de positrones, se obtiene la molécula 18F-FDG. De este modo, la FDG es inyectada en los pacientes y un escáner de PET es utilizado para formar imágenes de la distribución de la FDG a lo largo del cuerpo.

MRI estructural. Esta modalidad es utilizada para detectar cambios en el tejido de la GM y WM. Respecto a su uso en el diagnóstico de AD, esta técnica es particularmente muy sensible para detectar cambios en el volumen de GM debido a la atrofia cerebral que se produce por la pérdida neuronal.

DTI. Es una técnica de MRI utilizada para evaluar los cambios microestructurales del cerebro dentro de las zonas fibrosas de la WM.

Imágenes de Resonancia Magnética Funcional (fMRI). fMRI es un procedimiento basado en MRI que permite obtener imágenes de las regiones cerebrales y medir la actividad cerebral al detectar cambios asociados con determinadas tareas en el flujo sanguíneo. En su forma básica, fMRI utiliza el contraste *blood-oxygen-level-dependent* (BOLD) para medir y mapear el flujo sanguíneo (respuesta hemodinámica) y la oxigenación de la sangre, los cuales tienen correlación con los cambios en la actividad neuronal en una escala de tiempo de pocos segundos. En investigaciones de AD, fMRI ha sido utilizada para evaluar cambios en la actividad cerebral relacionados con tareas cognitivas.

Nota 6 *En este trabajo nos hemos centrado en el estudio de los biomarcadores de AD basados en neuroimagen, principalmente en aquéllos que proporcionan información morfológica del cerebro afectado por atrofia, obtenidos de MRIs. Este tipo de biomarcadores son referenciados en este estudio como “Biomarcadores morfológicos obtenidos de MRIs”. El motivo de considerar MRI y descartar PET, fMRI, DTI, etc., es que estas últimas modalidades son más invasivas, más costosas y menos accesibles; por lo tanto existe mayor probabilidad de aplicar nuestro estudio y utilizar nuestros resultados en la práctica clínica diaria.*

1.5 Objetivos

Actualmente, el diagnóstico definitivo de DAT (demencia del tipo AD) requiere además de la presencia de un alto déficit cognitivo, la confirmación histológica (mediante autopsia) de la presencia de las placas seniles y proteínas τ y $A\beta$ en los tejidos cerebrales. Sin embargo, durante el tiempo de vida de los pacientes, los clínicos pueden obtener un diagnóstico de “probable” o “posible” AD basado en su experiencia y la utilización de otras herramientas de diagnóstico. Este diagnóstico puede ser generado para las tres etapas de AD (AD-preclínico, MCI_{AD} y DAT) basándose en características clínicas, neuropsicológicas y la información proporcionada por biomarcadores específicos de AD.

El objetivo principal de este trabajo ha sido establecer una metodología de soporte al diagnóstico y predicción de AD, de manera automática, principalmente en sus etapas MCI_{AD} y DAT, utilizando la combinación de biomarcadores morfológicos con características neuropsicológicas y sociodemográficas de los pacientes. Esta metodología estuvo centrada en dos tareas principales:

1. La selección de las variables predictoras más discriminantes a partir de la combinación de biomarcadores morfológicos con características socio-demográficas y neuropsicológicas. Estas variables deberían ser útiles para la discriminación de casos clínicos en los tres estados: (1) Estado normal (generalmente personas mayores sanas); (2) Estado MCI_{AD} ; y (3) Estado DAT.
2. Entrenamiento de modelos de clasificación de casos clínicos basados en máquinas de soporte vectorial (*Support Vector Machine*, SVM). Los modelos SVM estarían enfocados a dos tareas principales: (1) Diagnóstico de AD mediante la discriminación entre sujetos sanos y sujetos DAT; y (2) Predicción de AD, orientada a la discriminación de sujetos MCI_{AD} con riesgo de progresión a AD y sujetos MCI_{AD} sin riesgo de progresión. Asimismo,

1. INTRODUCCIÓN Y ALCANCES

estos modelos deberían garantizar resultados aceptables con respecto a la sensibilidad y especificidad de las tareas de clasificación.

Adicionalmente, se establecieron objetivos secundarios que ayudarían a alcanzar el objetivo principal:

1. Obtención de grupos de entrenamiento, validación y test a partir ensayos clínicos de AD validados y con reconocimiento de las principales investigaciones realizadas en el área.
2. Establecimiento de las técnicas de selección de variables predictoras más utilizadas en el diagnóstico de demencia, principalmente AD.
3. Estudio de las principales características de SVM y definición de su correcta configuración para nuestros problemas de clasificación.
4. Establecimiento de las métricas estadísticas para calificar el desempeño de los modelos de clasificación obtenidos y de esta forma garantizar la obtención de los mejores resultados.
5. Verificación de la viabilidad del uso de la metodología propuesta en un entorno clínico. Para este fin se estableció el desarrollo de una herramienta computacional que facilite el análisis de las imágenes de Resonancia Magnética y la utilización de los modelos de clasificación obtenidos de manera automática y robusta.

En los capítulos 4 y 5 son presentadas las etapas, procesos y algoritmos implicados en esta metodología, así como los resultados obtenidos en este estudio.

Capítulo 2

Biomarcadores morfométricos obtenidos de MRI.

2.1 Introducción

Tradicionalmente, las técnicas de diagnóstico basada en imágenes, principalmente en imágenes estructurales, han sido utilizadas para excluir otro tipo de enfermedades del cerebro, cuyos síntomas pudiesen ser confundidos con AD (por ejemplo: tumores cerebrales o hematomas epidurales), ver sección 1.3.5. Sin embargo, con el incremento del potencial de estas técnicas y el advenimiento de nuevas modalidades de imágenes médicas, su uso no sólo ha estado enfocado a mejorar la exactitud del diagnóstico clínico de esta enfermedad, también se ha visto beneficiada la potencialidad de estas imágenes para monitorear el progreso de la enfermedad y los efectos de tratamientos. Diversos estudios han demostrado que debido a presencia de ciertos patrones en el cerebro, es factible diagnosticar utilizando imágenes de Resonancia Magnética (MRIs), si un paciente se encuentra en la etapa preclínica, MCI o demencia por AD [Desikan et al., 2009; Ewers et al., 2011; Killiany et al., 2000].

En esta nueva era de la imagenología médica aplicada al diagnóstico de AD, uno de los principales campos de estudio es la combinación de las MRIs con técnicas de rastreo manual y/o herramientas automáticas de extracción volumétrica para obtener regiones de interés (ROIs) del cerebro. Entonces, diferentes mediciones morfométricas son obtenidas sobre estos ROIs, dando lugar a los llamados biomarcadores morfométricos. Los biomarcadores morfométricos describen y cuantifican el grado de atrofia de ciertas zonas del cerebro afectadas por AD, ayudan a la detección temprana y cambios asociados a lo largo de sus etapas, así como a predecir el grado de progresión de la enfermedad [Hampel et al., 2011].

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

Asimismo, valoraciones del poder de análisis de los biomarcadores morfométricos basados en MRIs han demostrado su eficiencia frente a los test clínicos o neuropsicológicos en la identificación de pacientes en las primeras etapas de AD, incluso cuando los primeros síntomas clínicos no son visibles. Técnicas de análisis de MRIs detectan la atrofia cerebral en etapas donde el deterioro cognitivo es muy pequeño y posiblemente reversible, y se ha demostrado que MRI es superior a los biomarcadores basados en líquido céfalo raquídeo (CSF) en la predicción del declinamiento cognitivo [Fjell and Walhovd, 2011]. Frente a estas evidencias y otras investigaciones relacionadas con los biomarcadores genéticos, químicos, etc., el NIA y *Alzheimer's Association* lideraron la revisión de los criterios de diagnóstico de AD [Albert et al., 2011; McKhann et al., 2011; Sperling et al., 2011], comentados en la sección 1.3. Uno de los hechos más resaltantes de la actualización de estos criterios, ha sido la incorporación de los biomarcadores de AD, entre ellos los biomarcadores morfométricos basados en MRIs, dentro de los criterios de diagnóstico de AD para los campos de investigación y ensayos clínicos, sin descartar la posibilidad de ser incluidos como herramientas de diagnóstico clínico en un futuro no muy lejano [McKhann et al., 2011].

En la sección 1.4.1 describimos las modalidades y métodos de neuroimagen comúnmente utilizados para detectar “*in vivo*” los cambios del cerebro asociados con la neurodegeneración y el deterioro cognitivo a causa de AD. En este capítulo presentamos una breve descripción de los estudios orientados a determinar biomarcadores morfométricos basados en MRI para las diferentes etapas de AD¹ y los enfoques de clasificación automática de sujetos construidos a partir de éstos. Asimismo, son presentadas las principales técnicas de procesamiento y análisis de MRIs para la obtención de los biomarcadores morfométricos. En la nota 6 se explicó el motivo de la selección de MRI.

2.2 Descripción.

La información anatómica obtenida de las MRIs puede ser utilizada para diferentes propósitos. Una de las corrientes actuales de mayor auge es su aplicación para conocer el impacto de la neurodegeneración, trauma y otras condiciones de las estructuras macroscópicas del cerebro. Dentro de las diversas investigaciones están encaminadas a proponer y utilizar descriptores morfométricos basados en MRIs como biomarcadores de diagnóstico y monitorización de enfermedades del cerebro; un grupo significativo dedica sus esfuerzos en determinar la progresión entre las etapas AD-preclínico, MCI y AD. En las publicaciones [Clark et al., 2008; Fjell and

¹Para mayor detalle sobre los estudios de AD basados en otras modalidades de imágenes médicas, ver [Ewers et al., 2011; Hampel et al., 2008, 2010a].

Walhovd, 2011; Hampel et al., 2010a, 2011] se presenta una revisión general de las principales investigaciones realizadas en el uso de biomarcadores para el diagnóstico de AD, incluyendo los biomarcadores basados en MRIs.

Antes de describir las técnicas más utilizadas para la obtención de los biomarcadores morfométricos a partir de esta modalidad de imagen médica, es necesario listar algunas de las conocidas ventajas del uso de las MRIs en el diagnóstico de AD y las características deseables en estos biomarcadores:

Ventajas del uso de MRIs en AD.

- MRI es una técnica de imagen que produce claras reproducciones del cerebro y es beneficiosa para descartar otras causas de demencia, tales como tumores o derrames cerebrales;
- Es útil para mostrar el resultado de los cambios estructurales en el cerebro relacionadas con AD.
- Proporcionan una visualización macroscópica de la atrofia de los tejidos cerebrales ocasionada por los cambios neuronales producidos durante AD, incluso desde etapas muy tempranas. Los neurólogos realizan estudios longitudinales a través de ciertos periodos de tiempo para conocer la cantidad de tejido cerebral que se ha perdido desde la primera imagen obtenida del paciente en los inicios del tratamiento. Posteriormente, mediante mediciones correlativas de ciertas áreas del cerebro, el médico puede entender los cambios del cerebro afectado por esta enfermedad e informarle de la eficacia del tratamiento aplicado.
- Resultados coherentes e uniformes en estudios con imágenes médicas como las MRIs son más fáciles de aplicar y mantener en las instituciones médicas y de investigación, a diferencia de la variabilidad de ensayos clínicos necesarios para la medición directa de las proteínas $A\beta$ y τ , como biomarcadores clínicos.

Características deseables biomarcadores morfométricos.

Con el fin de utilizar la información proporcionada por las MRIs como biomarcadores morfométricos de diagnóstico, monitoreo y predicción de AD, en [Desikan et al., 2009; Ewers et al., 2011] se han descrito las principales características que deberían presentar estas mediciones. Algunas de las cuales son resumidas a continuación:

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

1. Detectar y cuantificar específicamente las características fundamentales de la patología de AD en personas con un riesgo elevado para tener demencia ocasionada por AD (es decir, en los individuos en estado MCI) y en pacientes con un diagnóstico clínico de AD confirmado [Desikan et al., 2009];
2. Demostrar una gran exactitud en la discriminación de las características normales de las personas con edad avanzada y lo sujetos con MCI y demencia AD [Desikan et al., 2009];
3. Presentar un alto grado de coherencia y de reproductibilidad en el *test-retest* de múltiples estudios independientes [Desikan et al., 2009];
4. Tener fuerte correlación con las mediciones clínicas del declinamiento cognitivo así como las mediciones de la invasión de la patología celular [Desikan et al., 2009];

2.3 Aplicación de los biomarcadores morfométricos en la predicción y diagnóstico de AD.

La atrofia o pérdida del volumen cerebral es una de las primeras consecuencias de la muerte neuronal ocurrida en el proceso degenerativo de AD [Killiany et al., 2000; Morra et al., 2009]. Cuando la atrofia es severa, ésta puede ser distinguida con mayor seguridad a través de una inspección visual simple de imágenes estructurales como las MRIs o las CTs [Clark et al., 2008]. Sin embargo, la clave en el diagnóstico de enfermedades neurodegenerativas como AD consiste en la identificación de las características y procesos relacionados con los patrones de atrofia que se producen en las etapas iniciales, antes de que el proceso degenerativo de AD se generalice.

Existen numerosos estudios sobre morfometría del cerebro, los cuales buscan describir las diferencias de las estructuras cerebrales y el grado de atrofia entre pacientes afectados en alguna de las etapas de AD a partir de las MRIs [Karas et al., 2004; Liu et al., 2011; Risacher et al., 2010; Walhovd et al., 2010; Westman et al., 2011]. Debido a las características fisiopatológicas de AD: formación de nudos en los lóbulos temporal medial (*medial temporal lobes*, MTL) y la ocurrencia temprana de problemas de memoria; MTL se ha constituido en el foco de estudios de neuroimagen de AD. Sin embargo, los avances en el análisis estructural del cerebro han hecho posible nuevos estudios que sirven como evidencia que la atrofia debido a AD puede ser identificada en diferentes áreas corticales y subcorticales del cerebro, incluso en etapas muy tempranas de la enfermedad [Fjell and Walhovd, 2011].

Con la finalidad de relacionar las zonas afectadas por atrofia y el uso de biomarcadores para su detección con las etapas de AD, hemos asumido el modelo hipotético de biomarcadores

propuesto por Ewers et al. [2011]. Este modelo representa la utilidad de los principales biomarcadores basados en neuroimágenes durante el curso de AD, ver figura 2.1. Obsérvese que en la figura son consideradas las modalidades de imágenes descritas en la sección 1.4.1. Con respecto a los biomarcadores basados en MRIs, representado como una línea verde en la figura, el modelo refleja que es posible detectar la atrofia de la materia gris en la etapa preclínica de AD; siendo identificada (de manera no exclusiva) en el hipocampo. La pérdida de volumen del hipocampo continúa su declive a través de la progresión de la enfermedad. Asimismo, una atrofia generalizada de las zonas corticales y subcorticales es observada en la etapa de demencia.

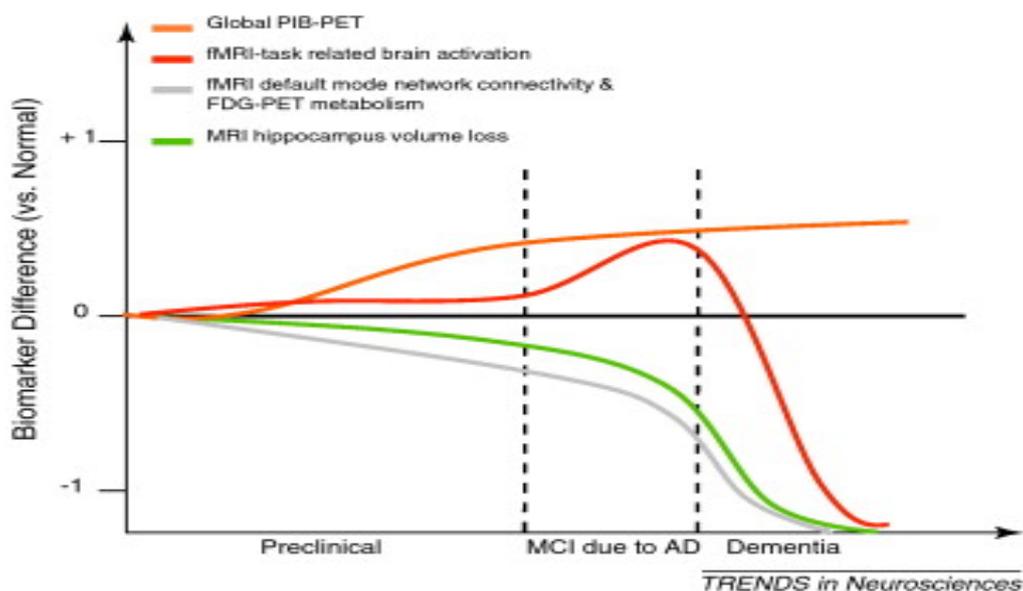


Figura 2.1: Modelo hipotético de los principales biomarcadores basados en neuroimágenes y su utilidad predictiva durante la etapas de AD. Fuente: [Ewers et al., 2011].

A continuación presentamos una breve descripción de los principales cambios estructurales de cerebro y la deposición de $A\beta$ observados en algunos estudios basados en MRIs. Los resultados de estos estudios han sido agrupados por cada etapa de AD.

2.3.1 Estudios de MRIs en la etapa AD-preclínica.

Existe evidencia que en la etapa preclínica de AD, la deposición de la proteína $A\beta$ en el cerebro está asociada a la atrofia de la GM; siendo los principales sitios de deposición temprana: las regiones neo-corticales, parietal y frontal. En consecuencia, altos grados de atrofia han sido encontrados en el hipocampo y la circunvolución del cíngulo [Ewers et al., 2011]. Sin embargo, a pesar de esta evidencia, la inclusión de estas investigaciones, la exacta definición de los niveles

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

patológicos de deposición de $A\beta$ en esta etapa, aún continúa siendo un importante tema de investigación.

2.3.2 Estudios de MRIs en la etapa MCI_{AD} .

La cuantificación de la morfometría cerebral basada en MRI puede ser utilizada para distinguir sujetos normales de control (sC) de sujetos en la etapa MCI (sMCI) o sujetos en la etapa de demencia debido a AD (sAD). Sin embargo, para ser considerado como biomarcadores tempranos, es importante que las MRIs cumplan dos tareas principales: (1) predecir cambios clínicos, esto es, distinguir los sujetos sMCI con riesgo de progresión a AD (sMCIc) de los sujetos sMCI estables o no convertido (sMCIinc); y, (2) predecir el declinamiento cognitivo.

Algunos estudios han demostrado que durante la etapa MCI_{DAT} , el incremento de los niveles de $A\beta$ en el cerebro o la reducción de este amiloide en el fluido cerebroespinal, están asociados con la atrofia de estructuras del MTL [Karas et al., 2004] como el hipocampo [Chupin et al., 2009] y la corteza entorrinal; y otras estructuras cerebrales del lóbulo parietal [Karas et al., 2004]. En el estudio de Heckemann et al. [2011], se comparó la morfometría de tres grupos de diagnóstico: sMCIinc, sMCIc y sAD; encontrándose diferencias altamente significativas en regiones individuales del cerebro. La mayoría de estas diferencias fueron ubicadas en el lóbulo temporal, una parte significativa en el lóbulo parietal y una asimetría incrementada en las regiones corticales posteriores.

[Karas et al., 2004] analizó patrones de pérdida de GM con la finalidad de determinar características propias del estado MCI_{DAT} y sus diferencias con el estado DAT. Su método consistió en obtener volúmenes locales y globales de la GM utilizando morfometría basada en vóxel (*voxel-based morphometry*, VBM) [Ashburner and Friston, 2000; Ridgway et al., 2008]. La aplicación de VBM sobre las MRIs de 33 sAD, 14 sC y 22 sMCI demostró que los sMCI presentan reducciones locales significativas en el lóbulo temporal medio, la corteza insular y el tálamo, en comparación con sujetos sC. El estudio también encontró que los sMCI presentan mayor cantidad de GM en el lóbulo parietal y en el cíngulo posterior y anterior, en comparación con los sAD; y que los sAD presentan mayor grado de atrofia en el lóbulo temporal medial. Asimismo, este estudio no encontró diferencias específicas del volumen de GM entre los sMCI y los otros dos grupos. Un estudio similar fue realizado por [Chételat et al., 2005].

2.3.3 Estudios de MRIs en la etapa de demencia AD.

En esta etapa, los biomarcadores morfométricos basados en MRI son utilizados principalmente para calificar los cambios de cantidad de GM en las diferentes estructuras cerebrales. Al respecto,

estudios han identificado cambios severos de volumen del lóbulo temporal medial, la cual podría incluir hasta una reducción del 20% del volumen del hipocampo; así como cambios pequeños y difícilmente visibles en zonas externas de este lóbulo. Por ejemplo, la reducción del volumen de los lóbulos temporales fue observada en los cerebros de pacientes con AD, de acuerdo a los estudios basados en VBM realizados por Chételat et al. [2005]; Karas et al. [2004]. La reducción del espesor cortical también ha sido observada en las superficies corticales, particularmente en los MTLs y la corteza entorrinal [Dickerson et al., 2009]. Otras de las áreas afectadas la constituyen el lóbulo temporal lateral y los lóbulos prefrontal y parietal.

Por otro lado, evaluaciones de la WM utilizando el método DTI (técnica basada en MRI para evaluar los cambios microestructurales de la materia blanca) han demostrado que las fibras que conectan el hipocampo con la circunvolución del cíngulo posterior son afectadas en sujetos con AD con una escala mucho mayor que en los sujetos mayores. Estos resultados sugieren que el daño de la WM podría tener relación con la atrofia de la GM [Ewers et al., 2011].

2.4 Obtención de los biomarcadores morfométricos.

Desde más de dos décadas atrás, las MRIs han sido utilizadas para obtener patrones o biomarcadores morfométricos *in vivo* que describen los cambios cerebrales producidos por AD, durante el curso de esta enfermedad. A partir de las numerosas investigaciones relacionadas con esta rama de las neurociencias, identificamos dos grupos claramente definidos:

1. El primero grupo está orientado a obtener y analizar mediciones manuales de ciertos ROIs del cerebro [Killiany et al., 2000]. Pérdida del volumen en regiones específicas, con el hipocampo y la corteza entorrinal, pueden ser detectadas *in vivo* utilizando técnicas manuales de extracción de ROIs. Sin embargo, estos métodos manuales constituyen una labor muy intensiva dado la complejidad característica de las estructuras cerebrales; en consecuencia estos métodos típicamente están centrados en unas pocas regiones definidas *a priori* [Dickerson et al., 2009].
2. El segundo grupo está orientado a calcular mediciones automatizadas basado en el enfoque *whole-brain* [Buckner, 2004; Desikan et al., 2006; Destrieux et al., 2010; Dickerson et al., 2009; Fischl et al., 2002; Heckemann et al., 2011; Hua et al., 2008; Klöppel et al., 2008].

Sin embargo, a pesar de que ambas técnicas ofrecen varias ventajas, su utilidad está sujeta a la obtención de biomarcadores de diagnóstico precisos y confiables. Por un lado, los métodos manuales de extracción de ROIs son inexactos, presentando variaciones entre diferentes sujetos,

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

y por otro lado, los métodos automáticos presentan el riesgo de disminuir su fiabilidad a la hora de aplicar el *test-retest*, debido a la imposibilidad de evaluar el estado de la enfermedad en un individuo particular.

Procesamiento y análisis automático de neuroimágenes.

La cuantificación morfométrica del cerebro para la obtención de biomarcadores de AD implica un conjunto de técnicas de procesamiento y análisis de imágenes. Las técnicas de procesamiento (basadas en técnicas de visión por ordenador) implican un conjunto de algoritmos que operan sobre la representación digital de estas imágenes con el fin de obtener ROIs y mejorar el proceso de obtención de información médica. Una vez procesadas estas imágenes, son aplicadas técnicas de análisis a fin de cuantificar los ROIs, obtener los biomarcadores y facilitar su posterior análisis. El progreso de estas técnicas de procesamiento y análisis de neuroimágenes, iniciado durante las pasadas décadas, permite obtener una mayor cantidad de información funcional y anatómica de cualquier individuo, con una enorme precisión en tiempo y en espacio. Sin embargo, en investigaciones y ensayos clínicos de AD, los cuales utilizan grandes bases de datos de imágenes, es imposible procesar y analizar directamente esta información, sin utilizar herramientas de cálculo que capturen los parámetros relevantes de soporte al diagnóstico médico.

2.4.1 Herramientas automáticas.

Para cubrir las deficiencias de los métodos manuales y garantizar las características deseadas en el uso de biomarcadores obtenidos de MRIs (ver sección 2.2), los continuos avances en algoritmos de análisis de imagen han facilitado el desarrollo de herramientas de software basadas en MRI. Estas herramientas son capaces de parcelizar automáticamente el cerebro en regiones anatómicas [Desikan et al., 2006; Fischl et al., 2002] e incluso cuantificar el grado de atrofia de ciertas regiones cerebrales de manera individual. A continuación listamos las principales herramientas utilizadas en el campo de AD.

- **FSL (*The FMRIB Software Library*)** es una herramienta de análisis de imágenes del cerebro para los tipos: MRI estructural, funcional y de difusión [Jenkinson et al., 2012]. FSL integra un conjunto de librerías escritas principalmente por el *Analysis Group*, FMRIB, *Oxford* y está libremente disponible en <http://www.fmrib.ox.ac.uk/fsl>. Principalmente incluye las librerías: BET (extracción del cerebro), FAST (segmentación de GM, WM y CSF). FIRST (segmentación subcortical), FSL-VBM (análisis de la densidad de la GM utilizando VBM) y SIENA-SIENEX (análisis de la atrofia cerebral en estudios

transversales y longitudinales); han sido utilizadas en numerosas investigaciones dedicadas a calificar el grado de atrofia del cerebro en sujetos MCI_{DAT} y DAT [Acosta et al., 2012; Bishop et al., 2011; Li et al., 2010].

- **Freesurfer.** Es un compendio de herramientas para el análisis de datos de neuroimágenes de código abierto, desarrollado por *Martinos Center for Biomedical Imaging* y libremente disponible en (<http://surfer.nmr.mgh.harvard.edu/fswiki>). Freesurfer integra diversos algoritmos para cuantificar las propiedades estructurales, funcionales y de conectividad del cerebro humano [Fischl, 2012]. Dentro de las principales funcionalidades de este software tenemos:
 - Segmentación volumétrica de las principales estructurales cerebrales macroscópicas [Fischl et al., 2002, 2004b].
 - Segmentación de los subcampos del hipocampo
 - Alineamiento inter-individual basado en los patrones de plegado cortical [Fischl et al., 1999a].
 - Parcelación de los patrones de plegado cortical [Desikan et al., 2006; Destrieux et al., 2010; Fischl et al., 2004b].
 - Mapeo del espesor de la corteza del cerebro humano [Fischl and Dale, 2000].
 - Reconstrucción de modelos de superficie de la corteza del cerebro humano [Dale et al., 1999; Fischl et al., 1999a].

La fiabilidad de las funcionalidades descritas, así como los detalles técnicos de los procedimientos de análisis de imagen realizados con Freesurfer han sido demostradas en anteriores investigaciones [Dale et al., 1999; Fischl, 2012; Fischl and Dale, 2000; Fischl et al., 1999a, 2002; Han et al., 2006; Jovicich et al., 2006; Segonne et al., 2004]. Asimismo, Freesurfer constituye una de las principales herramientas automáticas utilizada en investigaciones dedicadas a obtener biomarcadores morfométricos del cerebro para el diagnóstico y predicción en el curso de AD [Chételat et al., 2005; Cho et al., 2012; Chupin et al., 2009; Cuingnet et al., 2011; Desikan et al., 2009; Dickerson et al., 2009; Heckemann et al., 2011; Morra et al., 2009; Nestor et al., 2008; Salat et al., 2009].

- **SPM (*Statistical parametric mapping*).** Es un software de análisis de neuroimágenes funcionales como fMRI, PET, SPECT, EEG y MEG, basado en el método *Statistical parametric Mapping*. A pesar de que esta herramienta está dirigida a neuroimágenes

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

funcionales, en algunas investigaciones dedicadas a obtener biomarcadores morfométricos, ha sido utilizada en combinación con FSL o Freesurfer. Para mayor detalle visitar <http://www.fil.ion.ucl.ac.uk/spm>.

2.4.2 Principales Rutinas de Procesamiento y Análisis de MRIs.

En neurología (así como en otras ramas médicas), el procesamiento y análisis de imágenes del cerebro está orientado a corregir los problemas de calidad más comunes en este tipo de imágenes, obtener los ROIs y analizarlos a fin de obtener los biomarcadores morfométricos. Estas funcionalidades constituyen la base de las herramientas mencionadas en la sección 2.4.1.

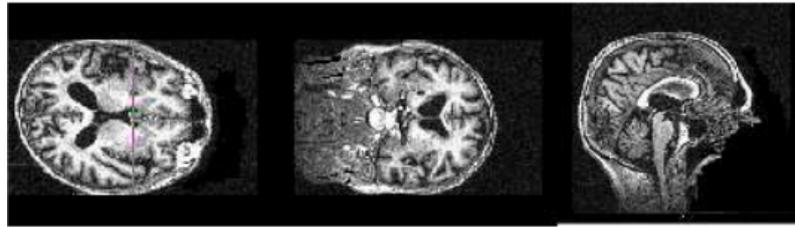
Con la finalidad de mencionar las principales técnicas aplicadas en la obtención de biomarcadores de AD, las hemos clasificado en tres etapas principales, cada una de las cuales agrupa procesamientos específicos:

1. Etapa de preprocesamiento.
 - (a) Eliminación de rasgos faciales.
 - (b) Alineamiento y corrección de las imágenes con respecto a un atlas del cerebro.
 - (c) Eliminación del bias.
2. Etapa de segmentación
 - (a) Eliminación de las estructuras no cerebrales.
3. Etapa de segmentación y cuantificación. Obtención de biomarcadores morfométricos.

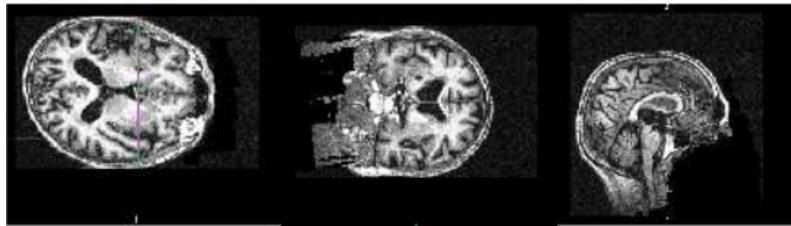
A continuación describimos brevemente estos procesamientos, presentando ejemplos de su aplicación sobre sesiones de MRIS.

2.4.2.1 Eliminación de rasgos faciales (*defacing*).

El objetivo principal de esta etapa es proteger la privacidad de los pacientes por lo que son localizadas las características faciales humanas (ojos, nariz, dientes, etc.) y removidas sin afectar el tejido cerebral. La figura 2.2 presenta un ejemplo de eliminación de rasgos faciales sobre sesiones de MRIs, cortesía del departamento de Biomedicina del *Centre Internacional de Mètodes Numèrics en Enginyeria* (CIMNE) (<http://www.cimne.upc.edu>), realizada con la rutina *mri_deface* del software Freesurfer (<http://surfer.nmr.mgh.harvard.edu/fswiki>).



(a)



(b)



(c)

Figura 2.2: Vistas ortogonales de tres sesiones de MRI (con cortos intervalos de tiempo) sin rasgos faciales. 2.2(a): sesión mpr-1. 2.2(b): sesión mpr-2. 2.2(c): sesión mpr-3. Fuente: Sesiones de MRIs del cerebro proporcionadas por el departamento de Biomedicina del *Centre Internacional de Mètodes Numèrics en Enginyeria* (CIMNE) (<http://www.cimne.upc.edu>) y procesadas con Freesurfer (<http://surfer.nmr.mgh.harvard.edu/fswiki>)

2.4.2.2 Alineamiento de multisesiones de MRIs y corrección respecto a un atlas del cerebro

En los estudios de imágenes realizados en investigaciones y ensayos clínicos (incluso podría extenderse en el diagnóstico clínico), es común obtener varias sesiones de imágenes (multisesiones), ya sea para un mismo sujeto, diferentes sujetos, o ambos casos. Las multisesiones de MRIs pueden aumentar la sensibilidad de los estudios o permitir efectuar conclusiones generalizadas a una población de estudio. Sin embargo, la adquisición de las multisesiones de MRIs no es perfecta por una serie de razones: factores físicos propios de la instrumentaria utilizada, la naturaleza imperfecta de esta modalidad de imagen, factores sujetos a las personas estudiadas y los intervalos de tiempos entre cada sesión que no garantizan la obtención de MRIs en el mismo sistema

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

de coordenadas. Por ejemplo, cortos movimientos de la cabeza en una escala menor a 1 mm podría significar la causa de importantes errores en un futuro análisis estadístico del cerebro, si estos no son corregidos a tiempo.

Como solución, dos procesamiento son necesarios para garantizar un correcto análisis estadístico: (1) alineamiento de todas las MRIs individuales de múltiples sesiones obtenidas en cortos intervalos de tiempo; y, (2) alineamiento de todas las sesiones obtenidas en tiempos más largos a un común espacio. Ambas tareas utilizan algoritmos de registro de imágenes. Después que todas las imágenes son alineadas, una variedad de métodos estadísticos pueden ser aplicados para combinar los resultados con respecto a las sesiones de MRIs o sujetos, ya sea para crear un único resultado para un grupo de sujetos o para comparar diferentes grupos de sujetos.

Alineamiento de todas las MRIs individuales de múltiples sesiones obtenidas en cortos intervalos de tiempo. El principal objetivo de este procedimiento es identificar pequeños cambios espaciales en las intensidades de cada vóxel de reconstrucción tridimensional del cerebro a lo largo de varias sesiones de imágenes obtenidas en cortos intervalos de tiempos. En la figura 2.3 se muestra un ejemplo de MRI después de aplicar una técnica de alineamiento sobre las tres sesiones de MRIs iniciales presentadas en la figura ; obsérvese que la resolución de esta imagen ha mejorado en comparación con las tres sesiones iniciales.

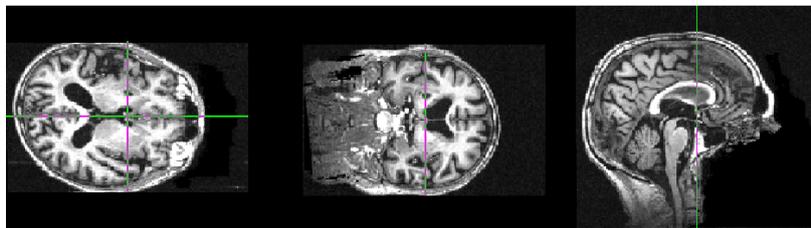


Figura 2.3: Vistas ortogonales de las MRIs resultantes después de aplicar las técnicas de alineamiento. Fuente: Sesiones de MRIs del cerebro proporcionadas por el departamento de Biomedicina del *Centre Internacional de Mètodes Numèrics en Enginyeria* (CIMNE) (<http://www.cimne.upc.edu>) y procesadas con Freesurfer (<http://surfer.nmr.mgh.harvard.edu/fswiki>)

Alineamiento de todas las sesiones MRIs a un común espacio. Esta última etapa implica técnicas de corrección de las MRIs a una “maqueta” o atlas del cerebro. De esta manera, las MRIs son transformadas a un “sistema espacial de cerebro estándar” como el sistema de coordenadas propuesto por [Talairach and Tournoux, 1988]. Por un lado, las maquetas del cerebro representan “un promedio” de muchos cerebros, los cuales han sido registrados en un común sistema de coordenadas. Por otro lado, un atlas, además de haber sido construido a

partir de varias sesiones de MRIs registrados en un común sistema de coordenadas, adicionan información más sofisticada de cada vóxel del cerebro; por ejemplo: tipo de tejido, estructura cerebral, área funcional, etc. En la figura 2.4 se presentan las vistas ortogonales de la sesión de MRI, después de aplicar corrección con el sistema de coordenadas espacial del atlas propuesto por *Talairach* [Talairach and Tournoux, 1988]. Obsérvese que la ubicación espacial de la imagen ha cambiado con respecto a la ubicación espacial presentada en la figura 2.3.

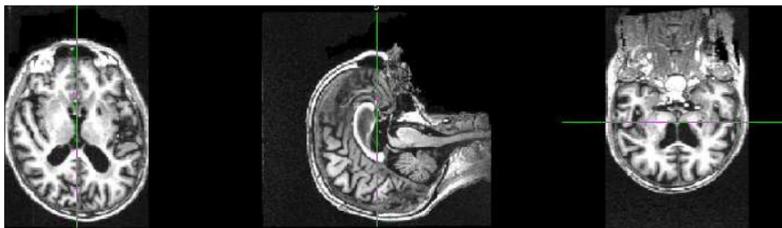


Figura 2.4: Vistas ortogonales de la sesión de MRI corregida con respecto a un sistema de coordenadas del atlas Talairach. Fuente: Sesiones de MRIs del cerebro proporcionadas por el departamento de Biomedicina del *Centre Internacional de Mètodes Numèrics en Enginyeria* (CIMNE) (<http://www.cimne.upc.edu>) y procesadas con Freesurfer (<http://surfer.nmr.mgh.harvard.edu/fswiki>)

2.4.2.3 Eliminación del Bias.

A menudo las MRIs se ven afectadas por variaciones de intensidad conocidas como “*bias*” presentes en diferentes zonas de las imágenes, las cuales son causadas principalmente por inhomogeneidades del campo de radiofrecuencia. La segmentación y el reconocimiento exacto de los tejidos puede verse comprometido por estas intensidades, al existir variaciones locales dentro de éstos. Por este motivo es recomendable aplicar técnicas de eliminación del *bias* antes de proceder con la segmentación u otros análisis estructurales.

2.4.2.4 Eliminación de las estructuras no cerebrales.

En los estudios de AD, la mayoría de los análisis realizados tienen como área de interés el cerebro y nos los otros tejidos de la cabeza (cuero cabelludo, cráneo, glóbulos oculares, etc.), por lo tanto resulta útil eliminar estas estructuras antes del análisis estructural. Es posible eliminar estas estructuras manualmente utilizando herramientas de software gráficas; sin embargo, este delimitado manual puede implicar una considerable cantidad de tiempo. Como solución alternativa, existen métodos automáticos y semiautomáticos, los cuales se basan en ubicar la capa alrededor del cerebro, caracterizada por una intensidad (generalmente un color en la escala de grises) significativamente diferente al tejido cerebral. Este cambio de intensidad es utilizado como una

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

barrera que delimita la superficie estimada del cerebro. En la figura 2.5 se presenta el resultado de aplicar un algoritmo de extracción de estructuras no-cerebrales en las MRIs presentada en la figura 2.4.

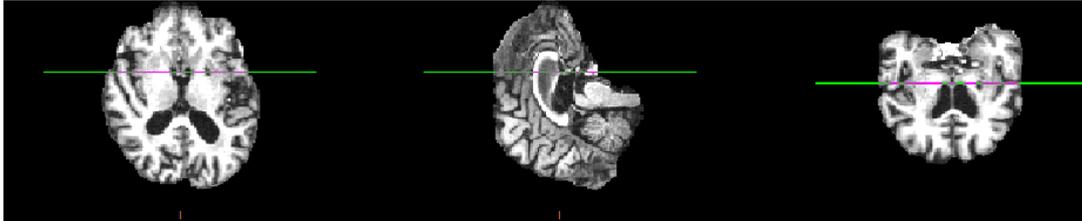


Figura 2.5: Vistas ortogonales de la imagen cerebral obtenida después de extraer las estructuras no-cerebrales.

2.4.2.5 Cuantificación Morfométrica.

Como ha sido mencionado con anterioridad, los desórdenes degenerativos como AD están asociados frecuentemente a cambios estructurales del cerebro. Los cambios morfométricos pueden incluir variaciones en el volumen o la forma de las regiones subcorticales, así como alteraciones en el espesor, el área y los patrones del plegado de la corteza cerebral [Fischl et al., 2002]. Para cuantificar estos cambios morfométricos existen dos tipos de análisis: (1) Análisis basado en la superficie y (2) Análisis basado en el volumen.

Análisis basado en superficie. Esta técnica depende de la posición y orientación de la *ribbon* cortical y es utilizado para calificar la variabilidad de la corteza cerebral. Este análisis es llevado a cabo a partir de la reconstrucción de modelos tridimensionales de la superficie de la WM y la superficie cerebral (*pial surface*); y el cálculo de la distancia entre ambas superficies. El cálculo de esta distancia implica el promedio de la distancia más corta desde cada punto de la superficie de la WM hasta la superficie pial y desde cada punto de la superficie pial hasta la superficie de la WM.

Análisis basado en volumen. Implica técnicas volumétricas utilizadas para reconstruir y detectar cambios en las estructuras corticales y subcorticales. Utilizando la segmentación “*whole-brain*”, una etiqueta neuroanatómica es asignada a cada vóxel de la reconstrucción tridimensional del cerebro.

2.5 Enfoques de clasificación basados en biomarcadores morfométricos

En la sección 2.3 se comentaron algunos estudios basados en morfometría cerebral para la obtención de biomarcadores de AD a partir de MRIs, los cuales aplicaron las técnicas de procesamiento y análisis de MRIs comentadas en la sección 2.4.2. Si bien estos estudios reportaron significativos resultados para determinar las zonas del cerebro afectadas durante el curso de esta enfermedad, incluso en etapas muy tempranas (sin la presencia de los síntomas clínicos), la mayoría de estas investigaciones no han proporcionado herramientas o modelos automáticos de clasificación para el diagnóstico y predicción de AD. Sin embargo, es importante resaltar que sus resultados han sido significativos para estudios posteriores sobre clasificación de pacientes con AD, debido a que los biomarcadores morfométricos propuestos son utilizados como características predictoras para tareas de clasificación.

Los biomarcadores morfométricos, asumidos como patrones de la atrofia cerebral, son complejos y altamente variables; y dependen de las etapas de AD. Asimismo, son concurrentes con los cambios estructurales relacionados a la edad y no son exclusivos de AD. En consecuencia, es necesario establecer métodos para identificar los biomarcadores o patrones más discriminativos de la enfermedad, obtenerlos de manera automática y utilizarlos en el diagnóstico y predicción de AD mediante herramientas de clasificación capaces de distinguir automáticamente los sujetos en la etapa DAT (sAD), sujetos MCI_{AD} (sMCI) y sujetos sanos de control (sC). Asimismo, estas herramientas deben ser capaces de discriminar entre los sujetos con MCI estable (sMCIc), es decir sin riesgo de progresión a AD; de aquellos sujetos con riesgo de conversión a AD (sMCIc).

Cuingnet et al. [2011] categorizó los enfoques de clasificación automática de pacientes basándose en el tipo de características extraídas de las MRIs. Tres diferentes categorías fueron establecidas: (1) Enfoque basadas en vóxeles, (2) basado en vértices y (3) basado en ROIs. Estas categorías son descritas en las siguientes secciones.

2.5.1 Clasificación basada en *vóxel*.

En esta categoría, las características o patrones discriminantes de AD (biomarcadores morfométricos) son definidos en el nivel de vóxel MRI. Específicamente, estas características definen la probabilidad de las diferentes clases de tejidos, como la GM, WM y el CSF, en un vóxel específico [Klöppel et al., 2008; Vemuri et al., 2008].

[Vemuri et al., 2008] desarrolló una herramienta de diagnóstico de AD utilizando máquinas de soporte vectorial (SVM) para la clasificación de dos tipos de sujetos sC y sAD. SVM fue

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

utilizado para la selección de los patrones o variables predictoras más discriminativas entre ambos tipos de sujetos. Densidades de vóxeles isotrópicos de zonas de la GM, WM y el CSF (obtenidas sobre el volumen de MRIs), características demográficas (edad y sexo) y mediciones de la apolipoproteína E (APOE); fueron obtenidas de 190 pacientes sAD (con probable AD) y 190 sujetos sC. Tres modelos de clasificación basados en SVM fueron entrenados: Modelo I basado en las densidades del tejido; Modelo II, densidades con características demográficas como covariante del modelo; y Modelo III, densidades con características demográficas y mediciones de APOE; las dos últimas características fueron utilizadas como covariantes del modelo. Como resultado del Modelo I, un vector de 237 variables basado en patrones anatómicos (vóxeles de la GM(124), WM (45) y CSF (68)) fueron seleccionadas como más discriminativas entre sC y sAD. Estos patrones estuvieron localizados en el lóbulo temporal medial-basal, ver figura 2.6. La figura presenta una muestra de MRI-T1 cuyos vóxeles han sido marcados de acuerdo a una escala de color. La escala indica la frecuencia de ocurrencia de cada vóxel en múltiples mapas de tejidos: Amarillo, vóxeles utilizados en los tres tejidos GM, WM y CSF; naranja, vóxeles localizados para los tejidos de la WM y CSF; y rojo, vóxeles localizados para un solo tejido (GM, WM o CSF). Asimismo, durante la etapa de test de los tres modelos, el modelo I obtuvo sensibilidad=86% y especificidad=86%; el modelo 2 obtuvo sensibilidad=88% y especificidad=90%; y el modelo 3 obtuvo sensibilidad=86% y especificidad=92.%

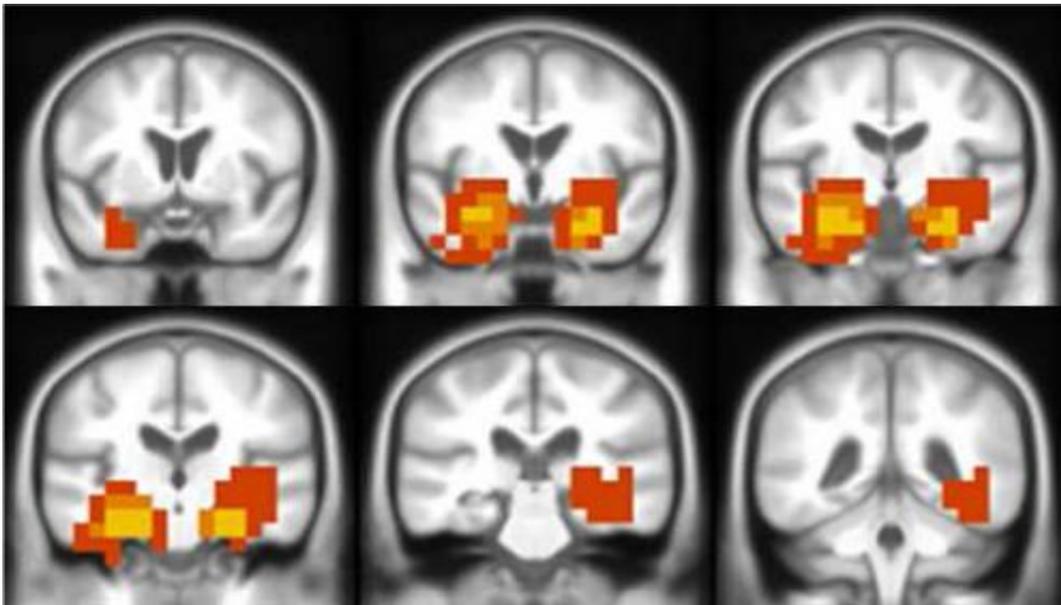


Figura 2.6: Patrones anatómicos con máximo poder de discriminación entre sujetos sAD y sC localizados con modelo de clasificación SVM obtenido por [Ewers et al., 2011].

[Klöppel et al., 2008] utilizó SVM para construir modelos de clasificación entre sujetos sC y sAD sobre patrones obtenidos de MRIs. La selección de las variables (representación de los patrones) más relevantes fue realizada con SVM. Este estudio demostró que las SVMs pueden clasificar con gran éxito los sC y sAD con un 89% de exactitud y tienen la ventaja de constituir modelos robustos, los cuales pueden ser aplicados en diferentes centros clínicos. Finalmente, [Klöppel et al., 2008] sugirió que las SVM juegan un rol importante en la práctica clínica.

2.5.2 Clasificación basada en vértices (enfoque cortical).

En la segunda categoría, los patrones de AD representan valores del espesor cortical calculados sobre cada vértice de la superficie cortical. El espesor cortical se ha constituido en un índice directo de atrofia cerebral y como un potencial candidato para asistir en el diagnóstico de AD [Cuingnet et al., 2011; Dickerson et al., 2009; Hua et al., 2008].

Dentro de este enfoque, encontramos el método de espesor cortical basado en atlas, en el cual todos los vértices de la corteza cerebral han sido agrupados en un atlas a fin de utilizarlo como “referencia” en posteriores clasificaciones. Los resultados de Desikan et al. [2006], posiblemente constituyen uno de los principales trabajos en esta área y han servido de base para numerosas publicaciones. Desikan et al. [2006] propuso un método de etiquetado automático para subdividir la corteza cerebral humana entre ROIs basados en la zona de la circunvolución (*gyrus*). Utilizando 40 sesiones de MRIs, identificaron 68 regiones de interés (*Region of Interest, ROI*) (34 ROIs para cada uno de los hemisferios izquierdo y derecho). Finalmente esta información fue codificada en un atlas, a través del cual es posible etiquetar automáticamente los ROIs del cerebro en casos reales. Este método es utilizado para estudios morfométricos y funcionales de la corteza cerebral humana, así como en investigaciones clínicas dirigidas para rastrear los cambios cerebrales causados por una enfermedad en el tiempo. Básicamente, estos estudios obtienen los 64 ROIs a partir de las neuroimágenes y por cada uno de ellos calculan mediciones como el promedio, la desviación estándar, el área, etc.; las cuales son utilizadas como patrones o características de la enfermedad.

Desikan et al. [2009] buscó determinar las mediciones sobre MRIs capaces de identificar con alto grado de exactitud sujetos en la etapa MCI. Para este fin, calcularon la media del espesor (para regiones neo corticales) y el volumen (para regiones neo corticales y no-neocorticales) sobre los 68 ROIs [Desikan et al., 2006] obtenidas de las sesiones de MRIs de más de 97 sujetos. En la figura 2.7 se muestra una representación tridimensional de 34 ROIs utilizadas en este estudio (para un hemisferio). Como resultado, Desikan et al. [2009] demostró que las mediciones automáticas del espesor de la corteza entorrinal (*“entorhinal cortex thickness”*), el volumen del

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

hipocampo y el espesor de la circunvolución supramarginal (“*supramarginal gyrus thickness*”) son altamente discriminativas para identificar sMCI y sAD con aceptables porcentajes de exactitud y especificidad de 94% y 74%, respectivamente. Adicionalmente, resultados similares han sido presentados en [Desikan et al., 2010], los cuales demostraron que mediciones automáticas sobre MRIs calculadas sobre la corteza temporal medial predicen con alta exactitud y seguridad la progresión de MCI_{AD} a DAT.

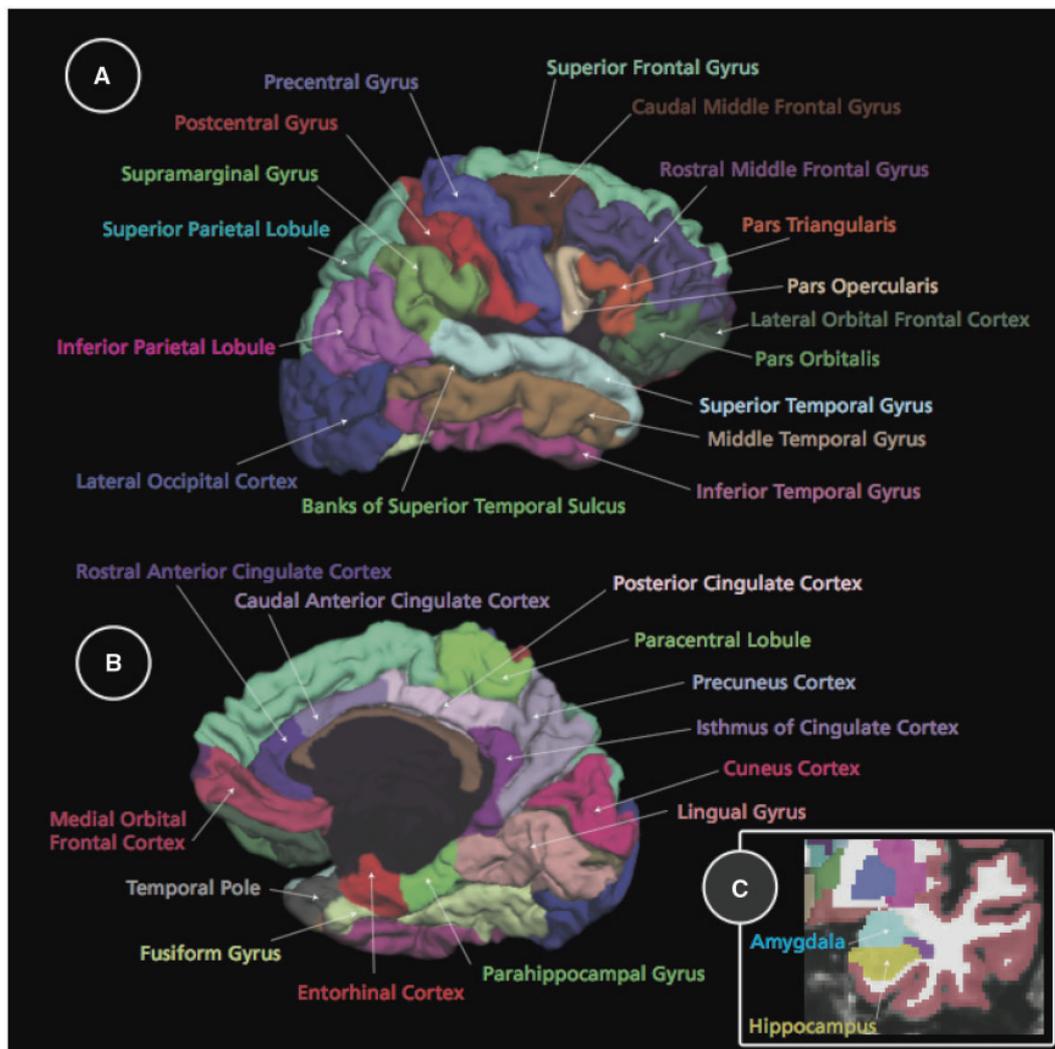


Figura 2.7: Representación tridimensional de los 34 ROIs (vista de un hemisferio) examinados en el estudio de [Desikan et al., 2009] y originalmente propuestos por el método de etiquetado de la corteza cerebral propuesto en [Desikan et al., 2006] .

2.5.3 Clasificación basada en ROIs.

La tercera categoría basada en ROIs engloba únicamente las investigaciones centradas en el análisis del volumen y/o la forma del hipocampo [Chupin et al., 2009; Mueller et al., 2010]. El hipocampo es una de las primeras regiones cerebrales afectadas por AD [Das et al., 2012; Dubois et al., 2007; Morra et al., 2009; Mueller et al., 2010].

En el trabajo de Chupin et al. [2009] se presentó un método de segmentación automática del hipocampo y la amígdala basado en la deformación de regiones simultánea, considerando *a priori* información anatómica y probabilística de ambos tejidos. El método de segmentación fue aplicado en MRIs de pacientes sC (166), sMCI 210: 76 progresaron a AD (sMCIc) y 134 no progresaron a AD (sMCIinc); y sAD (145). Con la finalidad de verificar si el método propuesto podría proporcionar biomarcadores de AD, los volúmenes del hipocampo obtenidos fueron utilizados para testear la discriminación de los tres tipos de sujetos. Como resultado, la clasificación de los pacientes como sAD y sC alcanzó una máxima exactitud de 80%, especificidad de 78% y sensibilidad de 74%. Asimismo, la discriminación entre sMCIc y sMCIinc obtuvo una máxima exactitud de 65%, especificidad de 64% y sensibilidad de 67%. Estos últimos resultados demostraron que a pesar de que el hipocampo proporciona información útil para el diseño de herramientas de pronóstico (progresión hacia el estado de demencia AD), nos es suficiente para discriminar, con altos valores de exactitud, entre sujetos sMCIinc y sMCIc; siendo necesario complementar estos métodos con otras técnicas de clasificación.

En [Mueller et al., 2010], basándose en la estructura heterogénea del hipocampo al estar conformado por algunas regiones o subcampos con características histológicas distintas: corteza entorrinal, el subículo, CA1, zona de transición CA1-2, CA3 -circunvolución dentada (DG); ver figura 2.8; buscaron determinar patrones de pérdida de volumen dentro del hipocampo de 91 sujetos: sC (53), sMCI (20) y sAD (18); y establecer si el volumen hipocampal era eficiente para discriminar entre los tres grupos. Como resultado encontraron que estos subcampos contenidos dentro del hipocampo, podrían resultar más discriminativos en el diagnóstico temprano de AD, en comparación con el uso del volumen completo del hipocampo.

[Das et al., 2012] estableció un método de morfometría basada en la deformación (*deformation-based morphometry*, DBM) para determinar la eficiencia del volumen del hipocampo (obtenido sobre MRIs) como discriminador de los cambios longitudinales producidos en el cerebro de sujetos sC y sMCI. Adicionalmente, [Das et al., 2012] analizó el grado de atrofia de los subcampos del hipocampo, encontrando importantes diferencias entre ambos grupos de sujetos.

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

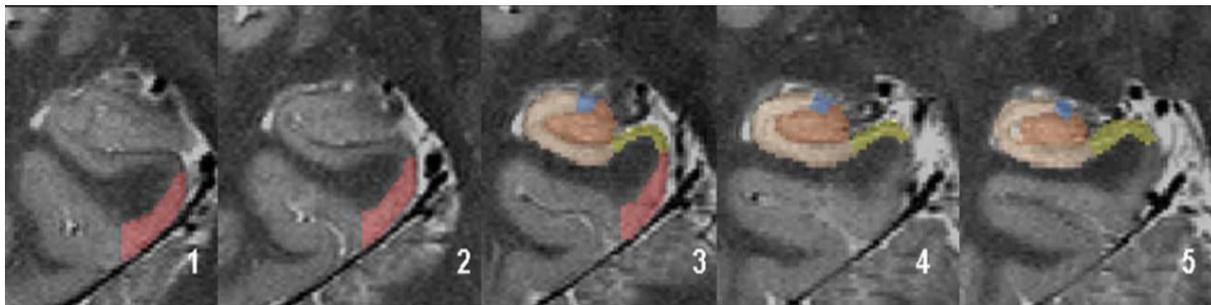


Figura 2.8: Ejemplo de los subcampos del hipocampo marcados manualmente: (1) Corte inicial; (2,3 y 4) Cortes intermedios; (5) Corte más posterior. Etiquetas: Roja, corteza entorrinal; Amarilla, subículo; Beige, CA1; azul, zona de transición CA1-2 ; Marrón, CA3&DG. Fuente:[Mueller et al., 2010].

2.5.4 Clasificación mixta.

Existen investigaciones sobre clasificación de pacientes de AD, las cuales han combinado 2 o más de las categorías anteriores con otro tipo de variables establecidas por datos neuropsicológicos y sociodemográficos de los pacientes. Por ejemplo, [Cui et al., 2011] buscó determinar un conjunto de variables o características predictoras de la conversión de sujetos MCI a demencia AD. Examinaron simultáneamente múltiples características de diferentes modalidades de datos (incluyendo biomarcadores morfométricos basados en MRIs) correspondientes a sujetos sC, sMCI y sAD. Como resultado, fueron seleccionadas como características más discriminantes las mediciones de biomarcadores morfométricos alusivos al volumen subcortical del hipocampo (hemisferio izquierdo y derecho), el volumen cortical de la circunvolución temporal media (*Middle temporal gyrus*) correspondiente al hemisferio derecho y la media del espesor cortical de las regiones: corteza entorrinal del hemisferio izquierdo, corteza parietal inferior (hemisferio derecho), corteza retrosplenial (hemisferio izquierdo) y la circunvolución temporal media (hemisferio izquierdo).

2.5.5 Ventajas y desventajas de los enfoques de clasificación.

Los enfoques de clasificación mencionados, presentan resultados con gran porcentaje de exactitud, la mayoría de ellos, superiores al 84%. Sin embargo, estos enfoques han sido evaluados sobre poblaciones de estudios diferentes, lo cual hace difícil comparar su poder de clasificación. Con la finalidad de calificar el desempeño de estos enfoques y su aporte en la clasificación sC vs. sAD, sC vs. sMCIc, sC vs. sMCIc; Cuingnet et al. [2011] ha realizado un estudio muy completo para comparar diferentes métodos de clasificación. Este estudio fue aplicado sobre una misma población y de esta manera fueron controlados factores como: la edad, género, nivel, de

educación, grado de deterioro, etc; los cuales podrían influenciar la evaluación del desempeño de la clasificación. Un total de 11 métodos fueron testeados: (i) 5 métodos basados en el enfoque de vóxel MRI; (ii) 3 métodos basados en el enfoque cortical; (iii) 3 métodos basados en el volumen (2) y la forma (1) del hipocampo. Los resultados y conclusiones más resaltantes de este estudio son comentados a continuación. Para mayor detalle de los métodos analizados, los datos utilizados y los resultados de clasificación, ver [Cuingnet et al. \[2011\]](#) .

- Discriminación entre sujetos en la etapa AD (sMCIc y sAD) y sC.
 - Todos los métodos de clasificación testeados alcanzaron una exactitud significativamente alta en la discriminación de sujetos sAD de los sujetos sC.
 - Todos los métodos, excepto un método basado en el enfoque de vóxel y los métodos basados en el hipocampo obtuvieron resultados significativamente buenos en la discriminación entre sujetos sMCIc y sC.
 - Para la clasificación sAD *vs.* sC, la mayoría de los métodos alcanzaron alta sensibilidad y especificidad; sin embargo, en la discriminación de sMCIc, su sensibilidad fue sustancialmente baja.

- Predicción de la conversión en pacientes MCI (sMCIc *vs.* sMCIc).

 - Ninguno de los 11 métodos estudiados fue capaz de predecir la conversión de sujetos en la etapa MCI_{AD} a DAT con resultados aceptables; sin embargo uno de los métodos basado en el hipocampo alcanzó los valores más altos en sensibilidad y especificidad de 62% y 69%, respectivamente.
 - [Cuingnet et al. \[2011\]](#) explica que el pobre desempeño de estos métodos en la discriminación sMCIc *vs.* sMCIc se justifica por el hecho que los sujetos sMCIc son un grupo muy heterogéneo; es decir, algunos sMCIc se convierten en corto tiempo a AD, mientras que otros se mantienen estables por un largo periodo de tiempo. Por este motivo, recomiendan que los métodos de clasificación deberían enfocarse en la detección de sujetos MCI convertidos, los cuales constituyen un grupo más homogéneo.

2. BIOMARCADORES MORFOMÉTRICOS OBTENIDOS DE MRI.

Capítulo 3

Modelos de Clasificación de AD basada en SVM.

3.1 Introducción.

En los capítulos anteriores detallamos estudios muy importantes basados en la utilización de las MRIs en pacientes con AD para la identificación de biomarcadores morfométricos, los cuales pudiesen ser utilizados para el diagnóstico temprano, la planificación del tratamiento y el monitoreo de la progresión de esta enfermedad. Sin embargo, la utilización de estos biomarcadores presenta dos problemas: (1) su disposición e interpretación no resulta una tarea fácil y directa debido a la complejidad de las estructuras cerebrales y las grandes cantidades de datos que por lo general necesitan ser analizadas; y (2) su repercusión se ha visto limitada porque típicamente reportan diferencias entre pacientes y sujetos de control a nivel grupal; mientras que en la práctica clínica, los médicos tienen que tomar decisiones sobre un paciente específico. En consecuencia, para que los biomarcadores puedan ser útiles en ambientes clínicos y ser considerados en los criterios de diagnóstico clínico, es necesario (además de numerosos estudios que sustenten su confiabilidad) proporcionar herramientas de cálculo que permitan obtenerlos directamente y ser capaces de hacer inferencias de diagnóstico y predicción a nivel de individuo.

Desde hace unos pocos años atrás, ha habido un creciente interés dentro de la comunidad de las neurociencias por el uso de métodos analíticos que permitan hacer inferencias a nivel individual a partir de los biomarcadores. Uno de estos métodos está basado en el uso de máquinas de aprendizaje automático (*Machine Learning*, ML), . ML es un método multidisciplinario basado en el desarrollo de algoritmos y técnicas capaces de extraer automáticamente información a partir de datos. En estudios de demencia, ML proporciona una efectiva forma para automatizar el análisis y diagnóstico de datos clínicos e imágenes radiológicas. Por ejemplo, ML puede reducir

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

considerablemente la carga de los médicos al identificar automáticamente patrones complejos en las imágenes médicas y brindarles soporte en la toma de decisiones inteligentes. Una completa revisión de la aplicación de ML en radiología ha sido presentada en [Wang and Summers, 2012]. Las principales aplicaciones de ML en demencia están estrechamente relacionadas con dos enfoques: (1) la extracción de patrones (representadas como variables) más discriminativas; por ejemplo, volúmenes de ROIs del cerebro; y (2) el diseño de modelos de diagnóstico, los cuales predicen si un sujeto es o no demente, a partir de las variables extraídas [Chen and Herskovits, 2010].

La máquina de soporte vectorial (*Support Vector Machine*, SVM) es un tipo específico de ML que ha tenido gran aceptación en el diagnóstico clínico. En el campo de la neurología relacionada a AD, ha constituido la base de estudios [Davatzikos et al., 2008; kai Shen et al., 2012; Klöppel et al., 2008; Plant et al., 2010; Vemuri et al., 2008] dedicados al análisis de bases de datos con alta dimensionalidad para el desarrollo de nuevas herramientas de diagnóstico basadas en biomarcadores morfométricos de MRIs (así como otro tipos de biomarcadores). La mayoría de estos estudios utilizaron SVM en dos principales tareas. La primera dedicada a la reducción de la dimensionalidad de los patrones espaciales del cerebro representadas como variables¹, utilizando diferentes técnicas de extracción y selección de variables. La segunda tarea dedicada a construir modelos de clasificación² de pacientes a partir de las variables seleccionadas en la tarea anterior. De esta manera, los modelos de clasificación SVM son utilizados como herramientas de diagnóstico y pronóstico, capaces de clasificar automáticamente grupos de pacientes con altos porcentajes de sensibilidad y especificidad durante las etapas de test. Es decir, estos modelos son capaces de dar soporte al diagnóstico de AD mediante la discriminación de sujetos sanos sC, sMCI y sAD; y dar soporte a la predicción de AD al ser capaces de discriminar dentro de los sujetos sMCI, los sujetos sMCI_{inc} (con MCI estable) y los sujetos sMCI_{inc} (con riesgo de conversión a DAT).

En este capítulo presentamos una breve introducción a los principales conceptos de SVM y los componentes necesarios para su configuración. Comentamos algunos de los modelos basados en SVM que han sido propuestos para clasificar sujetos como sC, sMCI o AD. Asimismo hacemos una breve introducción a las mediciones estadísticas utilizadas para calificar el desempeño de estos modelos.

¹En adelante se utilizará el término de variables morfométricas o variables predictoras morfométricas para hacer referencia a las características o patrones espaciales del cerebro utilizados como biomarcadores morfométricos de AD.

²En este trabajo hacemos referencia a modelos de clasificación como una estructura y su interpretación correspondiente que engloba un conjunto de datos y configuraciones generados por algoritmos de aprendizaje utilizados como clasificadores

3.2 Máquinas de soporte vectorial (SVM): Visión general

ML es el estudio de algoritmos informáticos, los cuales pueden aprender relaciones complejas o patrones a partir de datos empíricos y tomar decisiones acertadas. Puede ser considerado como un campo interdisciplinario en estrecha relación con la inteligencia artificial, el reconocimiento de patrones, minería de datos, estadística, teoría de la probabilidad, optimización, física estadística y ciencia de la computación. Debido a esta característica interdisciplinaria, ML abarca diferentes tópicos [Bishop, 2006] como: *kernel learning*, modelos probabilísticos, modelos lineales, redes neuronales artificiales (*Artificial Neural Networks*, ANNs), reducción de la dimensionalidad y selección de variables, entre otros. En la figura 3.1 se presenta un esquema de los principales tópicos de ML presentada en [Wang and Summers, 2012], los cuales han sido ampliamente descritos en el libro de C.M. Bishop [Bishop, 2006]. Obsérvese que el diagrama representa la interconexión entre los tópicos, lo cual explica la base interdisciplinaria de ML.

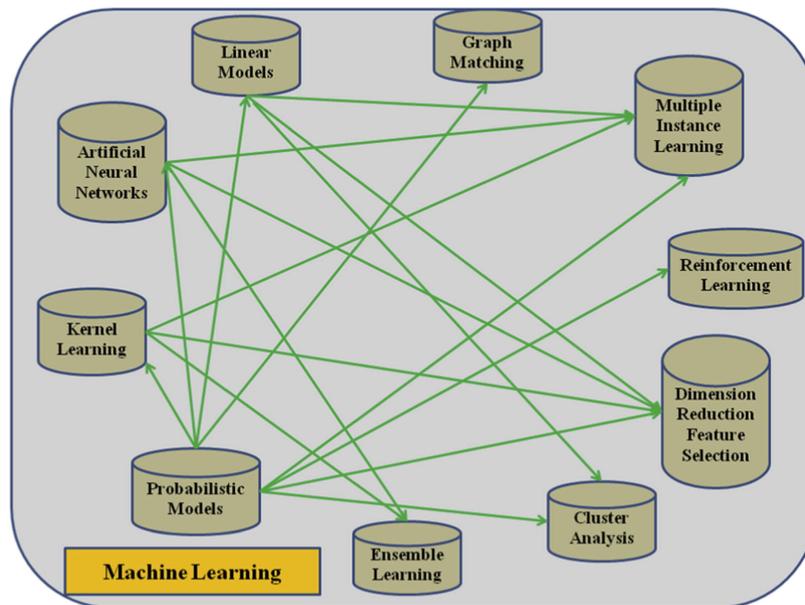


Figura 3.1: Principales tópicos de ML. Fuente:[Wang and Summers, 2012].

Dentro de ML, existen dos principales enfoques que pueden ser asumidos: el aprendizaje supervisado y aprendizaje no supervisado. En el aprendizaje supervisado se busca desarrollar una función, la cual mapea dos o más conjuntos de observaciones con dos o más categorías de definidas por operadores a través de un procedimiento iterativo que reduce gradualmente la diferencia entre el resultado predicho y el resultado esperado. Posteriormente, el algoritmo puede ser utilizado para asignar nuevas observaciones (datos nunca antes tratados por el algoritmo),

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

a una de las categoría predefinidas con una determinada precisión. Por otro lado, en el aprendizaje no-supervisado, se busca determinar cómo los datos son organizados sin la disponibilidad de información *a priori* proporcionada por el operador, siendo el objetivo principal descubrir lo desconocido.

3.2.1 Clasificadores.

Una forma específica del algoritmo de reconocimiento de patrones supervisado es la solución de problemas de clasificación. Los problemas de clasificación están relacionados con el descubrimiento automático de regularidades en los datos de observaciones, las cuales pueden ser utilizadas para clasificar los datos en diferentes categorías predefinidas. Un clasificador es una función de decisión que toma los valores de distintas variables (*features*) pertenecientes a una instancia (*example*) y predice la “clase-etiqueta” (*class label*) a la que esta muestra pertenece. De este modo, el objetivo de un clasificador es capturar con el mayor porcentaje de exactitud, la relación entre cada instancia y su correspondiente clase [Orrù et al., 2012; Pereira et al., 2009].

Nota 7 *En un problema de clasificación (también en problemas de regresión), las características, también denominadas como patrones y predictoras; son variables independientes, las cuales forman parte de una instancia. Las instancias, también llamadas muestras, son conjuntos de variables independientes. Y las clases-etiquetas, son variables dependientes, las cuales constituyen la respuesta del clasificador.*

El clasificador está formado por un número de parámetros que son “aprendidos” a partir de datos de entrenamiento y el aprendizaje es evaluado a partir de nuevos datos durante una etapa de test. El común criterio utilizado en algoritmos de aprendizaje de clasificadores es que las instancias que constituyen los datos de entrenamiento y test han sido extraídos de manera independiente a partir de una “distribución de instancias” y que cuando evaluamos un clasificador sobre un conjunto de test, estamos obteniendo un estimado de su rendimiento sobre algún conjunto de datos de test a partir de la misma distribución [Pereira et al., 2009].

Con la finalidad de formalizar el aprendizaje de un clasificador y el modelo de relaciones entre las características y las clases en un conjunto de entrenamiento, podemos denotar una instancia como un vector de columnas $x = [x_1, \dots, x_v]$ y su clase-etiqueta como y ; entonces, dado una instancia x , el clasificador es una función f que predice la etiqueta de la siguiente manera: $y = f(x)$. Después de ser entrenado el clasificador, éste es utilizado para determinar si las variables de un nuevo conjunto de datos contienen información sobre la clase de la instancia. Esta relación es comprobada mediante una etapa de test, en la cual el clasificador es utilizado para predecir un diferente conjunto de instancias (no vistos antes), conocido como datos de test.

En la figura 3.2 se presenta un esquema del proceso de aprendizaje de un clasificador (representación tomada y modificada de [Pereira et al., 2009]). Obsérvese que en el esquema un conjunto de datos iniciales ha sido dividido en un conjunto de entrenamiento y un conjunto de test; los parámetros del clasificador son aprendidos a partir de las instancias del conjunto de entrenamiento; y entonces, el clasificador es evaluado sobre las instancias de un independientes conjunto de instancias del grupo de test. Finalmente, las etiquetas predecidas son comparadas con etiquetas verdaderas (clases reales) a fin de evaluar el desempeño de la predicción del clasificador. Exactitud (*Accuracy*) [Kohavi and Provost, 1998] es la medición estadística comúnmente utilizada para evaluar el rendimiento de un clasificador sobre los datos de test, la cual simplemente representa la fracción de instancias en el conjunto de test para las cuales han sido predecidas correctamente su etiquetas, es decir es una comparación con las etiquetas reales. Esta medición y otras mediciones muy importantes para esta tarea de evaluación, son descritas con mayor detalle en la sección 3.3.3.

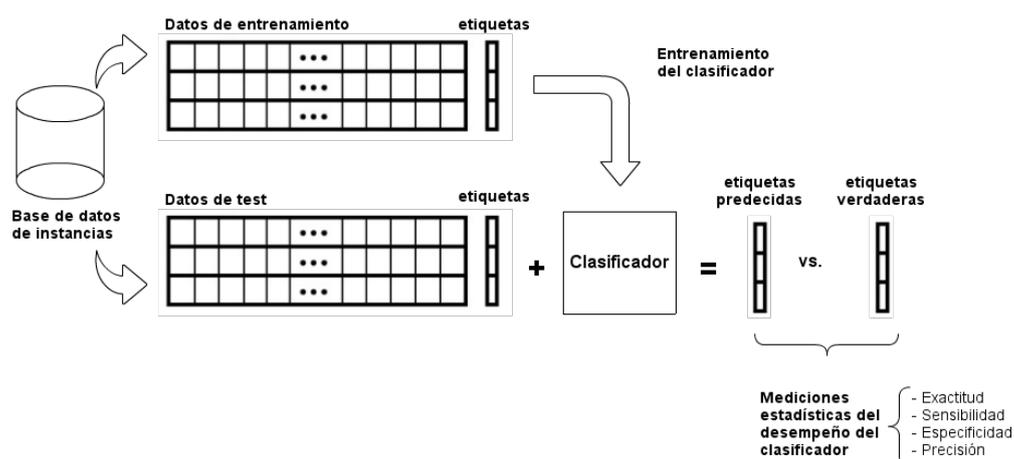


Figura 3.2: Representación del aprendizaje de un clasificador.

Existen diferentes tipos de clasificadores, la mayoría de los cuales se basan en técnicas de reconocimiento de patrones, como las ANNs, árboles de decisión y SVM, entre muchas otras técnicas [Bishop, 2006]. Sin embargo por razones que serán comentadas más adelante, en este estudio nos hemos enfocado en el método SVM como un clasificador lineal, donde la función de clasificación depende de una combinación lineal de los atributos y pesos denotados por w_i . Esta combinación adquiere la forma general $f(x) = g(w_1x_1 + \dots w_yx_y)$ y son los parámetros que deben ser aprendidos por el clasificador SVM.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

3.2.2 Función kernel

Antes de introducirnos en los conceptos de las SVMs, describimos las funciones kernel, las cuales constituyen la base de este tipo de ML. La manera más simple de realizar la separación de las clases es mediante una línea recta, un plano recto o un hiperplano N-dimensional. Sin embargo, en los problemas reales no se suelen presentar casos ideales de dos dimensiones como en el ejemplo presentado en la figura 3.3, en la cual los datos de dos clases, denotadas por círculos azules y círculos rojos son separadas por el contorno de decisión (curva magenta). El algoritmo de aprendizaje de los clasificadores debe tratar con problemas con: (1) más de dos variables predictoras, (2) curvas no lineales de separación, (3) casos donde las instancias no pueden ser completamente separadas y (4) clasificaciones para más de dos categorías (multiclases). En consecuencia, algunas veces las ML lineales presentan limitaciones computacionales y no pueden ser utilizadas.

La representación de un clasificador por medio de funciones Kernel ofrece una solución a estos problemas, proyectando la información a un espacio de características de mayor dimensión, lo cual aumenta la capacidad computacional de las máquinas de aprendizaje lineal.

Existen diversos tipos de funciones kernel propuestos para problemas lineales y no lineales. Las funciones kernel más utilizados para las SVMs serán mencionadas más adelante.

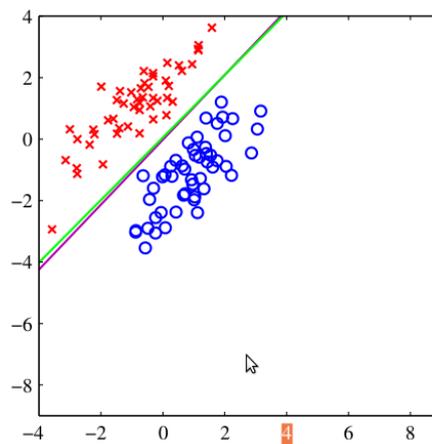


Figura 3.3: Ejemplo de separación lineal de un problema de clasificación ideal.

3.2.3 Definición de SVM.

SVM es un tipo específico de ML desarrollado por Vapnik [1995], el cual ha alcanzado gran popularidad en la solución de problemas de clasificación, regresión y detección. Las SVMs in-

cluyen un conjunto de algoritmos de aprendizaje supervisado muy útiles para resolver problemas de clasificación de datos; en algunos casos su utilización ha sido considerada más fácil que las ANNs. Una mayor descripción sobre SVM y sus diferencias con las ANNs puede ser encontrado en [Burges, 1998; Kecman, 2004, 2005].

Dado un conjunto de instancias de entrenamiento es posible etiquetar las clases y entrenar una SVM a fin de construir un modelo que “prediga” la clase de una nueva instancia. Intuitivamente, una SVM es un modelo que representa a las instancias como puntos en el espacio, separandolas en clases con un espacio (margen) lo más amplio posible, conocido como hiperplano [Vapnik, 1995]. De este modo, cuando las nuevas muestras se ponen en correspondencia con dicho modelo, pueden ser clasificadas a una u otra clase en función de su proximidad.

Definición formal.

De manera más formal, una SVM construye un hiperplano o conjunto de hiperplanos en un espacio de dimensionalidad muy alta (o incluso infinita) que puede ser utilizado en problemas de clasificación (o regresión). De este modo, una buena separación entre las clases permitirá una clasificación correcta.

En un problema de clasificación de dos clases, SVM se describe como: dado un conjunto de entrenamiento de tamaño l conformado por los pares instancia-clase etiqueta $(x_i, y_i), i = 1, \dots, l$; donde $x_i \in R^n$ y $y \in [0, 1]$ o $[-1, 1]$, las SVMs requieren la solución del problema de optimización presentado en la ecuación 3.1 [Chih-Wei et al., 2003].

$$\begin{aligned} \min_{\omega, b, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \\ \text{suje}to \quad & y_i (w^T \phi(x_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0 \end{aligned} \tag{3.1}$$

Los vectores de entrenamiento (vectores de soporte) x_i son mapeados en un espacio dimensional alto (tal vez infinito) por la función ϕ . SVM busca un hiperplano o conjunto de hiperplanos proyectados en este espacio de dimensionalidad muy alta que tenga la máxima distancia (margen) con los puntos que estén más cerca del hiperplano y separe de forma óptima a los puntos de una clase de la de otra. $C > 0$ es conocido como un parámetro de penalización en términos del error. Es ese concepto de “separación óptima”, donde reside la característica fundamental de las SVMs. De esta forma, los puntos de los vectores de entrenamiento que estén etiquetados con una categoría estarán a un lado del hiperplano y los puntos que se encuentren en la otra categoría estarán al otro lado.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

Además, la relación definida como $K(x_i, x_j) \equiv \phi(x_i^T) \phi(x_j)$ es conocida como la función kernel, introducida en la sección 3.2.2. Existen diversos tipos de kernel utilizados en SVM, siendo los más utilizados:

- *Linear:*

$$K(x_i, x_j) = x_i^T(x_j) \quad (3.2)$$

- *Polinomial:*

$$K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \quad (3.3)$$

- *Radial Basis Function:*

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2), \gamma > 0 \quad (3.4)$$

- *Sigmoid:*

$$K(x_i, x_j) = \tanh(\gamma x_i^T x_j + r) \quad (3.5)$$

donde γ , r y d son los parámetros del kernel.

Problema no lineal.

Como se acaba de describir, la base de SVM la constituyen los algoritmos basados en *kernel*, en los cuales, las predicciones para nuevos datos de entrada depende únicamente de una función *kernel* evaluada en el subconjunto de los puntos de datos de entrenamiento [Bishop, 2006]. En la sección 3.2.2 se presentó una breve descripción de las características de las funciones kernel y en la sección 3.2.3 se presentaron los tipos de kernels más utilizados en SVM.

Existen dos principales ventajas del desarrollo de kernels en un contexto SVM [Orrù et al., 2012]. Primero, el kernel es una representación eficiente de la similitud entre las instancias, lo que significa que la complejidad computacional de las SVMs es establecida por el número de muestras en lugar del número de variables, lo cual es ventajoso en configuraciones de alta dimensionalidad. Segundo, los kernels pueden ser utilizados para “aprender” contornos de decisión no lineales, en cuyo caso, los kernels no lineales son utilizados para mapear implícitamente los datos a partir de un espacio de entrada (donde la separación no lineal es posible) hacia un espacio de variables de altas dimensiones (donde los datos pueden ser separados por un espacio lineal).

3.3 Etapas de construcción de un clasificador SVM

Las principales etapas de la construcción de clasificadores SVM son: (1) Preparación de los datos de entrenamiento y test, (2) Entrenamiento y test del clasificador y (3) Evaluación de su desempeño.

3.3.1 Preparación de datos para entrenamiento y test.

Antes de proceder con el desarrollo de un clasificador, tres pasos preliminares son requeridos: (1) extracción de instancias y variables, (2) selección de variables y (3) preprocesamiento de los datos.

Las extracción de instancias implica la transformación de los datos originales en un conjunto de instancias integradas por variables (características) y sus respectivas clases-etiqueta. Las características y las clases-etiquetas son utilizadas como “variables de entrada” y “variables de salida”, respectivamente en un SVM. Estos datos son representados en una matriz, en la cual las filas son las instancias y las columnas son las variables de entrada; más una columna (ubicada por lo general al inicio o al final de la matriz) que representa la variable de salida. En el contexto de diagnóstico de AD, este paso podría consistir en la transformación de información de un paciente o sujeto de control a un vector de variables de entrada y salida adecuado, en el cual, cada una de las variables de entrada represente una característica del individuo, como información de biomarcadores y características clínicas; y la variable de salida represente el diagnóstico de AD .

La selección de variables implica la selección de un subconjunto de variables a partir de las variables obtenidas en el paso anterior, con el objetivo de optimizar el aprendizaje del clasificador. Principalmente, se busca eliminar las variables más redundantes y seleccionar las variables más relevantes y discriminantes para el problema de clasificación. En el contexto del diagnóstico de AD, este paso podría implicar la selección de un conjunto de variables con mediciones de determinadas ROIs del cerebro de los sujetos a clasificar; por ejemplo, variables conteniendo información de biomarcadores morfométricos, las cuales sean más discriminantes para clasificar diferentes clases de sujetos con valores de exactitud altos. En la sección 3.4 describimos con mayor detalle este paso.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

Preprocesamiento de las instancias es un paso final a considerar durante la preparación de las instancias e implica tareas de transformación de los datos al formato de números reales y funciones de normalización o escalamiento aplicados a cada columna de variables de entrada. La normalización o escalamiento de los datos, antes de ser utilizados en el entrenamiento de las SVMs, trae muchas ventajas; sin embargo, es importante tener en cuenta cuál es la forma de escalamiento más adecuada de acuerdo al tipo de datos que se está evaluando. Para mayor explicación de la importancia del escalamiento y cómo proceder para su selección, ver [Beale et al., 2011; Edwards and Raskutti, 2005]. En el contexto del diagnóstico de AD, el escalamiento de los datos es útil para tratar todas las variables de manera uniforme y resolver la diferencia de rango numérico y unidades de medición entre los diferentes tipos de variables morfométricas y clínicas. Una ventaja adicional es reducir las dificultades numéricas durante el cálculo, debido a que en la solución de problemas de clasificación utilizando SVM existen productos entre vectores, en donde los atributos con mayor rango numérico podrían haber afectado a los de menor rango.

3.3.2 Entrenamiento y Test.

Como comentamos en la sección 3.2.3, SVM tiene como objetivo desarrollar un algoritmo que permita la clasificación de observaciones individuales en clases distintas basada en datos sobre un espacio de dimensiones altas. Para alcanzar este objetivo, el algoritmo de aprendizaje del SVM es desarrollado mediante una etapa de entrenamiento y validado mediante una etapa de test, como fue descrito en la sección 3.2.1, y representado de manera general en la figura 3.2. Asimismo, este objetivo implica a su vez tareas iniciales como la selección de una función kernel adecuada para el tipo de problema y la optimización de los parámetros del algoritmo de aprendizaje y los parámetros propios del kernel, utilizando datos de entrenamiento.

3.3.2.1 Selección de la función kernel y búsqueda de los mejores parámetros.

Para que las SVMs encuentren la “separación óptima” de los grupos de datos es necesario la selección de una función *kernel* adecuada y encontrar los valores más óptimos para la configuración de sus parámetros. Por ejemplo, utilizando el kernel RBF, la configuración de los parámetros C y γ es requerida. C es un parámetro de costo del clasificador SVM, el cual permite mayor flexibilidad en la separación de las clases, por lo que es importante asignarle un margen adecuado. γ (*Gaussian kernel width*) es un parámetro propio del kernel RBF, ver ecuación 3.4. Inicialmente, los valores para ambos parámetros no son conocidos por lo que su asignación y mejor combinación es usualmente seleccionada utilizando una red de búsqueda (*grid-research*) con secuencias de crecimiento exponencial para C y γ .

3.3.2.2 Entrenamiento y test

Durante la fase de entrenamiento, un SVM es entrenado utilizando datos predefinidos (dentro de grupos) por el operador con la finalidad de estimar una “función de decisión” o “hiperplano” ϕ , ver ecuación 3.1, la cual distinga de mejor manera entre los grupos de interés (por ejemplo: pacientes y controles). La figura 3.4 ilustra un hipotético problema de clasificación binaria entre pacientes con AD (puntos rojos) y sujetos sanos de control (puntos azules), considerando únicamente dos variables predictoras. Cada eje representa el valor obtenido por cada variable, con cada instancia u observación de un diferente sujeto representada con un símbolo (puntos). La solución de este problema es realizada por una separación (delineamiento) de las clases, como es presentado en la figura 3.4(a). Obsérvese que existen diferentes posibilidades para este delineamiento (líneas discontinuas); sin embargo el objetivo de un clasificador SVM (y posiblemente la tarea más difícil del clasificador) es ser entrenado para maximizar el margen de separación entre los dos grupos, utilizando las instancias más próximas al plano de separación como puntos de definición. Estos puntos son llamados “vectores de soporte” (*support vectors*). En la figura 3.4(b) el clasificador óptimo (del tipo lineal) es representado por la línea roja discontinua y los vectores de soporte son encerrados por círculos de color negros.

Durante la etapa de test, el algoritmo SVM es utilizado para predecir un grupo de datos (o sólo una instancia) nunca antes evaluado por el clasificador. Para garantizar la generalización de la solución al problema de clasificación, es necesario que ninguna de las instancia de datos de test se superpongan a los datos utilizados durante el entrenamiento.

3.3.3 Evaluación del desempeño de los clasificadores

Antes de aplicar un clasificador en la práctica, es necesario que su desempeño sea evaluado. El desempeño de un algoritmo de clasificación es generalmente descrito por mediciones estadísticas como la exactitud, precisión, sensibilidad, especificidad y el área bajo la curva del Receptor de Funcionamiento Característico (“*Receiver Operating Characteristic*”, ROC) conocido como AUC; las cuales son calculadas sobre la matriz de confusión (o matriz de contingencia) obtenida por el clasificador durante la etapa de test. La selección de estas mediciones para un estudio en particular dependerá del objetivo del estudio [Alonzo and Pepe, 2007]. Alonzo and Pepe [2007] presenta una descripción más completa de las categorías de los estudios clínicos a partir de los objetivos que persiguen.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

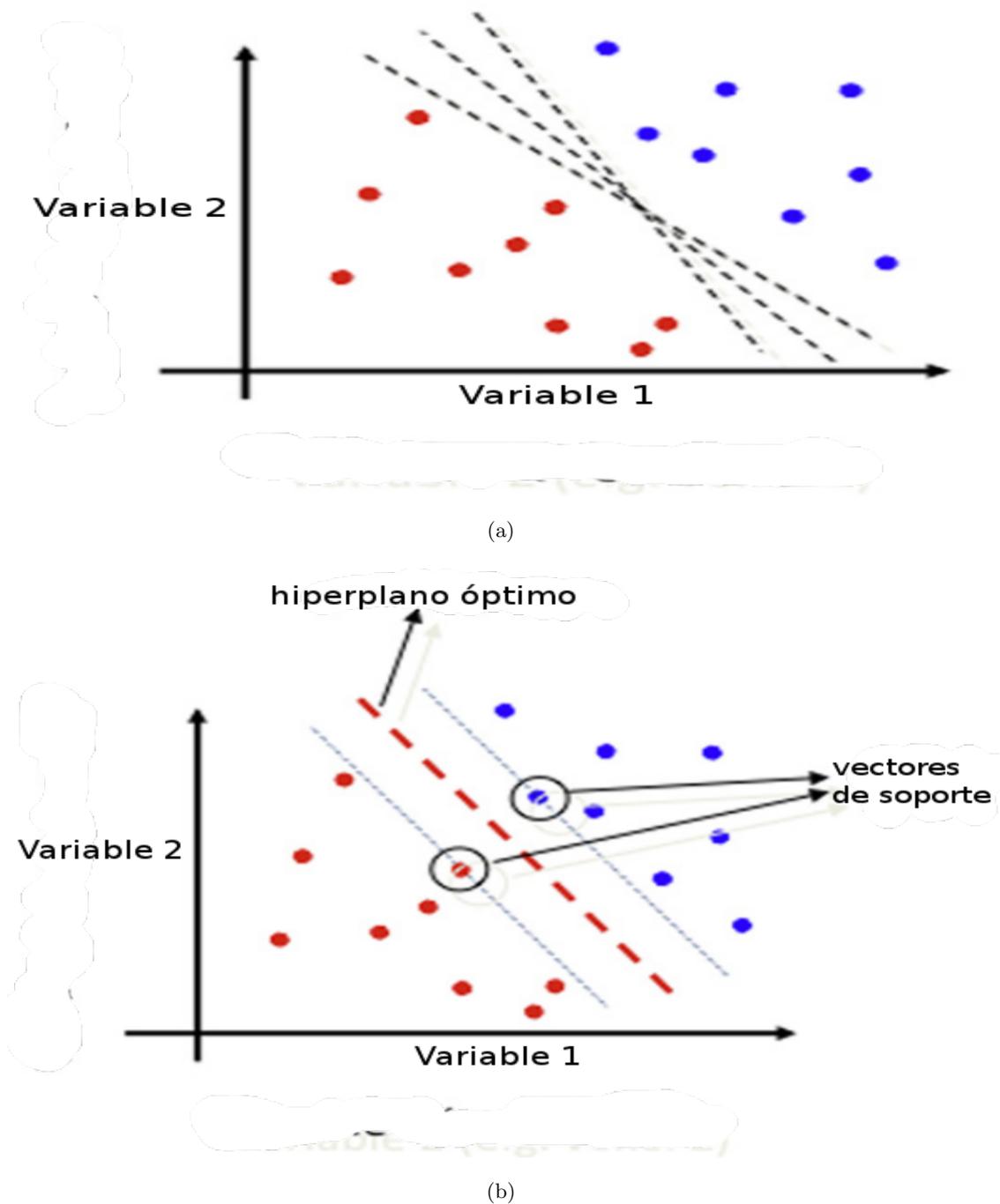


Figura 3.4: Hipotética representación de un problema de clasificación binario: sC (puntos rojos) vs. sAD (puntos azules): (a) Diferentes posibilidades de separación de las clases; (b) Separación óptima obtenida con SVM.

3.3.3.1 Matriz de confusión

La matriz de confusión representa las clasificaciones predecidas y actuales, donde las primeras son obtenidas por un clasificador [Kohavi and Provost, 1998]. Una matriz de confusión es de tamaño $l \times l$, donde l es el número de diferentes valores de la clase-etiqueta. Por ejemplo, dado un problema de clasificación binario donde l es 2 y las clases pueden tomar los valores positivo (p) y negativo (n), su matriz de confusión puede ser representada como en la tabla 3.1.

Tabla 3.1: Representación de una matriz de confusión

↓ actual \ predecido →	negativo (n)	positivo (p)
negativo (n)	TN	FP
positivo (p)	FN	TP

Considerando la matriz de confusión presentada en la tabla 3.1 para un problema de clasificación binario, en la que los resultados se etiquetan como p o n . Hay cuatro posibles resultados a partir de un clasificador binario como el propuesto. Si el resultado predecido por el clasificador es p y el valor real (original) es también p , entonces este resultado se conoce como un **Verdadero Positivo (True Positive, TP)**; sin embargo si el valor real es n entonces se conoce como un Falso Positivo (*False Positive*, FP). De igual modo, si el resultado predecido por el clasificador es n y el valor real también es n , este resultado es un Verdadero Negativo (*True Negative*, TN); y si el valor real es p , este resultado es un Falso Negativo (*False Negative*, FN).

En el contexto del diagnóstico clínico, los indicadores estadísticos obtenidos a partir de una matriz de confusión miden la capacidad de un clasificador en discriminar correctamente alternativos estados de salud. Considerando un ejemplo de clasificación de AD, entre las clases sC (sujeto de control) vs. sAD (sujeto con AD), el clasificador debe determinar a cuál de las dos clases pertenece un sujeto. Si el sujeto es diagnosticado con AD (valor real), un FN ocurre cuando el clasificador predice que el sujeto pertenece a la clase sC; y un TP ocurre si el clasificador etiqueta al sujeto como sAD. Por otro lado, si el sujeto no tiene AD (valor real), un FP ocurre si el clasificador predice que dicho sujeto pertenece a la clase sAD; y un TN ocurre en el caso que el clasificador etiquete como sC.

3.3.3.2 Mediciones basadas en la matriz de confusión

En esta sección describimos los descriptores estadísticos obtenidos a partir de la matriz de confusión a fin de evaluar un clasificador; asimismo estos descriptores también pueden ser extendidos

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

para la evaluación del poder predictivo de una variable [Guyon and Elisseeff, 2003]. Para entender el significado de cada descriptor, utilizamos el ejemplo de clasificación sC vs. sAD, descrito anteriormente. Las definiciones asumidas en este trabajo están basadas en Alonzo and Pepe [2007]; Kohavi and Provost [1998].

Exactitud (ACC) representa la proporción global de las correctas predicciones hechas por un clasificador sobre un conjunto de muestras, ver ecuación 3.6. ACC es usualmente estimada sobre datos de test independientes, los cuales no han sido utilizados durante el aprendizaje del clasificador (etapa de entrenamiento). Existen técnicas más complejas de estimación de ACC, como la validación cruzada, las cuales son comúnmente utilizadas sobre datos con un reducido número de instancias.

$$ACC = (TN + TP)/(TN + FP + FN + TP) \quad (3.6)$$

Precisión (PREC). Se refiere a la proporción de TPs identificados sobre el total de clasificaciones positivas (FP + TP), ver ecuación 3.7.

$$PRE = TP/(FP + TP) \quad (3.7)$$

Sensibilidad (TPR). También conocida como la razón de verdaderos positivos (*True positive rate*, TPR) es la proporción de TPs correctamente identificados por el clasificador, ver ecuación 3.8. TPR mide hasta qué punto un clasificador es capaz de detectar o clasificar los casos positivos correctamente dentro de todos los casos realmente positivos disponibles durante la evaluación; por ejemplo, el porcentaje de personas con AD quiénes han sido clasificados como sAD.

$$TPR = TP/(FN + TP) \quad (3.8)$$

Especificidad (TNR). También conocida como la razón de verdaderos negativos (*True negative rate*, TNR), se refiere a la proporción de TNs correctamente identificados, ver ecuación 3.9; por ejemplo, el porcentaje de sujetos sanos de control identificados como sC.

$$TNR = TN/(TN + FP) \quad (3.9)$$

Existen otras mediciones obtenidas a partir de la matriz de confusión como: (1) la razón de falsos positivos (*False Positive Rate*, FPR), que se refiere a la proporción de falsas alarmas y define cuántos resultados positivos son incorrectos dentro de todos los casos negativos reales

disponibles durante la evaluación; y la razón de falsos negativos (*False Negative Rate*, FNR); los cuales son calculadas utilizando las ecuaciones 3.10 y 3.11, respectivamente.

$$FPR = FP/(TN + FP) = 1 - TNR \quad (3.10)$$

$$FNR = FN/(FN + TP) \quad (3.11)$$

Curvas ROC y Valores AUC. Las curvas ROC son otro tipo de medición del desempeño de un clasificador muy aceptado en las diferentes áreas de aplicación de los clasificadores basados en ML. Una curva ROC representa las ventajas y desventajas (conocido como “trade-offs”) entre las sensibilidad y especificidad de un clasificador a través de una serie de puntos de corte (“*cutoff points*”). Una curva ROC es graficada sobre el espacio ROC, el cual está definido por FPR (1-especificidad) y TPR (sensibilidad) como sus ejes x y y , respectivamente. De este modo, la curva ROC representa FPR vs. TPR asociada con cada resultado binario como un *cutoff* c , el cual varía de $-\infty$ a $+\infty$. Considerando la evaluación de un clasificador en el contexto del diagnóstico de una enfermedad, la curva ROC mide la cantidad de separación entre la distribución de los resultados en la población enferma y la distribución de resultados en la población no enferma. Entonces las siguientes situaciones pueden presentarse:

- Cuando las distribuciones de los resultados para las poblaciones enfermas y no enfermas se solapan, entonces la curva ROC tiene una línea de 45° desde (0,0) hasta (1,1) con $FPR(c) = TPR(c)$ para todos los c . En este caso se dice que el clasificador es no discriminativo;
- Cuando estas distribuciones están más separadas, la curva ROC es más cerrada a la esquina superior izquierda del espacio ROC. Esto significa que una curva que llega hasta la esquina superior izquierda, con $FPR(c) = 0$ y $TPR(c) = 1$ para algún c ; corresponde a una clasificación perfecta.

Una típica forma de describir el desempeño de un clasificador es calcular el área bajo la curva ROC (AUC). El valor $AUC = 1$ representa una clasificación perfecta y $AUC = 0.5$ representa un clasificador no discriminativo. Existen diferentes interpretaciones del valor AUC, dos de ellas aplicadas al diagnóstico clínico: (1) AUC corresponde con la probabilidad de que un resultado de clasificación corresponda un sujeto enfermo seleccionado aleatoriamente supere a la clasificación de un sujeto no enfermo seleccionado aleatoriamente; y (2) el promedio de TPR sobre el rango completo de posibles FPR.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

Para explicar la relación de la curva ROC y el valor AUC en la figura 3.5 se presenta un ejemplo de 4 clasificadores representados por A , B , C y C' . En la figura, la línea diagonal (roja) del diagrama es una línea de estimación aleatoria conocida como “diagonal del espacio ROC”; los puntos por encima de la diagonal, como los puntos A y C' , representan los buenos resultados de clasificación y son considerados como “resultados mejores que el azar”; al contrario, los puntos por debajo de la diagonal, como el punto C corresponden a resultados pobres y comúnmente son definidos como “resultados peores que el azar”. Para resolver la existencia de puntos por debajo de la diagonal, como en el caso del punto C , se suele invertir los resultados de la matriz de confusión (invertir TP con FN y FP con TN), de este modo se aprovecha la capacidad de predicción y se obtiene un nuevo resultado que estaría ubicado por encima de la diagonal, ver punto C' . Por otro lado, el valor máximo de $AUC = 1$, el cual indica (teóricamente) una perfecta clasificación con $TPR=100\%$ y $FPR=100\%$, está representado por el punto (0,1) en el espacio ROC. Asimismo, un valor de AUC de 0.5 ($TPR=50\%$ y $FPR=50\%$) está representado por un punto ubicado dentro de la línea diagonal del espacio ROC, ver punto B . En conclusión, cuanto más cerca esté un punto de la esquina superior izquierda mejor será el desempeño del clasificador; sin embargo lo que en realidad determina el poder predictivo de un clasificador es la distancia de su curva ROC hacia la línea diagonal, así sea por arriba o por abajo de esta.

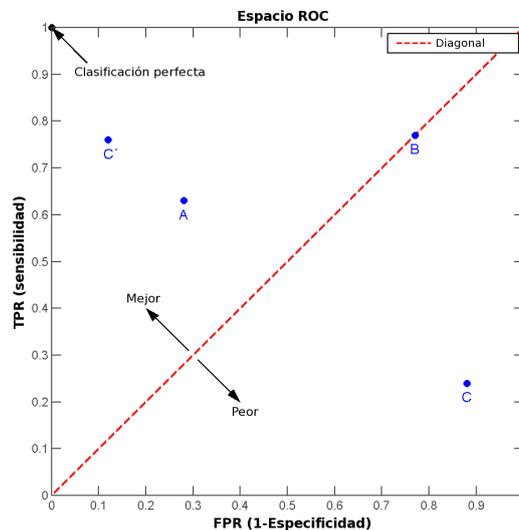


Figura 3.5: Espacio ROC y la representación de los ejemplos de valores AUC.

Adicionalmente, en la tabla 3.2 presentamos una común clasificación de los valores AUC comúnmente asumida en la evaluación de clasificadores o test clínicos.

Tabla 3.2: Clasificación AUC

Rango AUC	Clasificación
$AUC \leq 0.5$	no discrimininate
$0.5 < AUC \leq 0.6$	malo
$0.6 < AUC \leq 0.7$	suficiente
$0.7 < AUC \leq 0.8$	bueno
$0.8 < AUC \leq 0.9$	muy bueno
$0.9 < AUC \leq 1.0$	excelente

3.3.3.3 Factores que afectan el desempeño de un clasificador

Existen diversos factores que pueden afectar el rendimiento de un clasificador, los cuales pueden ser categorizados como factores propios del entrenamiento del SVM y factores propios de las muestras de datos utilizadas. La primera categoría está estrechamente ligada a los procesos comentados en la sección 3.3; por ejemplo, inadecuada asignación de los datos de entrenamiento y test, representación no-balanceada de las muestras para cada clase-etiqueta del problema de clasificación, existencia de variables predictoras poco discriminativas (redundantes o irrelevantes), ineficiente configuración de los parámetros del clasificador, etc. Por otro lado, la segunda categoría de factores podría incluir atributos demográficos (por ejemplo, edad, género, estatus social), características o severidad de la enfermedad (por ejemplo, estadio de AD), condiciones bajo las cuales han sido obtenidas las muestras, entre otras.

Es necesario identificar y comprender la influencia de estos factores debido a: (1) Se puede determinar la mejor manera de entrenar un clasificador a partir de la determinación de la población y entornos de configuración, en los cuales el clasificador presente resultados más o menos exactos, por ejemplo, utilizando análisis de regresión; y, (2) los resultados del clasificador pueden no ser relevantes para poblaciones con diferentes condiciones y características.

3.4 Selección de variables

La selección de variables (características o predictores) consiste en la selección de un subconjunto de variables más relevantes y la eliminación de las que no son redundantes, a partir de un conjunto inicial de ellos. Matemáticamente, este proceso de selección se puede entender como: Dado un conjunto de m muestras $(x_k, y_k)(k = 1, \dots, m)$ conformado por n variables $x_{k,i}(i = 1, \dots, n)$ y una variable de salida (clase-etiqueta) y_k , los métodos de selección buscan un subconjunto $x_{k,s} = (x_{k,1}, \dots, x_{k,s})$, con $s < n$, el cual optimice una función objetivo. Este proceso de selección requiere una búsqueda estratégica para seleccionar un subconjunto de candidatos y una función objetivo para evaluar estos candidatos.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

[Guyon and Elisseeff, 2003] describe los tres objetivos de la selección de variables como: (1) Mejorar el rendimiento de la predicción de los clasificadores o predictores, (2) proporcionar clasificadores más rápidos y efectivos, y (3) proporcionar un mejor entendimiento de los procesos implícitos que generan los datos. En consecuencia, esta etapa proporciona las siguientes ventajas: (1) facilita la visualización y el entendimiento de los datos, (2) reduce los requerimientos de almacenaje, (3) reduce los tiempos de entrenamiento y utilización, y (4) redefine el curso de la dimensionalidad para mejorar el desempeño de la predicción.

Un primer enfoque que se puede utilizar para la selección es la calificación de las variables mediante un *ranking*, el cual califica a cada variable (de manera individual) según su poder de discriminación con respecto a un clasificador. En este enfoque algún criterio que sirva como umbral, es utilizado para la inclusión/exclusión de las variables. Sin embargo, otro enfoque basado en la selección de subconjuntos de variables, las cuales en conjunto tienen una buena capacidad de predicción puede resultar mucho más útil, en contraste a la clasificación de variables de acuerdo a su poder de predicción individual.

Para abordar este problema, los métodos de selección: (1) Métodos de filtro y (2) Métodos Wrapper; son los más utilizados. Sin embargo antes de describirlos, hacemos una breve introducción al ranking de variables, las cual, dependiendo de su utilización en el proceso de selección de variables, puede ser establecida como una etapa de preprocesamiento o ser incluida dentro de los dos métodos de selección.

3.4.1 Ranking de variables

Algunos algoritmos de selección de variables incluyen el ranking de variables como un mecanismo principal o auxiliar de selección debido a su sensibilidad, escalabilidad y buen acierto empírico Guyon and Elisseeff [2003]. Esto significa que el ranking de variables no es necesariamente utilizado para construir predictores, también puede ser utilizado para determinar las variables más discriminativas con respecto a las clases.

Considerando la formulación presentada en la sección anterior, con un conjunto de m muestras (x_k, y_k) ($k = 1, \dots, m$) conformado por n variables $(x_{k,i})$ ($i = 1, \dots, n$) y y_k como variable de salida, el ranking de variables hace uso de una función de puntuación (*score*) $S(i)$ calculada a partir de los valores $x_{k,i}$ y y_k , $k = 1, \dots, m$. Generalmente, puntuaciones más altas indican variables mayor calificadas, entonces las variables pueden ser ordenadas en orden decreciente de $S(i)$.

Utilizado para construir clasificadores.

Cuando el ranking de variables es utilizado para construir clasificadores, las variables son agregadas progresivamente de acuerdo a su relevancia y entonces un clasificador (por ejemplo alguna ML) es entrenado en cada iteración y el desempeño de la variable actual en combinación con las variables ya seleccionadas es evaluado a fin de decidir la exclusión/inclusión de dicha variable, este procedimiento constituye la base de los métodos wrapper y será comentado más adelante. [Guyon and Elisseeff \[2003\]](#) ha demostrado que una variable que es completamente útil por si misma para construir un clasificador puede proporcionar una significativa mejora del desempeño del clasificador cuando es considerada con otras variables; asimismo dos variables que son útiles por si mismas, también son útiles al ser utilizadas en conjunto.

Utilizado como etapa de preprocesamiento.

Por otro lado, el ranking de variables también puede ser utilizado como una etapa de preprocesamiento, independiente de la elección de un clasificador, en cuyo caso es denominado un “método de filtrado” [[Ron and George H, 1997](#)]. EL ranking de variables establecido como un método de filtrado, no intenta optimizar directamente el desempeño del predictor, sólo busca evaluar cada variable individualmente para conocer su influencia en los resultados de clasificación. Computacionalmente, estos métodos son eficientes debido a que únicamente requieren cálculos de n puntuaciones y ordenar las puntuaciones; Estadísticamente, este método es robusto frente a problemas de *overfitting* debido a que introduce un *bias* que puede tener considerablemente menor varianza [Guyon and Elisseeff \[2003\]](#).

A continuación presentamos dos algoritmos de ranking de variables muy utilizados en el campo del diagnóstico de AD.

3.4.1.1 Método de Mínima Redundancia y Máxima Relevancia

El Método de Máxima Relevancia y Mínima Redundancia (Minimum Redundancy and Maximum Relevance, mRMR) es un método filtro desarrollado por [[Ding and Peng, 2005](#)] para la selección de variables. mRMR calcula la información mutua de dos variables basado en su función de densidad probabilística y evalúa todas las variables al observar las características intrínsecas de los datos con respecto al las clases. mRMR es obtenido al optimizar dos criterios: (1) máxima relevancia y (2) mínima redundancia ; simultáneamente. El primer criterio determina cómo una variable discrimina entre las clases y el segundo criterio mide la similitud entre la distribución de las variables y la distribución de las clases.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

En el proceso de selección de variables es importante ordenar y seleccionar las variables de acuerdo a su grado de relevancia en la predicción, pero al mismo tiempo es necesario tener un conjunto de variables no redundantes, con la finalidad de incrementar la robustez de los clasificadores. Para más detalle sobre la implementación de mRMR ver [Ding and Peng, 2005; Peng et al., 2005] y sobre la aplicación de conceptos de redundancia y relevancia en la selección de variables ver [Cui et al., 2011; Yu and Liu, 2004]

3.4.1.2 Método F-score

El método F-score [Chen and Lin, 2006] implica un algoritmo básico y simple que mide la distinción entre dos grupos de números reales. En el método F-score, el valor F-score de cada variable es calculada de acuerdo a la ecuación 3.12.

$$F(i) \equiv \frac{(\bar{x}_i^{(+)} - \bar{x}_i)^2 + (\bar{x}_i^{(-)} - \bar{x}_i)^2}{\frac{1}{n_+ - 1} \sum_{k=1}^{n_+} (x_{k,i}^{(+)} - \bar{x}_i^{(+)})^2 + \frac{1}{n_- - 1} \sum_{k=1}^{n_-} (x_{k,i}^{(-)} - \bar{x}_i^{(-)})^2} \quad (3.12)$$

donde $\bar{x}_i, \bar{x}_i^{(+)}$ y $\bar{x}_i^{(-)}$ son los promedios de la i^{th} variable del conjunto completo, positivos y negativos conjuntos de datos, respectivamente; $x_{k,i}^{(+)}$ es la i^{th} variable de la k^{th} instancia positiva y $x_{k,i}^{(-)}$ es la i^{th} variable de k^{th} instancia negativa. El numerador indica la discriminación entre los conjuntos positivos y negativos; y el denominador indica uno dentro de cada uno de los dos conjuntos.

Entonces, con la finalidad de seleccionar el subconjunto de variables a partir de los datos globales, un valor umbral es obtenido calculando el promedio de los valores F-score de todas las variables. Si el valor F-score de alguna variable es mayor que el umbral, la variable es seleccionada y agregada al espacio de variables; en otro caso, la variable es eliminada.

3.4.2 Métodos Wrapper

Los métodos wrapper utilizan las máquinas de aprendizaje como una “caja negra” para calificar subconjuntos de variables de acuerdo a su poder predictivo [Ron and George H, 1997]. Utilizando un formulación más general, estos métodos consisten en utilizar el desempeño de predicción de de una ML para calificar la utilidad relativa de un subconjunto de variables.

En la práctica, la utilización de los wrapper es determinada por las siguientes condiciones: (1) cómo buscar el espacio de todos los posibles subconjuntos de variables, [Ron and George H, 1997] presenta un amplio estudio sobre las estrategias de búsqueda más utilizadas, incluyendo algoritmo genéticos; (2) cómo calificar el desempeño de predicción de una ML a fin de orientar el proceso de búsqueda de variables y detenerlo cuando sea necesario, para lo cual comúnmente

es utilizada la validación cruzada; y (3) qué clasificador o predictor utilizar, siendo los más utilizados SVM, árboles de decisión, entre otros.

3.4.2.1 SVM como un método wrapper de selección de variables

Adicionalmente a la utilización de SVM como un efectivo método de clasificación, SVM también es utilizado para la selección de variables importantes para la construcción de un óptimo modelo predictor [Chen and Lin, 2006; Cui et al., 2011; Cuingnet et al., 2011; Mesrob et al., 2008]. Sin embargo, SVM no obtiene directamente la relevancia de las características, para ello es necesario modificar el algoritmo para seleccionar las variables más útiles a partir del entrenamiento y testeo de los datos. Por otro lado, SVM puede ser combinado con otros métodos de selección como los métodos de filtrado, independientes de SVM. De este modo, las variables importantes son seleccionadas por esos métodos y entonces SVM es aplicado para clasificación.

En este estudio, SVM ha sido utilizado para la selección de características importantes y como un método de clasificación, en ambos casos, con la finalidad de construir eficientes modelos de diagnóstico y predicción de AD.

3.5 Revisión de los estudios SVM aplicados en el diagnóstico de AD

En el campo de la neurología relacionada con estudios de AD, SVM ha constituido la base de los principales estudios [Davatzikos et al., 2008; kai Shen et al., 2012; Klöppel et al., 2008; Plant et al., 2010; Vemuri et al., 2008] dedicados al análisis de bases de datos con alta dimensionalidad a fin de desarrollar de nuevas herramientas de soporte al diagnóstico, basadas en biomarcadores morfométricos de MRIs (así como otro tipos de biomarcadores) y su combinación con otro tipo de características. Una de las principales características de SVM para el diagnóstico clínico es su capacidad para construir modelos de clasificación a partir del aprendizaje con datos de entrenamiento iniciales, conocida como etapa de entrenamiento; y clasificar con óptima exactitud nuevos casos clínicos, etapa conocida como *test*. Para ejemplo de diagnóstico de AD, los datos de entrenamiento y test corresponden a casos clínicos y el problema de clasificación está relacionado con la predicción entre grupos o clases previamente establecidas, por ejemplo, sAD vs. sC.

Dado que la eficiencia de los modelos depende de los datos utilizados para su construcción, una de las tareas más críticas es la selección de variables predictoras e instancias significativas para la conformación de los grupos de datos utilizados durante el entrenamiento, validación y test del modelo.

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

Con respecto a la utilización de SVM para el diagnóstico de AD, la mayoría de los estudios se han centrado en discriminar entre sMCI y sAD, lo cual se justifica por el creciente interés en predecir, en una etapa más temprana posible como la etapa MCI, los sujetos que pueden progresar hacia AD, a fin de iniciar un tratamiento temprano que mejore la calidad de vida del paciente y se preserve y prolongue en lo posible sus capacidades cognitivas. Asimismo, otros estudios se han centrado en comparar individuos sMCI con sC, por razones similares a las mencionadas en el primer caso. Biomarcadores basados en neuroimagen obtenidos a partir de diferentes modalidades (MRI, PET, DTI), combinados con otro tipo de biomarcadores como los basados en CSF y otras características (sociodemográficas y neuropsicológicas); han sido utilizados para construir modelos de clasificación con alto desempeño.

En la tabla 3.3 se resume una investigación de los principales estudios relacionados con la discriminación entre sC vs. sAD, sC vs. sMCI, sMCI vs. sAD y sMCIc vs. sMCIinc; algunos de los cuales ya han sido mencionados en los capítulos anteriores. La mayoría de estos estudios utilizaron SVM en dos principales tareas. La primera dedicada a la reducción de la dimensionalidad de los patrones o características espaciales del cerebro utilizando diferentes técnicas de extracción y selección de variables ¹. La segunda tarea dedicada a construir modelos de clasificación de pacientes a partir de las variables seleccionadas en el tarea anterior. De esta manera, los modelos de clasificación SVM fueron utilizados como herramientas de diagnóstico y pronóstico, capaces de clasificar automáticamente grupos de pacientes con altos porcentajes de sensibilidad y especificidad durante las etapas de test. Principalmente, estos modelos fueron capaces de dar soporte al diagnóstico de AD mediante la discriminación de sujetos sanos sC y sujetos sMCI y sAD; y dar soporte a predicción de AD al ser capaces de discriminar dentro de los sujetos sMCI, los sujetos sMCIinc (con MCI estable) y los sujetos sMCIc (con riesgo de conversión a AD). Davatzikos et al. [2005] demostraron que es posible discriminar entre sujetos utilizando respuestas “verdaderas” y “no verdaderas” con una exactitud del 99.3% basado en patrones discriminativos de la actividad cerebral, utilizando fMRIs. Más tarde, Davatzikos et al. [2008] desarrollaron un método automático para detectar los patrones de atrofia a través de la deformación de neuroimágenes de altas dimensiones, robustos métodos de extracción de variables y SVM. kai Shen et al. [2012]; Klöppel et al. [2008]; Plant et al. [2010]; Vemuri et al. [2008] utilizaron SVM para determinar el potencial de las MRIs en el diagnóstico y predicción de demencia del tipo AD, mediante modelos de clasificación de sAD vs. sC, encontrándose resultados muy prometedores. [Chen and Lin, 2006] aplicaron siete métodos de ML (incluyendo

¹En adelante se utilizará el término de variables morfométricas o variables predictoras morfométricas para hacer referencia a las características o patrones espaciales del cerebro utilizados como biomarcadores morfométricos de AD.

SVM) para construir cinco modelos de clasificación entre sujetos sanos y sujetos con demencia, encontrando diferencias de desempeño significativas entre estos modelos.

Tabla 3.3: Investigaciones en AD relacionadas con biomarcadores basados en imágenes y SVM

Autor(es)	Clasificación	Grupos	Biomarcadores	Desempeño clasificación (%)
Mesrob et al. [2008]	sAD vs. sC	sAD = 17 sC = 13	MRI estructural (GM)	ACC=90.2; TPR=84.3;TNR=96.0
Arimura et al. [2008]	sAD vs. sC	sAD = 29 sC = 25	MRI estructural	ACC=82.7
Davatzikos et al. [2008]	sMCI vs. sC	sMCI = 15 sC = 15	MRI estructural	ACC=90
Duchesne et al. [2008]	sAD vs. sC	sAD = 75 sC = 75	MRI estructural	ACC=92
Fan et al. [2008a]	sMCI vs. sC	sMCI = 15 sC = 15	MRI estructural & PET	ACC=100
Ferrarini et al. [2008]	sAD vs. sC	sAD = 58 sC = 28	MRI estructural	ACC=84
Klöppel et al. [2008]	sAD vs. sC	sAD = 18 FTLD = 19	MRI estructural	ACC=95
	sAD vs. FTLD	sAD = 190 sC = 190	MRI estructural	ACC=89.2
Vemuri et al. [2008]	sAD vs. sC	sAD = 190 sC = 190	MRI estructural & datos genéticos	ACC=89.3
Gerardin et al. [2009]	sAD vs. sC	sAD = 23 sC = 23	MRI estructural	ACC=94
	sMCI vs. sC	sMCI = 23 sC = 25	MRI estructural	ACC=83
Magnin et al. [2009]	sAD vs. sC	sAD = 16 sC = 22	MRI estructural, DTI	ACC=94.5
Haller et al. [2010]	sC vs. sMCI	sC = 35 sMCI = 67	MRI estructural, DTI	ACC=91.4
Nho et al. [2010]	sAD vs. sC	sAD = 182 sC = 226	MRI estructural	ACC=90.5
Oliveira et al. [2010]	sAD vs. sC	sAD = 14 sC = 20	MRI estructural	ACC=88.2
[Plant et al., 2010]	sAD vs. sC	sAD = 32 sC = 18	MRI estructural (GM + WM)	ACC=90
	sMCI vs. sC	sMCI = 24 sC = 18	MRI estructural (GM + WM)	ACC=97.62
Salas et al. [2010]	sAD + sMCI vs. sC	sAD + sMCI = 167 sMCI = 52	PET	ACC=86
	sAD vs. sC	sAD = 53 sC = 52	PET	ACC=95
	sMCI vs. sC	sMCI = 114 sC = 53	PET	ACC=88
Abdulkadir et al. [2011]	sAD vs. sC	sAD = 95 sC = 95	MRI estructural (GM)	ACC=84.5
Chen et al. [2011]	sAD vs. sC	sAD = 21 sC = 20	MRI funcional	ACC=87
	sMCI vs. sC	sMCI = 15 sC = 20	MRI funcional	ACC=95
Chincarini et al. [2011]	sAD vs. sC	sAD = 144 sC = 189	MRI estructural	TPR. = 89;TNR = 94
	sMCI vs. sC	sMCI = 79 sC = 204	MRI estructural & DTI	ACC=71.09
Cui et al. [2011]				
Dukart et al. [2011]	sAD vs. FTLD vs. sC	sAD = 21 FTLD = 14 sC = 13	MRI estructural (GM) & PET	ACC=91.7
Grana et al. [2011]	sAD vs. sC	sAD = 20 sC = 25	DTI	ACC=100
Hinrichs et al. [2011]	sAD vs. sC	sAD = 48 sC = 66	MRI estructural & PET & CSF(tau, amygdaloid-beta 142, p-tau 181P, t-tau, genotipo APOE) & NMs	ACC=88.2
Zhang et al. [2011a]	sAD vs. sC	sAD = 45 sC = 50	MRI estructural & PET & CSF(A β 42, t-tau and p-tau)	ACC=92
	sMCI vs. sC	sMCI = 91 sC = 50	MRI estructural & PET & CSF(A β 42, t-tau and p-tau)	ACC=80
Zhang et al. [2011b]	sAD vs. sC	sAD = 51 sC = 52	MRI estructural & PET & CSF(A β 42, t-tau and p-tau)	ACC=93.2
	sMCI vs. sC	sMCI = 99 sC = 52	MRI estructural & PET & CSF(A β 42, t-tau and p-tau)	ACC=76.4
Cuingnet et al. [2011]	sAD vs. sC;	sAD = 68 sC = 81	MRI estructural	TPR=81.0;TNR=95.0
	sC vs. sMCIc;	sC=81 sMCIc=37	MRI estructural	TPR=65.0;TNR=94.0
	sMCIc vs. sMCIcInc;	sMCIc=67 sMCIc=37	MRI estructural	TPR=62.0;TNR=69.0
kai Shen et al. [2012]	sAD vs. sC	sAD = 39 sC = 138	MRI estructural (hipocampo)	TPR=87.2;TNR=79.5

Debido a las ventajas de SVM presentadas en la sección 3.2.3 y los excelentes resultados obtenidas por las diversas investigaciones presentadas en esta sección, nuestro estudio estuvo centrado en SVM como método de ML para construir modelos de clasificación de casos clínicos

3. MODELOS DE CLASIFICACIÓN DE AD BASADA EN SVM.

de AD. En los siguientes capítulos se describe ampliamente la utilización de SVM y todos los conceptos aplicados, los cuales han sido presentados en este capítulo y los anteriores.

Capítulo 4

Construcción de los modelos de diagnóstico y predicción de AD basados en SVM.

4.1 Introducción

En el capítulo 1 se comentaron dos características resaltantes en el manifiesto de los estados AD-preclínico, (MCI_{AD}) (MCI ocasionado por AD)¹ y DAT (demencia ocasionada por AD)². La primera relacionada con el diagnóstico de AD, el cual es realizado después de la exclusión de otras formas de demencia, pero su diagnóstico definitivo es realizado únicamente después de un estudio post-muerte del tejido cerebral. La segunda característica relacionada con la ocurrencia de cambios neuropatológicos entre la transición de las tres etapas, causados por la densidad y distribución de las neurofibras de τ y $A\beta$, las cuales están presentes mucho antes de la aparición de primeros síntomas clínicos de AD.

En consecuencia, en el mundo de las neurociencias, existe una amplia necesidad de desarrollar metodologías que sirvan como sustitutos in vivo para la detección de estos cambios patológicos y, por tanto, sirvan para identificar con precisión a los individuos que se encuentren en las primeras etapas de AD. Es así, que la identificación de factores claves para el diagnóstico clínico de AD resulta uno de los principales avances conseguidos en el área de la neurología. El continuo descubrimiento de nuevas variables de diagnóstico y predicción basadas en algún tipo de biomarcador, presenta oportunidades y retos para los estadísticos y clínicos a la hora de evaluar su influencia en el desarrollo de esta enfermedad.

¹En adelante el estado “MCI ocasionado por AD” (MCI_{AD}) será referenciado en este trabajo como MCI.

²El estado de demencia ocasionada por AD (DAT) en adelante será referenciada como AD.

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

Diversas recomendaciones han surgido desde algo más de 25 años, a partir de que el primer caso de AD fuese descrito [Knopman et al., 2001], las cuáles han tratado de abordar principalmente: (1) la confiabilidad de los criterios de diagnóstico actuales,(2) su capacidad para establecer un diagnóstico confiable en la evaluación de las personas mayores, (3)la capacidad de las pruebas de laboratorio para mejorar el diagnóstico clínico. Asimismo, estas consideraciones han demostrado que la asociación estadística de una variable con el grado de demencia de un paciente, no es suficiente. Una de las cuestiones clave a resolver es cómo abordar la forma de evaluar y cuantificar la mejora en el diagnóstico y predicción de la enfermedad utilizando los diversos biomarcadores propuestos.

Para abordar este problema es necesario considerar que una simple búsqueda y clasificación de todas las variables más relevantes en la solución de un problema de clasificación como el diagnóstico clínico de AD, usualmente tiene un bajo desempeño para la construcción de un buen predictor (particularmente si las variables son redundantes). En consecuencia, obtener un subconjunto de variables útiles podría excluir aquellas variables que son redundantes pero también relevantes ¹.

En este capítulo se describe una metodología para la selección de variables predictoras útiles y el aprendizaje de dos modelos de clasificación basados en máquinas de soporte vectorial (SVM) de soporte al diagnóstico clínico de AD. Estos modelos fueron entrenados a partir de las variables predictoras seleccionadas y son clasificados como: modelo de diagnóstico capaz de discriminar sujetos sanos o de control (sC) de los sujetos con AD (sAD) y modelo de predicción capaz de discriminar dentro de los sujetos diagnosticados con deterioro cognitivo leve (MCI por sus siglas en inglés), aquellos sujetos con MCI que se convertirían a AD (sMCIc) y los sujetos con MCI que no se convertirían a AD (sMCInc) .

La metodología presentada está orientada a dar soporte al diagnóstico clínico de AD, principalmente basado en un enfoque de investigación o en ensayos clínicos de AD. En el capítulo 1 se especificó que el diagnóstico clínico definitivo de esta enfermedad sólo puede ser hecho después de su muerte (diagnóstico histológico). Sin embargo, en vida, es posible hacer un diagnóstico clínico denominado como “probable” o “posible”, incluso en las etapas iniciales.

4.2 Metodología propuesta

Existe una amplia investigación centrada en la determinación de las variables más útiles para el diagnóstico clínico de demencia a lo largo de las diferentes etapas de esta enfermedad, sin

¹En los trabajos de Ron and George H [1997] y Blum and Langley [1997] se presenta una discusión de relevancia versus utilidad de variables, así como varias definiciones de relevancia.

embargo, sólo desde unos pocos años atrás, algunas de estas investigaciones se han centrado en el uso de las máquinas de aprendizaje para la identificación de estas variables [Chen and Herskovits, 2010], por lo que la obtención de resultados óptimos en esta área se constituye en un problema abierto.

En este estudio proponemos una metodología basada en los datos (*“data-driven method”*) enfocada principalmente a dos tareas: (1) La selección de las variables y (2) El entrenamiento de modelos de clasificación SVM para el diagnóstico y predicción de AD. El diagnóstico de AD está dirigido a la discriminación automática entre sujetos con AD (sAD) y sujetos de control (sC); y la predicción de AD está dirigida a la discriminación entre los sujetos con MCI no convertidos a AD (sMCI_{nc}) de los sujetos con MCI convertidos a AD (sMCI_c).

La metodología propuesta está basada en las tres etapas principales de la construcción de modelos de clasificación, descritas en el capítulo 3, sección 3.3: (1) Preparación de los datos de entrenamiento y test, (2) Entrenamiento y test de los modelos y (3) Evaluación de su desempeño. En la figura 4.1 se presentan los principales procesos implicados en esta metodología.

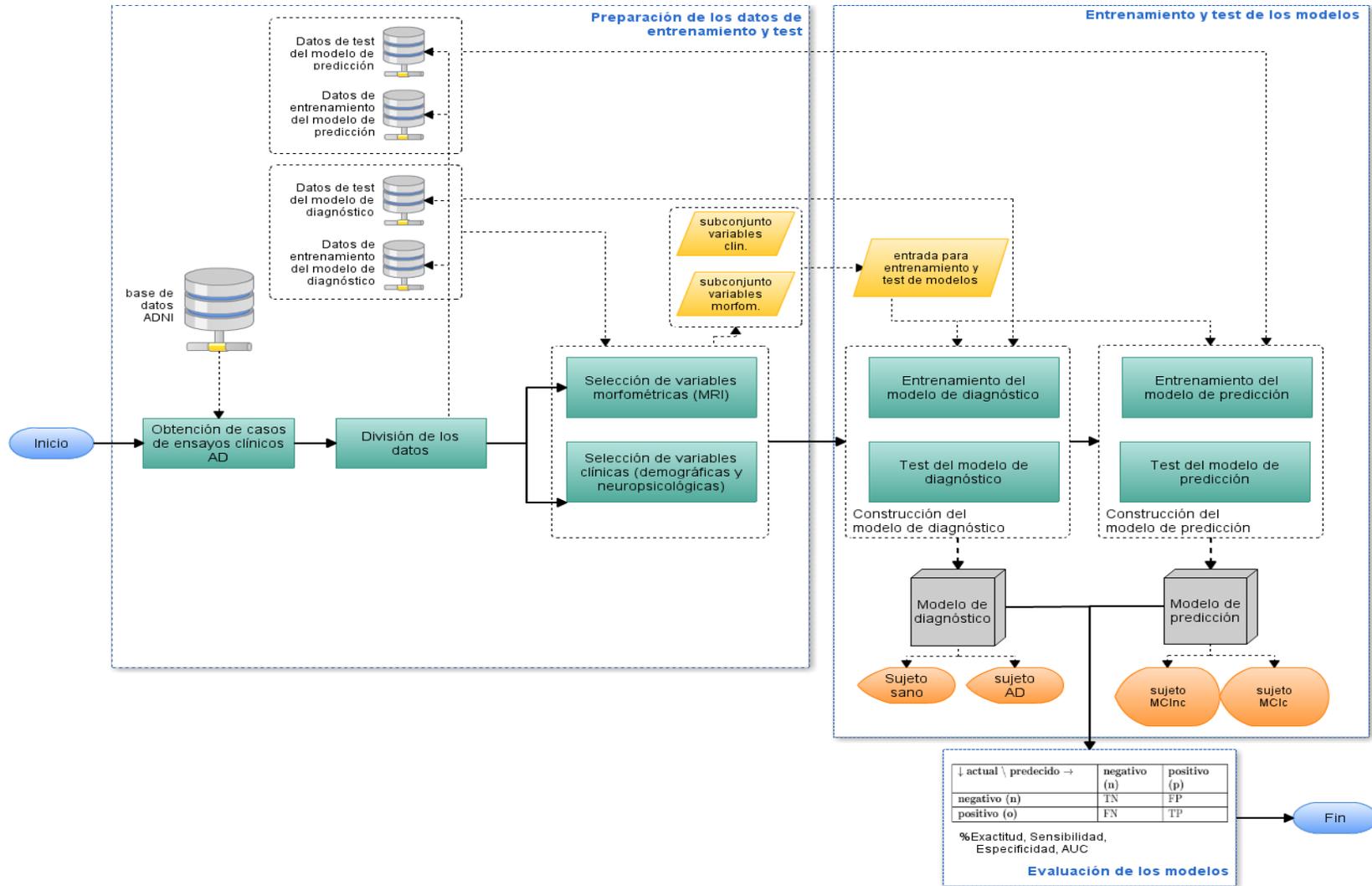


Figura 4.1: Principales etapas implicadas en la metodología de selección de variables y construcción de los modelos SVM para el diagnóstico y predicción de AD

Las preparación de los datos incluye principalmente: (1) La obtención de los datos de ensayos clínicos de AD para la conformación de los grupos destinados al entrenamiento, validación y test de los modelos. Estos datos estuvieron conformados por 324 variables iniciales integradas por biomarcadores morfométricos basados en imágenes médicas, biomarcadores neuropsicológicos y variables sociodemográficas, las cuales fueron obtenidas de datos clínicos del proyecto de investigación en la enfermedad de Alzheimer denominado “*Alzheimer’s Disease Neuroimaging Initiative*” (ADNI) (www.loni.ucla.edu/ADNI) [Weiner et al., 2012], el cual ofrece una amplia información sobre biomarcadores de AD; y (2) la selección del subconjunto de variables de mayor significancia para la discriminación entre sujetos de control (sC) y sujetos con AD (sAD)¹, aplicando la combinación de un método de *ranking* de variables con dos métodos *wrapper* basados en Support Vector Machine (SVM).

Para la etapa de entrenamiento y test de los modelos de diagnóstico y predicción se determinó el comportamiento no lineal del problema de clasificación y se establecieron los parámetros de configuración de los modelos SVM a entrenar: el tipo de la función kernel y la elección de los mejores parámetros para SVM. Después de obtener estos parámetros, los modelos fueron entrenado y testados utilizando los datos obtenidos en la etapa anterior.

Finalmente, la etapa de evaluación de los modelos implicó el cálculo de las mediciones estadísticas descritas en la sección 3.3.3 sobre las matrices de confusión obtenidas durante la etapa de test, con la finalidad de conocer el desempeño de los modelos SVM en las tareas de diagnóstico y predicción de AD.

En este capítulo es descrita la primera etapa de la metodología. Las etapa de entrenamiento y test de los modelos es presentada en el capítulo y las etapa de evaluación es presentada en el capítulo 6.

4.3 Obtención de datos de ensayos clínicos de AD

Para la obtención de datos válidos fue utilizada la base de datos de ensayos clínicos en AD del proyecto *Alzheimer’s Disease Neuroimaging Initiative* (ADNI) (www.loni.ucla.edu/ADNI) [Weiner et al., 2012], obtenida online en junio del 2011. ADNI es un proyecto dedicado a la investigación de Alzheimer (AD) y deterioro cognitivo leve (MCI), lanzada en el 2003 por el *National Institute on Aging* y otras reconocidas instituciones dedicadas a la investigación de AD. ADNI ofrece una amplia información sobre biomarcadores de AD, así como información

¹La selección de variables juega un rol importante en la construcción de sistemas de clasificación, lo cual no sólo ayuda en la reducción de los datos, también disminuye los costos de computación y beneficia el buen rendimiento de la clasificación

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

demográfica y neuropsicológica de los participantes. La base de datos completa incluye información de 800 participantes: 200 sujetos de control (sanos), 400 pacientes con MCI y 200 pacientes con AD leve, todos provenientes de 58 lugares de los Estados Unidos y Canadá. Los procedimientos de selección de los participantes y el protocolo del estudio completo utilizado por ADNI, fueron descritos por [Jack et al. \[2008\]](#); [Mueller et al. \[2005a\]](#) y se puede encontrar en línea en <http://www.nia.nih.gov/Alzheimers/ResearchInformation/ClinicalTrials/ADNI.htm>. Asimismo, una completa revisión de los principales resultados y publicaciones alcanzados hasta el momento por este proyecto ha sido presentado recientemente por [Weiner et al. \[2012\]](#).

ADNI proporciona los datos de sus estudios en diversas tablas clasificadas según el tipo de información: datos demográficos, resultados de análisis de imágenes, resultados test clínicos, diccionario de variables utilizadas, etc, las cuales incluyen una o más instancias por participante, registrados en cada visita de los participantes ¹.

4.3.1 Fiabilidad de los datos de ensayo clínicos utilizados

ADNI es uno de las plataformas más utilizados y reconocidas en investigación de AD, en los recientes años, y el uso de su base de datos ha implicado la publicación de más de 100 trabajos científicos. Asimismo, investigaciones recientes han surgido sobre cómo implementar un diseño similar a ADNI para el estudio de otro tipo de enfermedades neurodegenerativas, como la enfermedad de Parkinson [[Frisoni and Weiner, 2010](#)].

La fiabilidad de utilizar la base de datos ADNI como datos de ensayos clínicos de AD para investigaciones en AD queda demostrada por investigaciones europeos dedicadas a testear la viabilidad del uso de la plataforma americana ADNI en investigación de AD en Europa [[Frisoni, 2010](#); [Frisoni and Weiner, 2010](#); [Frisoni et al., 2008](#)]. En [[Frisoni et al., 2008](#)] se describe una de las primeras iniciativas con el desarrollo de un estudio piloto llevado a cabo para establecer la plataforma E-ADNI (*European Alzheimer's Disease Neuroimaging*), para lo cual, siete centros académicos del Consorcio Europeo de la enfermedad de Alzheimer (EADC: *European Alzheimer's Disease Consortium*), enrolaron un grupo de 59 sujetos europeos clasificados como sanos (18), sujetos MCI (19) y sujetos con AD (22), los cuales fueron estudiados siguiendo los criterios de análisis y controles de ADNI. Las características sociodemográficas, clínicas, neuropsicológicas y morfométricas, entre otras, fueron comparadas con otro grupo semejante de

¹ Los datos descargados contenían información demográfica, neuropsicológica y morfométrica de los participantes obtenidas en sus visitas: *screen visit* (sc), sexto mes (m06), doceavo mes (m12), dieciochoavo mes (m18), veinticuatroavo mes (m24) y treinta y seisavo mes (m36). Las variables morfométricas corresponden a MRIs estructurales de los estudios transversal y longitudinal.

sujetos norteamericanos, obtenidos de la base de datos ADNI. Como resultados más destacables de ese estudio se confirmó la similitud cognitiva y estructural entre ambos grupos y la viabilidad de la utilización de la plataforma ADNI para el enrolamiento de pacientes y la colección de datos clínicos, neuropsicológicos y morfométricos (datos volumétricos del cerebro) por centros del EADC. Por un lado, las características cognitivas, conductuales y neuropsicológicas fueron muy similares entre el grupo europeo y el grupo norteamericano. Asimismo, en el análisis de los biomarcadores morfométricos relacionados con la atrofia cerebral de los sujetos (calculados a partir de sus imágenes de Resonancia Magnética), se obtuvo que los valores del grupo de los sujetos norteamericanos en la mayoría de los grupos de diagnóstico, tiende a ser ligeramente mayor que su contraparte europea, sin embargo esta diferencia no es significativa. Una revisión más detallada de las principales iniciativas en investigación de AD basadas en ADNI, llevadas a cabo en Europa, es presentada en [Frisoni, 2010].

4.3.2 Criterios de inclusión de los casos clínicos

A fin de incluir instancias (casos clínicos)¹ y variables ² útiles para nuestro estudio, sólo fueron consideradas las instancias que cumplieran los siguientes criterios:

1. Muestras o instancias de participantes con MRIs estructural de ambos estudios transversal y longitudinal, correctamente preprocesadas y analizadas, es decir con un análisis completo.
2. Muestras de participantes con información demográfica y neuropsicológica completa, como mínimo para la primera visita (*screen visit*) .
3. Muestras de participantes que cumplieran los criterios de clasificación clínica de ADNI mencionadas en la tabla 4.1.

Los criterios de inclusión y exclusión de los participantes establecidos por ADNI y asumidos en este trabajo para la conformación de los grupos de entrenamiento, validación y test, son detallados en la tabla 4.1. En la tabla se observa que la clasificación de los participantes estuvo sujeta principalmente a dos características cognitivas muy utilizadas en el diagnóstico de demencia: el índice de *Mini-Mental State Examination* (MMSE) y el índice *Clinical Dementia Rating* (CDR). En el diagnóstico clínico, el CDR califica el grado de deterioro, utilizando un algoritmo estándar y sus posibles valores pueden ser: CDR=0: no deterioro; CDR=0.5: muy leve; CDR=1: leve; CDR=2: moderado; y CDR=3: demencia severa. Sin

¹En este trabajo consideramos instancia de datos o casos clínicos a cada una de las muestras con un conjunto de ocurrencias de variables únicas por sujeto presentadas en filas en la base de datos

² En nuestro contexto, consideramos variable a cada tipo de *feature* extraído de las base de datos

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

embargo, en el conjunto de datos de entrenamiento considerados en este estudio, solamente contamos con ocurrencias de CDR=0, 0.5 y 1. Para un mayor detalle de estos criterios, ver <http://www.adni-info.org/Scientists/ADNIGrant/ProtocolSummary.aspx>.

Tabla 4.1: Clasificación de los participantes según sus características clínicas

Variable clínica	Características Clínicas
Sujetos de control (sC)	Sujetos normales con valores de <i>Mini-Mental State Examination</i> (MMSE) [Folstein et al., 1975] entre 24 y 30 (inclusive), un índice de <i>Clinical Dementia Rating</i> (CDR) [Morris, 1993] de 0, y calificados como no depresivos, no MCI y no dementes
Sujetos con MCI (sMCI)	Pacientes MCI tienen un índice MMSE entre 24 y 30 (inclusive), reportan síntomas de problemas con la memoria, dificultades con la memoria objetiva, un CDR de 0.5, ausencia de niveles significativos de deterioro en otros dominios cognitivos, preservan las actividades esenciales de su vida diaria y poseen ausencia de demencia
sAD	Pacientes con Alzheimer tienen un índice de MMSE entre 20 y 26 (inclusive), un índice de CDR de 0.5 o 1.0, y reúne criterios del <i>National Institute of Neurological and Communicative Diseases and Stroke-Alzheimer's Disease and Related Disorders Association</i> (NINCDS/ADRDA) [McKham et al., 1984] para probable caso AD.

4.3.3 Definición de la variable de clasificación clínica

Para cada una de las instancias de ADNI fue agregada una variable de clasificación, establecida como variable de salida. Esta variable contenía información del diagnóstico clínico de cada caso realizado por ADN. De este modo, las instancias fueron etiquetadas como: sujetos sanos o de control (sC), sujetos con deterioro cognitivo leve (sMCI) o sujetos con AD leve (sAD), ver tabla 4.1.

Posteriormente, se identificó que algunos sujetos clasificados como sC fueron clasificados como sMCI o sAD en una visita posterior, y por otro lado, algunos sujetos clasificados como sMCI, fueron clasificados como sAD en una visita posterior. Ambos sucesos son explicados por el hecho de que el diagnóstico de MCI y AD es progresivo. En algunos casos, los sujetos pueden ser

clasificados como sC y estar desarrollando la enfermedad sin presentar los síntomas, e incluso, los pacientes con MCI pueden progresar o no para AD. En consecuencia, las instancias etiquetadas originalmente como sMCI fueron re-etiquetadas como sMCIc (para sujetos que se convirtieron a sAD) y como sMCIInc (para sujetos que no se convirtieron a AD). Esta subclasificación de los sujetos con MCI también ha sido asumida en otras investigaciones dedicadas a la predicción temprana de la conversión de MCI a AD [Cho et al., 2012; Cuingnet et al., 2011; Risacher et al., 2010].

4.3.4 Distribución de los datos

El siguiente paso fue la distribución de los casos clínicos en grupos destinados a tareas específicas para la construcción de los modelos de clasificación. Las instancias clasificadas como sC y sAD fueron asignadas para la tarea de diagnóstico, conformándose dos grupos: $sCsAD_{ADNI-train}$, utilizado para la fase de entrenamiento y validación; y $sCsAD_{ADNI-test}$, destinados para el testeo. Por otro lado, las instancias clasificadas como sMCI y sMCIInc fueron asignadas a la tarea de predicción de AD y se conformaron los grupos: $sMCI_{ADNI-train}$, utilizado para la fase de entrenamiento y validación; y $sMCI_{ADNI-test}$, utilizados para el testeo.

Asimismo, los cuatro grupos incluyeron 324 variables de entrada clasificadas como: variables morfométricas (318), demográficas (4) y neuropsicológica (2); y 1 variable de salida o de clasificación.

A continuación se detallan las principales características de los 4 grupos clínicos, etiquetados convenientemente para referencias en este estudio:

1. Grupos de entrenamiento y validación:

- $sCsAD_{ADNI-train}$: Grupo de entrenamiento y validación del modelo de diagnóstico de AD. Este grupo estuvo conformado por 1265 instancias: 631 instancias correspondientes a 216 sujetos de control (sC) y 634 instancias correspondientes a 287 sujetos con AD leve (sAD). Las variables morfométricas fueron calculadas a partir del análisis de imágenes de MR transversales ¹.
- $sMCI_{ADNI-train}$: Grupo de entrenamiento y validación del modelo de predicción de AD. Este grupo estuvo conformado por 894 instancias: 219 instancias correspondi-

¹Más adelante, este grupo fue utilizado para seleccionar un subconjunto de variables de mayor significancia para la discriminación entre sujetos de control (sC) y sujetos con AD (sAD)

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

entes a 106 sujetos, las cuales fueron etiquetadas como sMCIc y 675 instancias de 223 sujetos, etiquetas como sMCIInc ¹.

2. Grupos de test

- $sCsAD_{ADNI-test}$: Grupo de test del modelo de diagnóstico de AD, conformado por 1145 instancias: 591 instancias correspondientes a 204 sujetos de control (sC) y 554 instancias correspondientes a 246 sujetos con AD leve (sAD). Las variables morfométricas de este grupo fueron obtenidas de las MRIs estructurales del estudio longitudinal;
- $sMCI_{ADNI-test}$: Grupo de test del modelo de predicción de AD. Este grupo está conformado por 814 instancias clasificadas inicialmente como sMCI, las cuales corresponden a 297 pacientes e incluyen variables morfométricas de MRIs del estudio longitudinal. Asimismo, las instancias etiquetadas inicialmente como sMCI fueron clasificadas como sMCIc (pacientes con MCI quienes posteriormente se convirtieron a AD) y sMCIInc (pacientes con MCI quiénes no se convirtieron en AD).

En la tabla 4.2 se presenta la distribución de sujetos e instancias en los cuatro grupos descritos.

Tabla 4.2: Distribución de sujetos e instancias en los grupos de entrenamiento, validación y testeo de los modelos

	$sCsAD_{ADNI-train}$		$sMCI_{ADNI-train}$		$sCsAD_{ADNI-test}$		$sMCI_{ADNI-test}$	
	sC	sAD	sMCIc	sMCIInc	sC	sAD	sMCIc	sMCIInc
Nº sujetos	216	287	106	223	204	246	97	200
Nº instancias	631	634	219	675	591	554	181	633

4.3.5 Variables morfométricas basadas en imágenes de Resonancia Magnética Estructural

La obtención de estas variables estuvo relacionada con las mediciones morfométricas obtenidas por ADNI sobre las imágenes de MRIs estructurales (de ambos estudios transversal y longitudinal) de cada participante. Para la colección inicial de estas mediciones se tuvo en cuenta sólo aquellas que estuviesen presentes en los cuatro grupos de estudio: $sCsAD_{ADNI-train}$,

¹En una etapa posterior, estos datos fueron utilizados para seleccionar un subconjunto de variables de mayor significancia para la discriminación entre sujetos MCI que se convertirían a AD (sMCIc) de los sujetos con MCI, los cuales no se convertirían (sMCIInc)

$sMCI_{ADNI-train}$, $sCsAD_{ADNI-test}$ y $sMCI_{ADNI-test}$; siendo descartadas las mediciones etiquetadas como “*Unknown*” o con valores nulos. De este modo, fueron incluidas un total de 318 mediciones correspondientes a la zona cortical y subcortical del cerebro; y calculadas sobre 34 regiones de interés (ROIs) [Desikan et al., 2006]. Estas variables morfométricas corresponden a a cinco tipos de medición: *Cortical thickness average* (TA), *Cortical thickness standard deviation*(TS), *Surface area* (SA) y *Cortical volume* (CV) calculados para la zona neocortical; y el *Sub cortical volume* calculado para la zona no-neocortical. Asimismo, con el fin de conocer las diferencias en el tamaño del cerebro de cada sujeto y analizar el grado de la atrofia cerebral, el volumen intracranial total de cada sujeto (eTIV: *estimate of the total intracranial volume*) [Buckner, 2004] también fue considerado como variable.

En la tabla 4.3 se detalla la distribución de las 318 variables de acuerdo a la clasificación de las clase de ROIs así como los tipos de mediciones calculados para cada clase. Por otro lado, en las tablas 4.4 y 4.5 son presentadas las 318 variables morfométricas de acuerdo al tipo de lateralidad: hemisferio izquierdo (hi), hemisferio derecho (hd) y bilateral (bi) y los tipos de medición. En la siguiente etapa de esta metodología, un subconjunto de variables basadas en MRI (vMRI) sería seleccionado a partir de este grupo inicial de variables en base a su grado de utilidad y significancia para el diagnóstico entre sujetos sC y sAD.

Tabla 4.3: Tipos de mediciones morfométricas consideradas en este estudio

Clase ROI	Mediciones	Código	Nº variables
Neocortical=35	- Cortical thickness average	TA	68
	- Cortical thickness standard deviation	TS	68
	- Surface area	SA	70
	- Cortical volume	CV	69
No-Neocortical=28	-Subcortical volume	SV	42
Volumen Total=1	- eTIV: estimate of the total intracranial volume	ICV	1

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

Tabla 4.4: Variables morfométricas basadas en MRIs

	ROI	Lateralidad			Tipos de mediciones				
		Hi	Hd	bi.	CV	SA	TA	TS	SV
Variables de la zona cortical (n = 276)	Banks Superior Temporal Sulcus	✓	✓		✓	✓	✓	✓	
	Caudal Anterior Cingulate Cortex	✓	✓		✓	✓	✓	✓	
	Caudal Middle Frontal Gyrus	✓	✓		✓	✓	✓	✓	
	Cuneus Cortex	✓	✓		✓	✓	✓	✓	
	Entorhinal Cortex	✓	✓		✓	✓	✓	✓	
	Frontal Pole	✓	✓		✓	✓	✓	✓	
	Fusiform Gyrus	✓	✓		✓	✓	✓	✓	
	Inferior Parietal Cortex	✓	✓		✓	✓	✓	✓	
	Inferior Temporal Gyrus	✓	✓		✓	✓	✓	✓	
	Insula	✓	✓		✓	✓	✓	✓	
	Retrosplenial Cortex	✓	✓		✓	✓	✓	✓	
	Lateral Occipital Cortex	✓	✓		✓	✓	✓	✓	
	Lateral Orbital Frontal Cortex	✓	✓		✓	✓	✓	✓	
	Lingual Gyrus	✓	✓		✓	✓	✓	✓	
	Medial Orbital Frontal Cortex	✓	✓		✓	✓	✓	✓	
	Middle Temporal Gyrus	✓	✓		✓	✓	✓	✓	
	Paracentral Lobule	✓	✓		✓	✓	✓	✓	
	Parahippocampal Gyrus	✓	✓		✓	✓	✓	✓	
	Pars Opercularis	✓	✓		✓	✓	✓	✓	
	Pars Orbitalis	✓	✓		✓	✓	✓	✓	
	Pars Triangularis	✓	✓		✓	✓	✓	✓	
	Pericalcarine Cortex	✓	✓		✓	✓	✓	✓	
	Postcentral Gyrus	✓	✓		✓	✓	✓	✓	
	Posterior Cingulate Cortex	✓	✓		✓	✓	✓	✓	
	Precentral Gyrus	✓	✓		✓	✓	✓	✓	
	Precuneus Cortex	✓	✓		✓	✓	✓	✓	
	Rostral Anterior Cingulate Cortex	✓	✓		✓	✓	✓	✓	
	Rostral Middle Frontal Gyrus	✓	✓		✓	✓	✓	✓	
	Superior Frontal Gyrus	✓	✓		✓	✓	✓	✓	
	Superior Parietal Cortex	✓	✓		✓	✓	✓	✓	
	Superior Temporal Gyrus	✓	✓		✓	✓	✓	✓	
	Supramarginal Gyrus	✓	✓		✓	✓	✓	✓	
	Temporal Pole	✓	✓		✓	✓	✓	✓	
Transverse Temporal Cortex	✓	✓		✓	✓	✓	✓		
Hemisphere	✓	✓			✓				
Total Intracranial Volume			✓	✓					

Tabla 4.5: Variables morfométricas de la región no-cortical consideradas en este estudio

	ROI	Lateralidad			Tipos de mediciones				
		Hi	Hd	bi	CV	SA	TA	TS	SV
Variables de la zona sub-cortical (n = 42)	Accumbens Area	✓	✓						✓
	Amygdala	✓	✓						✓
	Caudate	✓	✓						✓
	Cerebellum Cortex	✓	✓						✓
	Cerebellum White Matter	✓	✓						✓
	Cerebral Cortex	✓	✓						✓
	Cerebral White Matter	✓	✓						✓
	Choroid Plexus	✓	✓						✓
	Hippocampus	✓	✓						✓
	Inferior Lateral Ventricle	✓	✓						✓
	Lateral Ventricle	✓	✓						✓
	Pallidum	✓	✓						✓
	Putamen	✓	✓						✓
	Thalamus	✓	✓						✓
	Ventricle Diencephalon	✓	✓						✓
	Vessel	✓	✓						✓
	Brain Stem			✓					✓
	Corpus Callosum Anterior			✓					✓
	Corpus Callosum Central			✓					✓
	Corpus Callosum Middle Anterior			✓					✓
	Corpus Callosum Middle Posterior			✓					✓
	Corpus Callosum Posterior			✓					✓
	Cerebrospinal Fluid			✓					✓
	Fourth Ventricle			✓					✓
	Non White Matter Hypointensities			✓					✓
	Optic Chiasm			✓					✓
	Third Ventricle			✓					✓
	White Matter Hypointensities			✓					✓

Las imágenes de MR estructural obtenidas por ADNI corresponden a múltiples adquisiciones de imágenes (entre 2 y 4 imágenes por sesión) de adultos con edades en el rango[18-96], las cuales fueron utilizadas previamente en diversas investigaciones dedicadas al desarrollo de biomarcadores de AD basados en neuroimagen [Cho et al., 2012; Chupin et al., 2009; Cui et al., 2011; Cuingnet et al., 2011; Desikan et al., 2009; Heckemann et al., 2011; Hua et al., 2008; Nestor et al., 2008; Okonkwo et al., 2010; Risacher et al., 2010; Westman et al., 2011]. Los participantes fueron sometidos a evaluaciones por imagen en sus visitas con intervalos diferenciados para los sujetos de control, los pacientes con AD leve y los pacientes con MCI. Estas evaluaciones fueron etiquetadas como: *Screen visit* (sc), *baseline visit* (base), *month 6* (m06), *month 12* (m12), *month 18* (m18), *month 24* (m24), *month 36* (m36) y *month 48* (m48). Mayor

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

detalle sobre la adquisición de MRIs estructurales de los participantes puede ser encontrada en: <http://adni.loni.ucla.edu/about-data-samples/image-data>.

Respecto al análisis morfométrico aplicado, estuvo basado en un flujograma de rutinas de procesamiento y análisis automático mencionadas en el capítulo 2, sección 2.4.2, para las cuales utilizaron el software libre de análisis de imagen Freesurfer [Fischl, 2012], ver sección 2.4. Utilizando Freesurfer, las ROIs del cerebro fueron obtenidas mediante la segmentación de la materia blanca subcortical y estructuras volumétricas profundas de la materia gris (incluyendo el hipocampo, la amígdala, núcleo caudado, putamen y los ventrículos) [Fischl et al., 2002, 2004a]; y la parcelación de la corteza cerebral en regiones basadas en las zonas gyral y suncal [Desikan et al., 2006; Destrieux et al., 2010; Fischl, 2012; Fischl et al., 2004a], entre otros métodos.

Los principales razones del estudio ADNI de utilizar Freesurfer en la obtención de los biomarcadores basados en imagen fueron:(1) Creación de modelos computarizados del cerebro a partir de MRIs, (2) Medición de las diversas propiedades morfométricas del cerebro de los participantes y (3) la normalización espacial de la zona cortical inter- sujetos basada en el alineamiento de los patrones del plegado cortical de los individuos para establecer correspondencia entre las regiones anatómicas homólogas. Con respecto a las versiones de Freesurfer utilizadas, en el estudio transversal de ADNI, el análisis de las imágenes fue realizado con la versión 4.3 y en el estudio longitudinal, el análisis de las imágenes fue realizada con la versión 4.4.

Básicamente, las técnicas aplicadas con Freesurfer consistieron en:

1. Corrección del movimiento, normalización de intensidades y corrección de inhomogeneidades de múltiples sesiones de MRIs para obtener una sola imagen volumen con un contraste alto de ruido [Reuter et al., 2010].
2. Localización del contorno de la materia gris/blanca (sWM: *white matter surface*) y localización del contorno de la materia gris/CSF (sGM:*grey matter surface*) a partir del primero;
3. Medición del espesor cortical, obtenida al calcular la distancia entre sWM y sGM en cada punto (por hemisferio) de toda la capa cortical;
4. Aplicación de técnicas de registro de imágenes para alinear los patrones del pliegue cortical y asignarlos probalísticamente a cada punto de la superficie cortical. De esta manera el neocórtex cerebral fue dividido en 32 ROIs basados en la zona gyral por cada hemisferio (izquierdo y derecho), de acuerdo las regiones propuestas por Desikan et al. [2006].

-
5. Adición de nuevas regiones no-cortical del cerebro, incluyendo el hipocampo, la amígdala, núcleo caudado, putamen y los ventrículos. Para conseguir esto, se utilizó un algoritmo que examina variaciones en las intensidades de los vóxeles y sus relaciones espaciales para clasificar las regiones no-neo-corticales sobre MRIs [Fischl et al., 2002].
 6. Obtención de mediciones morfométricas de las regiones neo-corticales y no-neo-corticales obtenidas en los pasos anteriores.

4.3.6 Variables clínicas: demográficas y neuropsicológicas

Estudios dedicados al progreso de la vejez, así como enfermedades del tipo MCI, AD y otros tipos de demencia, destacan un especial interés por analizar la influencia de las características socio-demográficas en el deterioro cognitivo. En [Stern et al., 1994] se estudió la relación entre el nivel de educación y el desarrollo de AD, encontrando que existe un riesgo de demencia mayor en sujetos con bajo nivel de educación. Anderson et al. [2007] y Wu et al. [2011] estudiaron en sujetos de Australia y Taiwan, respectivamente, la influencia de variables socio-demográficas como: edad, nivel de educación, lengua, nivel socioeconómico, sexo (junto a variables biológicas y clínicas) sobre el desempeño del MMSE (Mini-Mental State Examination) [Folstein et al., 1975]. Ambos estudios sumados a otras investigaciones [Moraes et al., 2010; Stern et al., 1994; van Hooren et al., 2007], han demostrado la influencia de estos factores sobre el deterioro cognitivo. En este estudio fueron incluidas inicialmente como variables demográficas: el sexo, edad, años de educación y *Handedness* (tipo de escritura), las cuales fueron documentadas en la primera visita de los participantes (*screen visit*). Las diferencias entre estas variables fue testada en los cuatro grupos sin encontrarse diferencias significativas en la edad, sexo y años de educación. Estas variables demográficas han sido consideradas en otras investigaciones de AD, realizadas sobre los datos de ADNI [Cui et al., 2011; Cuingnet et al., 2011; Landau et al., 2010] y otros base de datos sobre AD.

Con respecto a las variables neuropsicológicas, sólo fueron incluidas las variables MMSE y CDR [Morris, 1993], documentadas por ADNI en cada una de las visitas de los participantes ¹. Ambas variables tienen estrecha relación en el criterio de inclusión/exclusión de ADNI, ver sección 4.3.3. En la tabla 4.6 se presenta con detalle las variables demográficas y neuropsicológicas de los participantes considerados en este estudio.

Adicional a las variables demográficas y neuropsicológicas consideradas en nuestro estudio, ADNI incluye otras variables referentes al reclutamiento, exámenes físicos, concentración

¹Adicionalmente a estas dos características cognitivas, ADNI proporciona otro tipo de test cognitivos, los cuales no fueron considerados en este estudio

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

Tabla 4.6: Variables clínicas: características demográficas y neuropsicológicas de los grupos $sCsAD_{ADNI-train}$, $sMCI_{ADNI-train}$, $sCsAD_{ADNI-test}$ y $sMCI_{ADNI-test}$

	Entrenamiento y validación				Test			
	sC	sMCIc	sMCIc	sAD	sC	sMCIc	sMCIc	sAD
Nº sujetos	216	223	106	287	204	200	97	246
Mº instancias	631	675	219	634	591	633	181	554
Edad	78.68 ± 5.35	77.74 ± 7.36	75.89 ± 7.25	77.18 ± 7.56	78.65 ± 5.32	77.70 ± 7.08	75.91 ± 7.32	77.23 ± 7.50
% Femenino	46.76	33.63	39.62	41.81	45.10	36.00	37.11	41.06
% Masculino	53.24	66.37	60.38	58.19	54.90	64.00	62.89	58.94
Educación (años de educación)	15.93 ± 2.85	15.64 ± 3.11	15.92 ± 2.71	15.28 ± 2.98	15.99 ± 2.74	15.80 ± 3.02	15.88 ± 2.77	15.30 ± 2.98
MMSE	26.16 ± 8.40	23.83 ± 8.78	24.03 ± 6.25	19.22 ± 7.49	25.92 ± 8.80	23.65 ± 9.01	23.82 ± 6.48	19.17 ± 7.38

de biomarcadores de CSF, etc, las cuales no fueron considerados en este estudio. Para mayor información visitar: <http://adni.loni.ucla.edu/about-data-samples/clinical-data/>.

4.3.7 Preprocesamiento de las variables

Antes de aplicar los enfoques de selección de variables, las instancias de los cuatro grupos fueron preprocesados a través de tres tareas fundamentales aplicadas a los datos de los cuatro grupos clínicos $sCsAD_{ADNI-train}$, $sCsAD_{ADNI-test}$, $sMCI_{ADNI-train}$ y $sMCI_{ADNI-test}$:

1. Conversión de los datos de entrada y las clases de los targets a formato de números reales, esto debido a la presencia de variables con valores no-num-ericos: por ejemplo los valores de la variable Género que inicialmente eran F (femenino) y M (masculino) fueron convertidos a 0 y 1, respectivamente. Asimismo, los valores de los targets de las instancias de los grupos $sCsAD_{ADNI-train}$ y $sCsAD_{ADNI-test}$ fueron reescritos hacia los valores [0,1] de la siguiente manera: sC=0; sAD=1. El mismo procesamiento fue aplicado para los targets de los grupos $sMCI_{ADNI-train}$ y $sMCI_{ADNI-test}$, reescribiendo su clases como: sMCIc=0; sMCIc=1.
2. Conversión a formado de matriz dispersa, en vista de existía un gran número de campos con valores de cero, las cuatro matrices de datos fueron convertidas a formato de matrices dispersas.
3. Escalamiento de los datos, debido a la notable diferencia entre las escalas de las variables, encontrándose casos en dónde algunas escalas estaban en el rango de las milésimas, y en otros casos estaban en los miles. En este caso, tomando como base los rangos inferior y superior del grupo $sCsAD_{ADNI-train}$ por cada una de sus columnas de datos (correspondiente a una variable de entrada), los valores de cada columna fueron escalados linealmente a los rangos [0,1]. El mismo procedimiento fue aplicado a los parámetros de los otros tres

grupos, considerando para su escalamiento, los rangos de cada variable calculados sobre el primer grupo.

El escalamiento de los datos, antes de utilizarlos en el aprendizaje de las SVM trae varias ventajas, ver sección 3.3.1. En nuestro caso, el escalamiento de los datos fue útil para resolver las diferencias del rango numérico de los volúmenes de la zona no-neocortical del cerebro, cuyos valores iniciales oscilaban en los miles, con respecto a las otras mediciones como el espesor de las zonas del neocórtex cerebral, las cual es encontraban en el rango de los decimales. Otra de las ventajas del escalamiento de los datos fue evitar las dificultades numéricas durante el aprendizaje de las SVMs.

4.4 Selección de variables

El proceso de clasificación y selección consistió en eliminar las variables redundantes combinando dos enfoques: un método de filtrado basado en un *ranking* de las variables para calificarlas y ordenarlas de acuerdo a un criterio, seguido de un método de selección de variables basado en SVM. Los métodos de selección basados en SVM son: i) SVM *Wrapper*, ii) F-score y iii) la combinación de ambos, F-score + SVM *Wrapper*. La combinación de estos enfoques ha sido aceptada como una técnica recomendada para los procedimientos de selección de características para resolver problemas de clasificación de datos [Guyon and Elisseeff, 2003]. El enfoque de selección fue aplicado por cada tipo de variables, obteniéndose 3 subconjuntos de variables morfométricas y 3 subconjuntos de variables clínicas.

4.4.1 Análisis inicial de las variables. Curvas ROC

Con el fin de conocer la influencia de cada una de las variables sobre los resultados de clasificación clínica (variable de salida), así como verificar la existencia de variables poco relevantes y redundantes que podrían afectar el desempeño de los modelos de clasificación, el análisis del ROC fue utilizado para evaluar cada una de las variables originales. Diversos investigadores consideran que el AUC debería ser el criterio principal para evaluar la efectividad de los biomarcadores de diagnóstico, es así, que la efectividad de uso de esta medición estadística ha sido demostrada en diversas publicaciones, varias de ellas enfocados en cardiología y neurología [Cook, 2007; Cui et al., 2011; Doniger et al., 2005; Faraggi and Reiser, 2002; Pencina et al., 2008]. De este modo, cada variable del grupo sCsAD_{ADNI-train} fue categorizada de acuerdo al valor AUC, siguiendo la clasificación presentada en 3.2. Asimismo, en la figura 4.2 se puede observar gráficamente las curvas ROC obtenidas por las variables morfométricas y clínicas. Obsérvese la existencia de

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

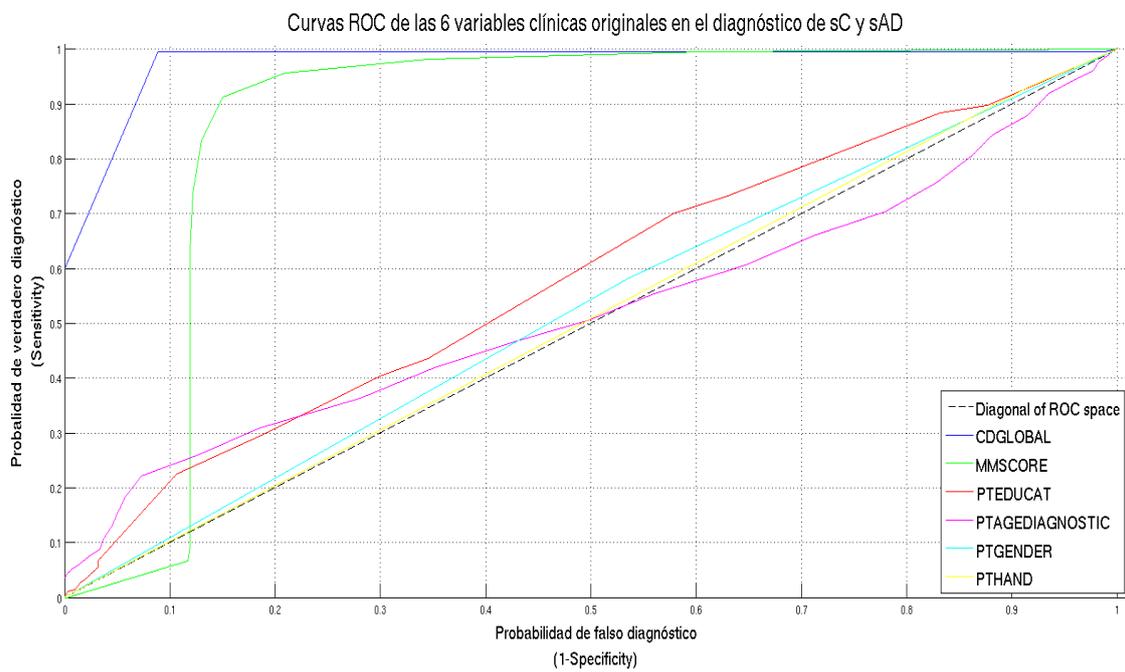
variables con valores AUC ≤ 0.6 (calificadas como ‘malas’ y ‘muy malas’) con curva ROC por debajo de la línea diagonal del espacio ROC, las cuales podrían afectar la construcción óptima de los modelos. Por otro lado, también se observa la existencia de variables con valores AUC altos como en caso de la variable morfométrica *Volume(WMParcellation)ofLeftHippocampus* (ST29SV) con un AUC de 0.8860, lo cual nos indica que dicha variable tiene la capacidad de discriminar un individuo seleccionado al azar en el grupo sAD con un valor de prueba mayor a la de un individuo elegido al azar entre el grupo sC en un 88.6% de las veces. Si la variable en cuestión no fuera útil para distinguir entre los dos grupos, su AUC hubiese sido igual a 0,5, en cuyo caso su curva ROC hubiese coincidido con la línea diagonal. Asimismo, en el caso que existiese una separación perfecta de los valores de los dos grupos, es decir, sin solapamiento de las distribuciones, el AUC de la variable hubiese sido igual a 1.0 y su curva ROC hubiese estado ubicada en la esquina superior izquierda del gráfico [Lasko et al., 2005]. Mayor descripción del análisis ROC y los valores AUC de variables predictoras y modelos de clasificación ha sido presentada en la sección 3.3.3.2.

En las tablas 3 y 4 del apéndice 6.5.2 son presentados los resultados completos de los valores AUC de las variables morfométricas y clínicas.

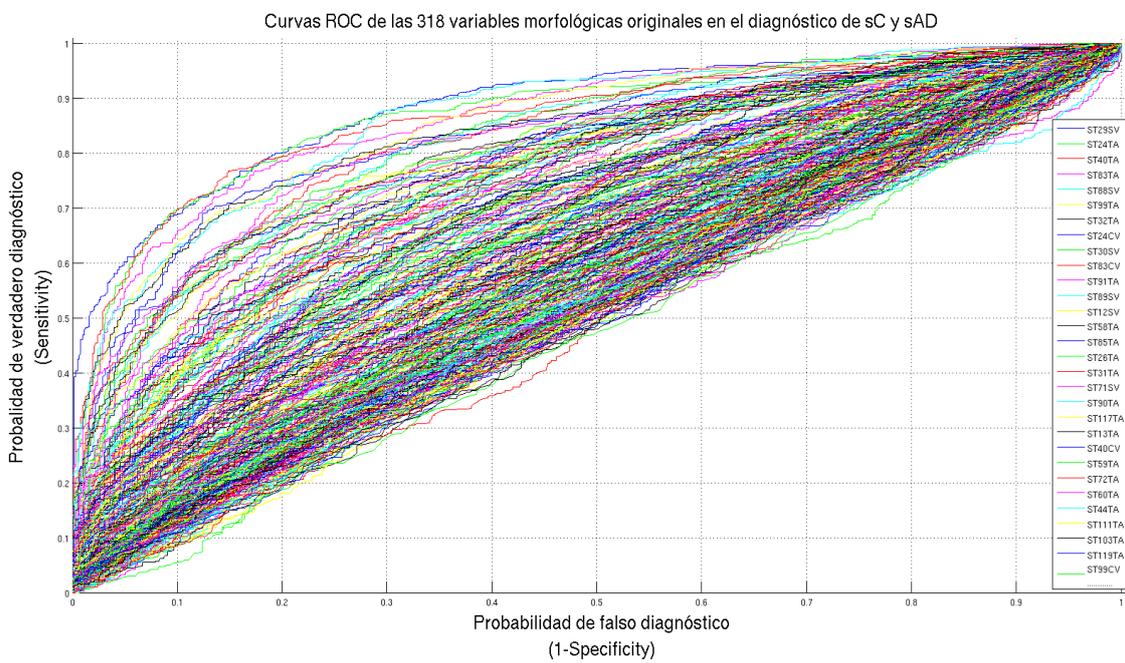
En las siguientes etapas nuestro estudio estuvo enfocado en la depuración de las variables redundantes y poco discriminativas, tratando de reducir aquellas con bajo AUC. La identificación de este subconjunto de variables no sólo ayudó a mejorar la exactitud en los resultados, también fue muy útil para la visualización e interpretación de los datos, reduciendo las mediciones y requerimientos de almacenamiento, reduciendo los tiempos de entrenamiento y definiendo el curso de la dimensionalidad para mejorar el desempeño de los modelos.

4.4.2 Ranking de variables

Una vez escalados los datos, el siguiente paso consistió en aplicar el *ranking* de variables, el cual fue utilizado como un método de filtrado (asumido como un paso de preprocesamiento) [Ron and George H, 1997]. Con este método, todas las variables fueron observadas, considerando las características intrínsecas de los datos con respecto a las evaluaciones clínicas (clasificación de los sujetos en sC o sAD). El ranking de ambos tipos de variables fue aplicado con el método de Mínima Redundancia y Máxima Relevancia (mRMR, por sus siglas en inglés *minimum redundancy and maximum relevance*) propuesto por Ding and Peng [2005]; Peng et al. [2005]. Este método calcula la información mutua de dos variables en base a una función de densidad probabilística, de este modo el *ranking* es obtenido al optimizar simultáneamente los criterios de máxima relevancia y mínima redundancia. Por convención asumimos que una alta puntuación indicaba



(a)



(b)

Figura 4.2: Gráfica de las curvas ROC (AUC) obtenidas por las 324 variables originales: (a) 6 variables clínicas; (b) 318 variables morfológicas originales

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

una variable importante y todas las variables fueron ordenadas de forma decreciente de $S(i)$. Para más detalle de este método ver la sección 3.4.1.1).

Los resultados completos del *ranking* de variables morfométricas y clínicas son presentados en la segunda columna de las tablas 3 y 4 del apéndice 6.5.2.

4.4.3 Selección del subconjunto de variables

Como fue descrito en la sección 3.4.2, SVM, además de ser utilizado como un método de clasificación, es también utilizado para la selección de variables basándose en la exactitud (ACC) [Kohavi and Provost, 1998] en la predicción ¹ obtenida al ser incluida (o excluida) una respectiva variable. Sin embargo, SVM no obtiene directamente la relevancia de las características, para ello es necesario complementar el método con otros cálculos. En este trabajo, la etapa de selección de variables fue aplicada sobre cada clase de variables (ordenadas de acuerdo al ranking obtenido en al etapa anterior) aplicando los siguientes métodos de selección basados en SVM:

1. Método F-score_SVM: Este método consistió en ordenar las variables de manera decreciente de acuerdo a un valor *F-score* (Fisher score) [Chen and Lin, 2006] calculado para cada una de ellas calculado con la ecuación 3.12. El valor *F-score* está basado en la discriminación de una variable individual con respecto a las clases, ver sección 3.4.1.2 para mayor detalle. El conjunto inicial de variables ordenadas es dividido en subconjuntos de variables con diferentes tamaños y cada subconjunto es utilizado entrenar un modelo clasificador basado en SVM. Finalmente es seleccionado el subconjunto de variables con las cuales el modelo alcanza la mayor exactitud en la predicción. Este método de selección fue aplicado utilizando la librería “fselect” de LIBSVM [Chang and Lin, 2011].
2. Método SVM Wrapper: Este método se basó en el análisis de la exactitud de predicción de un modelo clasificador al incluir progresivamente en su entrenamiento cada variable, según el ranking obtenido con el método mRMR. De este modo, el modelo basado en SVM empezó a ser entrenado con un subconjunto conformado por la primera variable del ranking, y el criterio de inclusión/exclusión de las siguientes variables dependió de la exactitud del modelo durante el test, al considerar la nueva variable.
3. Método Fscore_SVM + Wrapper: Con este método, la selección de variables se realizó a través de la combinación de los dos métodos de selección anteriores. De este modo, el

¹En este trabajo el término exactitud es utilizado para hacer referencia al porcentaje de las correctas predicciones hechas por un modelo sobre un dataset. En consecuencia, exactitud de la predicción (“Predictive accuracy”) se refiere al porcentaje de todas las muestras o instancias correctamente clasificadas en un dataset de test.

método SVM Wrapper fue aplicado sobre las variables previamente seleccionadas con la librería F-score. Del mismo modo, en cada paso del algoritmo, el criterio de inclusión o exclusión de una variable se basó en la máxima exactitud de clasificación del modelo al ser testeada con dicha variable.

Como resultado de la aplicación de los tres métodos, fueron obtenidos seis subconjuntos de variables distintos, con los cuales se construyeron modelos de clasificación auxiliares:

1. Modelo $SVM_{vMRI-Wrapper}$, construido con el subconjunto de variables morfométricas seleccionadas con el método SVM Wrapper.
2. Modelo $SVM_{vMRI-Fscore}$, construido con el subconjunto de variables morfométricas seleccionadas con el método Fscore.
3. Modelo $SVM_{vMRI-FscoreWrapper}$, construido con el subconjunto de variables morfométricas seleccionadas con la combinación de los métodos Fscore y SVM Wrapper.
4. Modelo $SVM_{vClin-Wrapper}$, construido con el subconjunto de variables clínicas (demográficas y neuropsicológicas) seleccionadas con el método SVM Wrapper.
5. Modelo $SVM_{vClin-Fscore}$, construido con el subconjunto de variables clínicas (demográficas y neuropsicológicas) seleccionadas con el método Fscore.
6. Modelo $SVM_{vClin-FscoreWrapper}$, construido con el subconjunto de variables clínicas (demográficas y neuropsicológicas) seleccionadas con la combinación de los métodos Fscore y SVM Wrapper.

El entrenamiento y test de estos modelos fue realizado sobre los grupos $sCsAD_{ADNI-train}$ y $sCsAD_{ADNI-test}$, respectivamente, considerando únicamente las variables seleccionadas en cada método. La selección del conjunto definitivo sería realizada evaluando el número de veces que fueron seleccionadas cada una de las variables; así como las métricas estadísticas calculadas a partir de la matriz de confusión [Kohavi and Provost, 1998] de cada clasificador, tales como exactitud, precisión, sensibilidad, especificidad; así como el valor AUC de cada variable.

4.4.3.1 Configuración del clasificador SVM

Antes de utilizar los tres métodos de selección de variables fue necesario configurar adecuadamente los clasificadores SVM utilizados por estos métodos durante la construcción del algoritmo

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

de aprendizaje y la evaluación de la ACC alcanzada con cada variable analizada. Esta configuración implicó determinar el tipo de kernel a utilizar y los parámetros más óptimos para su configuración, ver sección 3.3. De este modo fueron realizados dos análisis: (1) Se determinó el tipo de problema de clasificación a resolver, identificándose que en ambos problemas de clasificación sAD vs. sC, no existía una separación lineal de las clases, ver figura 4.3 (en la sección 3.2 fueron descritos los conceptos aplicados por SVM para la solución de problemas no lineales); y, (2) La proporción entre el número de instancias y el número de variables, determinándose que para el grupo de entrenamiento y validación ($sCsAD_{ADNI-train}$), el número de instancias era mucho mayor que el número de variables (1265 instancias vs. 29 variables). Considerando ambas características se procedió a determinar la función kernel más adecuada para este problema de clasificación. Fueron estudiados y aplicados los tipos de funciones más utilizadas para este tipo de problema, dentro de las cuales tenemos: polinomial, radial y sigmoid; siendo establecido que la función kernel *Radial Basic Function (RBF) kernel* se ajustaba mejor a nuestro problema. De este modo, todos los modelos fueron construidos utilizando el kernel RBF y validación cruzada *cross validation* (CV) para la búsqueda de los mejores parámetros de configuración SVM (ver sección 3.2 para la ampliación de estos conceptos).

4.4.3.2 Selección de variables morfométricas

Método F-score_SVM. Con este método todas las variables morfométricas fueron ordenadas de acuerdo a su valor F-score y un subconjunto de 159 variables fueron seleccionadas como más descriptivas, en función de la mayor exactitud de predicción en el test obtenida por el modelo auxiliar $SVM_{vMRI-Fscore}$. Durante la etapa de test, el modelo obtuvo la matriz de confusión conformada por: (1) Verdaderos positivos(TP)=554, es decir, 554 instancias originalmente clasificadas como sAD fueron clasificadas correctamente como sAD; (2) verdaderos negativos(TN)=573, es decir 573 instancias originalmente clasificadas como sC fueron clasificadas correctamente como tal; (3) falsos positivos(FP)=18; es decir 18 instancias diagnosticadas como sC fueron incorrectamente clasificadas como sAD y (4) falsos negativos(FN)=0, es decir es decir ninguna instancia diagnosticadas como sAD fueron incorrectamente clasificada como sC. A partir de la matriz de confusión fueron calculados los siguientes indicadores estadísticos: exactitud= 98.43%, precisión = 96.85%, sensibilidad=100%, especificidad=96.95%. En el capítulo 3 han sido descritas estas mediciones estadísticas en el contexto del diagnóstico de enfermedades.

En la tabla 4.7 son presentadas las 10 primeras variables seleccionadas con sus respectivos valores F-score y AUC, Obsérvese que con este método, las variables seleccionadas son ordenadas de manera decreciente, de acuerdo a su valor F-score. En la figura 4.4 se presentan las gráficas de

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

la clasificación AUC de estas variables para los grupos de entrenamiento y test. Obsérvese que la variable *Cortical Thickness Average of RightParahippocampal* (ST103TA) alcanzó el mayor F-score de 0.7329 y un valor óptimo de AUC de 0.7633.

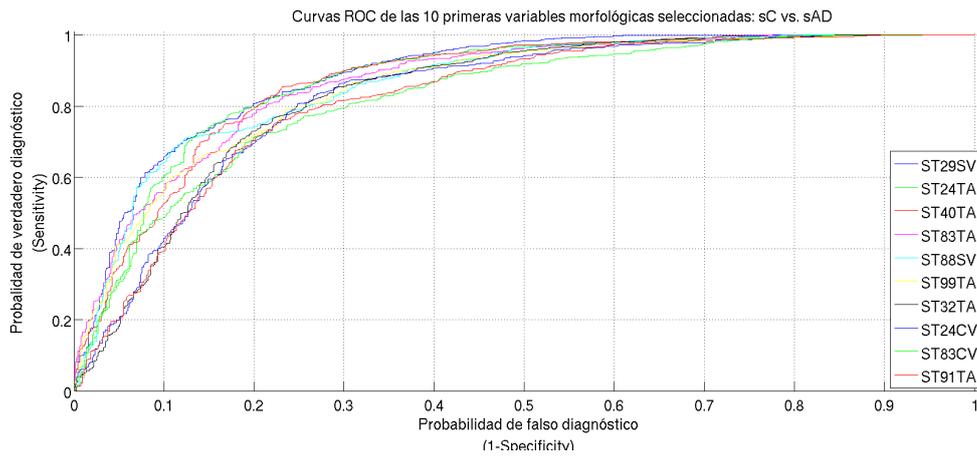
En la sexta columna de la tabla 3, apéndice 6.5.2, son presentadas las 159 variables morfométricas seleccionadas por este método.

Tabla 4.7: Lista de las 10 primeras variables morfométricas seleccionadas con el método F-score

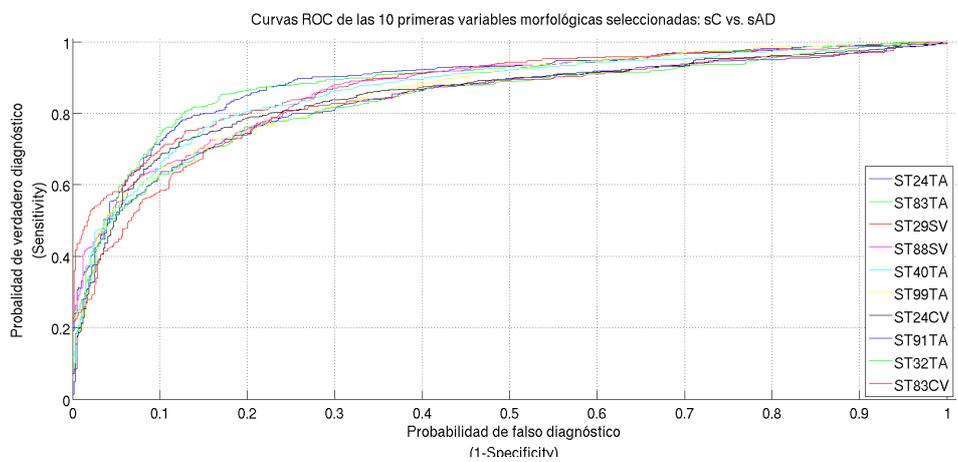
Ranking	Código	Nombre variables	F-score	AUC
12	ST103TA	Cortical Thickness Average of RightParahippocampal	0.7633	0.7329
95	ST101SV	Volume (WM Parcellation) of RightPallidum	0.5700	0.6709
217	ST102TS	Cortical Thickness Standard Deviation of RightParacentral	0.5109	0.6394
148	ST102SA	Surface Area of RightParacentral	0.5491	0.6057
88	ST104TS	Cortical Thickness Standard Deviation of RightParsOpercularis	0.5585	0.5966
132	ST103TS	Cortical Thickness Standard Deviation of RightParahippocampal	0.5589	0.5345
191	ST102TA	Cortical Thickness Average of RightParacentral	0.6248	0.4565
314	ST10CV	Volume (Cortical Parcellation) of Icv	0.5298	0.4283
31	ST103CV	Volume (Cortical Parcellation) of RightParahippocampal	0.7515	0.4283
187	ST105TS	Cortical Thickness Standard Deviation of RightParsOrbitalis	0.5126	0.4054

SVM Wrapper El método SVM Wrapper fue aplicado sobre los datos del grupo $sCsAD_{ADNI-train}$, cuyas variables fueron reordenadas de acuerdo a los resultados de la etapa de ranking. En la primera iteración del método, el modelo auxiliar $SVM_{vMRI-Wrapper}$ fue entrenado y validado considerando únicamente la variable *Cortical Thickness Average of LeftEntorhinal* (ST24TA), la cual alcanzó la mayor puntuación en el ranking mRMR, ver tabla 3, segunda columna. Asimismo, las demás variables fueron analizadas y agregadas/excluidas progresivamente, de acuerdo este ranking. En cada paso del algoritmo, se realizó el ajuste de los parámetros C y γ del modelo auxiliar y se verificó la exactitud alcanzada por el modelo durante el test. De este modo, sólo aquellas variables que mejoraban la exactitud de predicción del modelo fueron integradas definitivamente en este subconjunto de variables. Como resultado, 126 variables morfométricas fueron seleccionadas, a partir de las cuales el modelo $SVM_{vMRI-Wrapper}$ alcanzó el mayor desempeño. Durante el test con el grupo $sCsAD_{ADNI-test}$, considerando sólo las variables morfométricas seleccionadas, el modelo obtuvo la matriz de confusión conformada por los siguientes valores: TP=554, TN=587, FP=4 y FN=0. A partir de estos valores fueron calculados las siguientes mediciones estadísticas: exactitud= 96.65%, precisión = 99.28%, sensibilidad=100%, especificidad=99.32%.

En la séptima columna de la tabla 3, apéndice 6.5.2, pueden observarse las 126 variables morfométricas seleccionadas por este método.



(a)



(b)

Figura 4.4: Gráfica de las curvas ROC (AUC) obtenidas por las 10 primeras variables morfométricas seleccionadas con F-score: (a) Resultados sobre el grupo de entrenamiento y validación $sCsAD_{ADNI-train}$. (b) Resultados sobre el grupo de test $sCsAD_{ADNI-test}$

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

F-score_SVM + SVM Wrapper. Este método fue aplicado utilizando los datos de los grupo sCsAD_{ADNI-train} y sCsAD_{ADNI-test}, considerando únicamente las 159 variables morfométricas seleccionadas por el método F-score_SVM. Asimismo, los datos de ambos grupos fueron reordenados de acuerdo al orden decreciente del valor F-score de cada una de ellas. El algoritmo de selección fue iniciado con la variable *Cortical Thickness Average of RightParahippocampal* (ST103TA), la cual alcanzó la mayor puntuación, siendo agregadas y analizadas las demás variables progresivamente de acuerdo a su calificación con el ranking mRMR. En cada paso del algoritmo, el modelo fue entrenado con el subconjunto actual de variables y los parametros SVM (C y γ) fueron ajustados. Finalmente, 76 variables morfométricas fueron seleccionadas, a partir de las cuales el modelo SVM_{vMRI-FscoreWrapper} alcanzó la mayor exactitud de predicción sobre los datos del grupo de test sCsAD_{ADNI-test}. Durante la etapa final de testeo, el modelo obtuvo una matriz de confusión conformada por: TP=554, TN=585, FP=6 y FN=0. A partir de estos valores, fueron calculados las siguientes métricas estadísticas del modelo: exactitud= 99.48% (1139/1145), precisión = 98.93 %, sensibilidad%(554/554), especificidad=98.98%. En la octava columna de la tabla 3, apéndice 6.5.2 pueden observarse las 76 variables seleccionadas.

4.4.3.3 Selección de variables clínicas

Método F-score_SVM. Este método seleccionó como más discriminantes a tres de las seis variables clínicas originales, a partir de las cuales el modelo SVM_{vClin-Fscore} alcanzó la máxima exactitud de predicción del grupo sCsAD_{ADNI-test}, ver sexta columna de la tabla 4, apéndice 6.5.2. Durante la etapa de test del modelo con las 3 variables clínicas seleccionadas se obtuvo como matriz de confusión: TP=549, TN=568, FP=23 y FN=5. A partir de la matriz de confusión fueron calculados las siguientes indicadores: exactitud= 97.55%, precisión = 95.98%, sensibilidad=99.10%, especificidad=96.11%.

Método SVM Wrapper. Para la selección de las variables clínicas utilizando este método, el procedimiento fue similar al descrito para las variables morfométricas . El algoritmo fue iniciado con la variable CDGLOBAL(*Global Clinical Dementia Rating*), y las demás variables fueron agregadas según el *ranking* obtenido. Como resultado, este método seleccionó las seis variables clínicas iniciales, a partir de las cuales el modelo SVM_{vClin-Wrapper} alcanzó el mayor porcentaje de exactitud de predicción del grupo de test sCsAD_{ADNI-test}. Durante el test del modelo con las 1145 instancias del grupo de test sCsAD_{ADNI-test} considerando sólo las variables seleccionadas, se obtuvo como matriz de confusión: TP=548, TN=573, FP=18 y FN=6.

A partir de estos valores fueron deducidas las siguientes métricas estadísticas del modelo: exactitud= 97.90% (1121/1145), precisión = 96.82 % (548/566), Recall=98.92%(548/554), especificidad=96.95%(573/591). Los resultados de este método son presentados en la séptima columna de la tabla 4, apéndice 6.5.2.

F-score_SVM + SVM Wrapper. El procedimiento de este método para la selección de las variables clínicas fue similar al descrito para las variables morfométricas. El método SVM Wrapper fue aplicado sobre las tres variables clínicas seleccionadas y ordenadas por el método F-score_SVM. El algoritmo fue iniciado con la variable PTGENDER (Género del paciente) y las dos variables restantes fueron agregadas y analizadas progresivamente. Como resultado, el método SVM Wrapper seleccionó a las tres variables iniciales, ver octava columna de la tabla 4, apéndice 6.5.2. Con estas variables el modelo $SVM_{vClin-FscoreWrapper}$ alcanzó la mayor exactitud en la clasificación del grupo $sCsAD_{ADNI-test}$. Asimismo, en la etapa de test, el modelo obtuvo la matriz de confusión: TP=549, TN=568, FP=23 y FN=5; a partir de la cual se calcularon los indicadores estadísticos del modelo: exactitud= 97.55%, precisión = 95.98, sensibilidad=99.10, especificidad=96.11. Obsérvese que los valores de la matriz de confusión y los indicadores estadísticos son iguales a los resultados obtenidos por el modelo $SVM_{vClin-Fscore}$, lo cual se debe a que en ambos casos se seleccionaron las mismas variables clínicas, ver sección 4.4.3.3.

4.4.4 Discriminación entre los métodos de selección de variables.

Con la aplicación de las tres técnicas de selección de variables se obtuvieron subconjuntos de variables distinto por lo que fue necesario definir un criterio de discriminación entre ellos. Con cada uno de los métodos se consiguió reducir el número de variables originales, asimismo, existían variables que habían sido seleccionadas más de una vez y otras que no fueron seleccionadas por ningún método. Sin embargo, a fin de garantizar la robustez de los modelos frente a nuevos casos clínicos, era necesario obtener un subconjunto definitivo de características, las cuales no fueran redundantes.

A partir de estas consideraciones, fue necesario evaluar detalladamente cada subconjunto de variables y los modelos auxiliares construidos a partir de éstas, a fin de establecer un criterio definitivo de elección de las variables predictoras. Como se mencionó en la sección 3.3.3.2, los indicadores estadístico basados en la matriz de confusión son comúnmente utilizados para evaluar el poder de discriminación de los modelos y variables predictoras. En base a esto, fueron calculados los indicadores estadísticos: Exactitud(ACC), sensibilidad (TPR), especificidad (TNR), precisión (PRE) y AUC para los seis modelos auxiliares, ver figura 4.5; observándose que los

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

resultados fueron muy óptimos en todos los casos a pesar de las pequeñas diferencias. En la tabla 4.8 se resumen estos resultados.

Entonces, un primer criterio fue seleccionar como mejores subconjuntos de variables morfométricas y clínicas, aquéllos subconjuntos que permitieron construir modelos con mejores indicadores de desempeño. En este caso, los mejores resultados fueron obtenidos por el modelo $SVM_{vMRI-Wrapper}$ con las 126 variables morfométricas seleccionadas, y el modelo $vClin_Wrapper$, con las 6 variables clínicas seleccionadas. Sin embargo, en ambos subconjuntos, el número de variables seleccionadas aún era elevado y existía un gran número de variables con bajas calificaciones de $AUC \leq 0.6$ (clasificadas como “malas”, según clasificación AUC de la tabla 3.2). Estas mismas características se presentaron en los otros subconjuntos de variables.

Para abordar el riesgo de seleccionar un gran número de variables con estas características y garantizar la construcción de modelos robustos, se estableció un segundo criterio de discriminación. Se procedió a calificar cada una de las 324 variables originales de acuerdo al número de veces que fueron seleccionadas. De este modo, sólo serían incluidas las variables morfométricas y clínicas seleccionadas a la vez por los tres métodos (calificación igual a 3). Como resultado, fueron seleccionadas 29 variables conformadas por: i) Subconjunto conformado por 26 variables morfométricas y ii) Subconjunto conformado por 3 variables clínicas. En la tabla 4.9 se presentan las 29 variables seleccionadas con sus respectivos valores AUC y en la figura 4.6 se presenta gráficamente las curvas ROC de cada una de estas variables. Obsérvese que la mayoría de las variables seleccionadas presentan una curva ROC aceptable, por encima de la línea diagonal, con valores $AUC > 0.6$. Los resultados completos de calificación de las 318 variables morfométricas y las 6 variables clínicas son presentados en la décima columna de las tablas 3 y 4, apéndice 6.5.2, respectivamente.

En la siguiente etapa de la metodología, el desempeño de estas 29 variables sería analizado al construir los modelos finales de diagnóstico y predicción.

Tabla 4.8: Indicadores estadísticos de los seis modelos clasificadores auxiliares durante la etapa de test

Modelo	Entrada	Salida	ACC	PRE	TPR	TNR	AUC
$vMRI_{Fscore}$	318	159	98.43 (1127/1145)	96.85 (554/572)	100 (554/554)	96.95 (573/591)	0.9848
$SVM_{vMRI-Wrapper}$	318	126	99.65 (1141/1145)	99.28 (554/558)	100 (554/554)	99.32 (587/591)	0.9966
$vMRI_{Fscore+Wrapper}$	159	76	99.48 (1139/1145)	98.93 (554/560)	100 (554/554)	98.98 (585/591)	0.9949

Continued on next page

Tabla 4.8 – continued from previous page

Modelo	Entrada	Salida	ACC	PRE	TPR	TNR	AUC
vClin F_{score}	6	3	98.25 (1125/1145)	97.51 (548/562)	98.92 (548/554)	97.63 (577/591)	0.9827
vClin $Wrapper$	6	6	97.90 (1121/1145)	96.82 (548/566)	98.92 (548/554)	96.95 (573/591)	0.9794
vClin $F_{score}+Wrapper$	3	3	97.55 (1117/1145)	95.98 (549/572)	99.10 (549/554)	96.11 (568/591)	0.9760

Tabla 4.9: Lista del subconjunto definitivo de 29 variables predictoras seleccionadas

mRMR	Codigo	Nombre	Tipo	AUC	F-score
1	ST24TA	Cortical Thickness Average of LeftEntorhinal	Morfológica	0.8704	0.0322
7	ST13TA	Cortical Thickness Average of LeftBankssts	Morfológica	0.7916	0.0283
12	ST103TA	Cortical Thickness Average of RightParahippocampal	Morfológica	0.7633	0.7329
14	ST12SV	Volume (WM Parcellation) of LeftAmygdala	Morfológica	0.8056	0.0768
23	ST114TA	Cortical Thickness Average of RightRostralMiddleFrontal	Morfológica	0.751	0.1801
31	ST103CV	Volume (Cortical Parcellation) of RightParahippocampal	Morfológica	0.7515	0.4283
45	ST34TS	Cortical Thickness Standard Deviation of LeftIsthmusCingulate	Morfológica	0.651	0.0588
47	ST57TA	Cortical Thickness Average of LeftSuperiorParietal	Morfológica	0.704	0.0900
55	ST118TA	Cortical Thickness Average of RightSupramarginal	Morfológica	0.7551	0.0951
58	ST5SV	Volume (WM Parcellation) of CorpusCallosumMidPosterior	Morfológica	0.6527	0.0347
70	ST105TA	Cortical Thickness Average of RightParsOrbitalis	Morfológica	0.6849	0.2852
83	ST73TS	Cortical Thickness Standard Deviation of RightCaudalAnteriorCingulate	Morfológica	0.5534	0.0372
87	ST30SV	Volume (WM Parcellation) of LeftInferiorLateralVentricle	Morfológica	0.8267	0.0722
88	ST104TS	Cortical Thickness Standard Deviation of RightParsOpercularis	Morfológica	0.5585	0.5966
89	ST115TA	Cortical Thickness Average of RightSuperiorFrontal	Morfológica	0.7536	0.2771
95	ST101SV	Volume (WM Parcellation) of RightPallidum	Morfológica	0.57	0.6709
96	ST121TS	Cortical Thickness Standard Deviation of RightTransverseTemporal	Morfológica	0.5588	0.0029
118	ST127SV	Volume (WM Parcellation) of ThirdVentricle	Morfológica	0.6421	0.2154
172	ST109TS	Cortical Thickness Standard Deviation of RightPosteriorCingulate	Morfológica	0.5097	0.2516
173	ST13CV	Volume (Cortical Parcellation) of LeftBankssts	Morfológica	0.7055	0.1079
182	ST15TS	Cortical Thickness Standard Deviation of LeftCaudalMiddleFrontal	Morfológica	0.5006	0.0637
189	ST38CV	Volume (Cortical Parcellation) of LeftLingual	Morfológica	0.5764	0.0924
222	ST57SA	Surface Area of LeftSuperiorParietal	Morfológica	0.5356	0.0374
263	ST110TS	Cortical Thickness Standard Deviation of RightPrecentral	Morfológica	0.5506	0.0767
279	ST114CV	Volume (Cortical Parcellation) of RightRostralMiddleFrontal	Morfológica	0.6319	0.1246
314	ST10CV	Volume (Cortical Parcellation) of Icv	Morfológica	0.5298	0.4283
1	CDGLOBAL	Global CDR	Clínica	0.9777	0.1254
3	PTEDUCAT	Años de educacion	Clínica	0.5811	0.0205
4	PTGENDER	Genero	Clínica	0.5232	1.4259

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

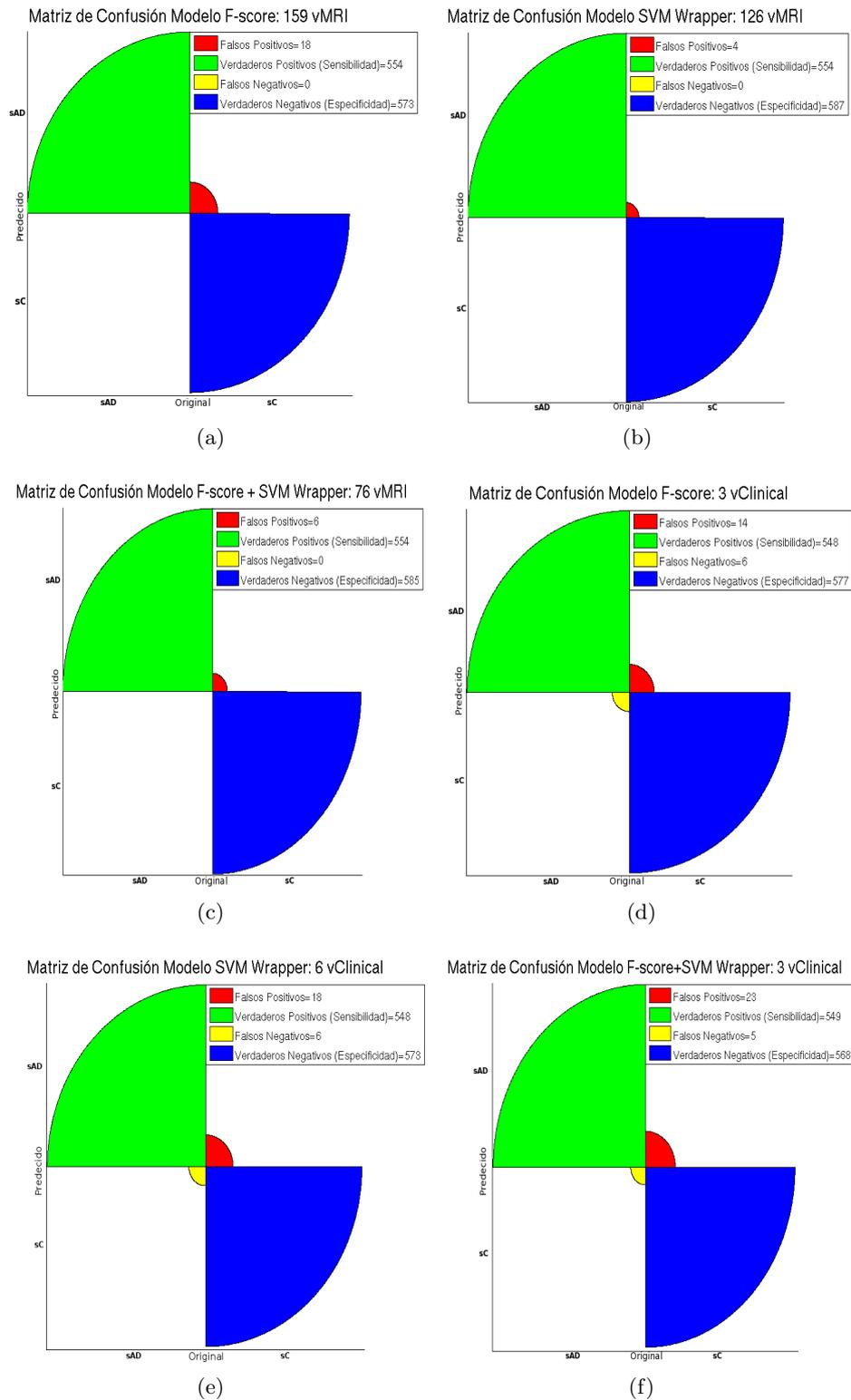


Figura 4.5: Matrices de confusión obtenidas con los métodos de selección de variables. Variables morfo-métricas: (a) Método F-score, (b) Método SVM Wrapper, (c) Método F-score + SVM Wrapper. Variables clínicas: (d) Método F-score, (e) Método SVM Wrapper, (f) Método F-score + SVM Wrapper

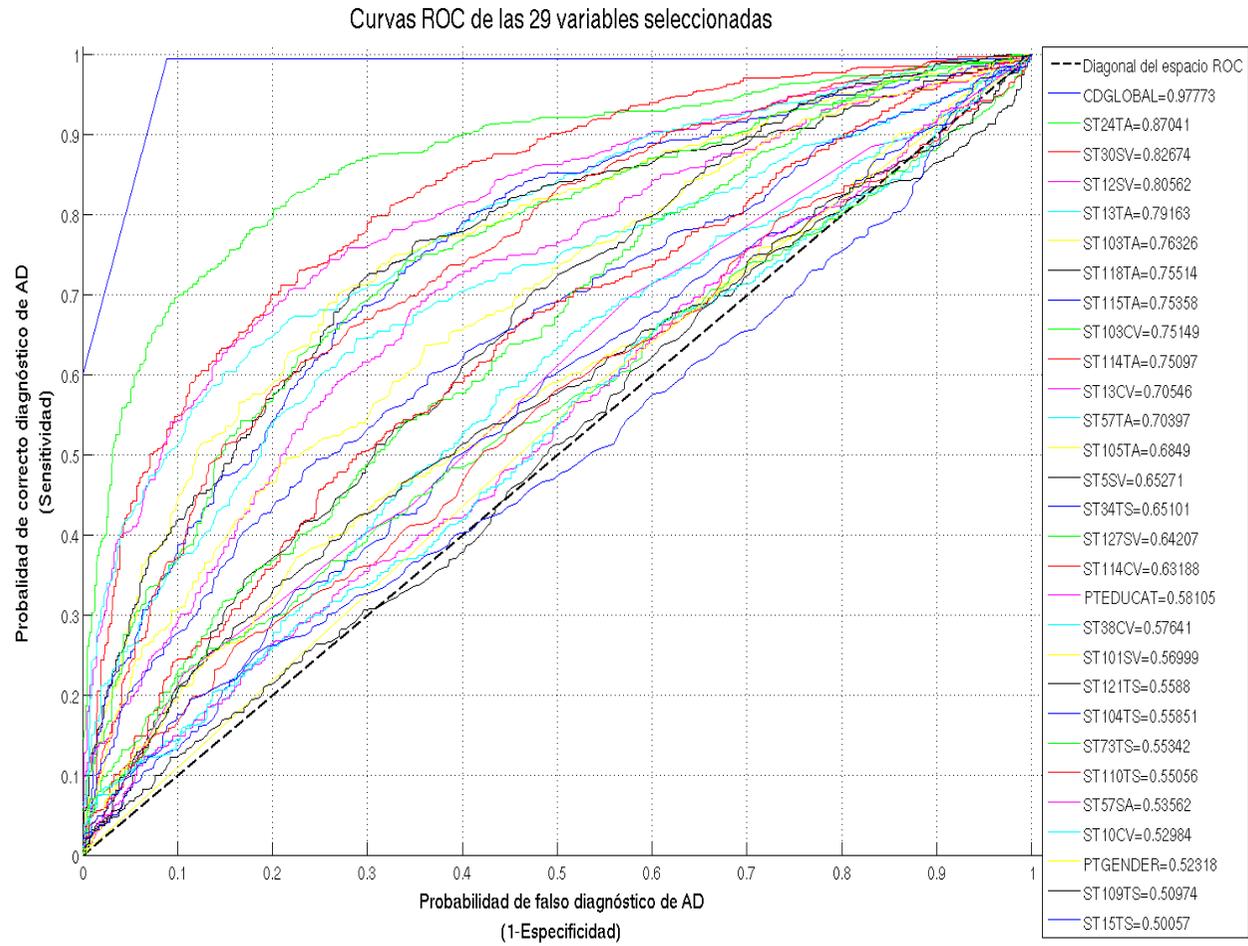


Figura 4.6: Curvas ROC de las 29 variables seleccionadas. Valores AUC calculados sobre el grupo $sCsAD_{ADNI-train}$

4. CONSTRUCCIÓN DE LOS MODELOS DE DIAGNÓSTICO Y PREDICCIÓN DE AD BASADOS EN SVM.

Capítulo 5

Entrenamiento, Test y Evaluación de los modelos

5.1 Introducción

En este capítulo describimos la etapa de entrenamiento y test para la construcción de modelos SVM de diagnóstico y predicción de AD a partir de los datos obtenidos en la etapa anterior. Asimismo, esta etapa implicó como tareas iniciales la determinación del comportamiento lineal o no lineal del problema de clasificación y el establecimiento de las características de configuración de los modelos SVM a entrenar (tipo de la función kernel y la elección de los mejores parámetros para SVM). La construcción de ambos modelos implicó seis etapas principales, como se presenta en la figura 5.1.

Dos modelos SVM: SVM-Diag₂₉ y SVM-Pred₂₉ fueron entrenados utilizando los grupos de entrenamiento de ADNI y las 29 variables predictoras seleccionadas. Durante la etapa de test, los modelos fueron evaluados con grupos de test obtenidas de la base de datos ADNI: sCsAD_{ADNI-test} y sMCI_{ADNI-test}; y de la base de datos *Open Access Series of Imaging Studies* (OASIS)(<http://www.oasis-brains.org/>) [Marcus et al., 2007]. Dos nuevos grupos fueron extraídos de la base de datos OASIS: un primer grupo de 45 casos clínicos seleccionados de la base de datos transversal (*cross-sectional database*) de OASIS y un segundo grupo de 7 casos de estudios obtenidos de la base de datos longitudinal. Ambos grupos de OASIS incluyeron casos de sujetos diagnosticados como sC o sAD, por lo que fueron al modelo de diagnóstico¹. La clasificación de estos nuevos grupos de test estuvo sujeta al re-entrenamiento del modelo de diagnóstico, debido a la exclusión de una de las variables predictoras del modelo original. Asimismo, en el segundo grupo de OASIS, las variables predictoras morfométricas fueron obtenidas a partir de las MRIs longitudinales de cada paciente, utilizando una herramienta desarrollada en MATLAB. Esta herramienta incluyó las librerías de procesamiento y análisis de imágenes de Freesurfer.

Los resultados obtenidos durante la etapa de test, confirman la eficiencia de la metodología introducida en el capítulo anterior y la robustez de los modelos SVM construidos en este trabajo.

¹El modelo de predicción no fue utilizado con la base de datos OASIS debido a que sus estudios transversales y longitudinales no tienen documentados casos de sujetos con MCI, los cuales se hayan convertido o no a AD

5. ENTRENAMIENTO, TEST Y EVALUACIÓN DE LOS MODELOS

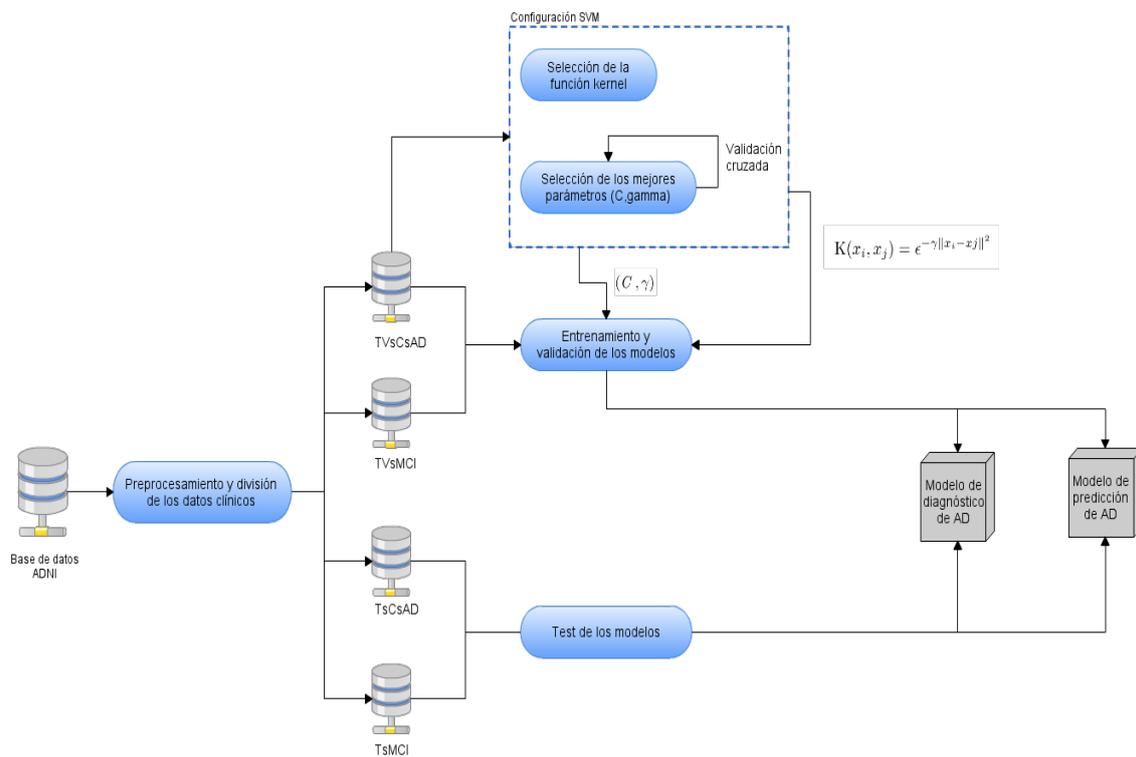


Figura 5.1: Etapas implicadas de la construcción de los modelos SVM para el diagnóstico y predicción de AD

5.2 Entrenamiento de los modelos SVM de diagnóstico y predicción de AD.

El principal objetivo en esta etapa fue el entrenamiento de dos modelos de clasificación basados en SVM: el primero para el diagnóstico de AD y el segundo para la predicción de la conversión de los pacientes MCI a AD. En nuestro estudio, los modelos clasificadores SVM fueron implementados utilizando la herramienta LIBSVM [Chang and Lin, 2011] con el kernel de configuración *Gaussian radial basics funtion* (RBF). Asimismo los datos utilizados correspondieron a los grupos creados en la primera etapa de la metodología, ver sección 4.3.4, considerando únicamente las variables seleccionadas, ver tabla 4.9 . Chih-Wei et al. [2003] presenta una fácil introducción a LIBSVM, detallando los pasos para preprocesar los datos, la selección del modelo utilizando el kernel RBF, así como sus beneficios y desventajas.

5.2.1 Selección del kernel de modelo SVM

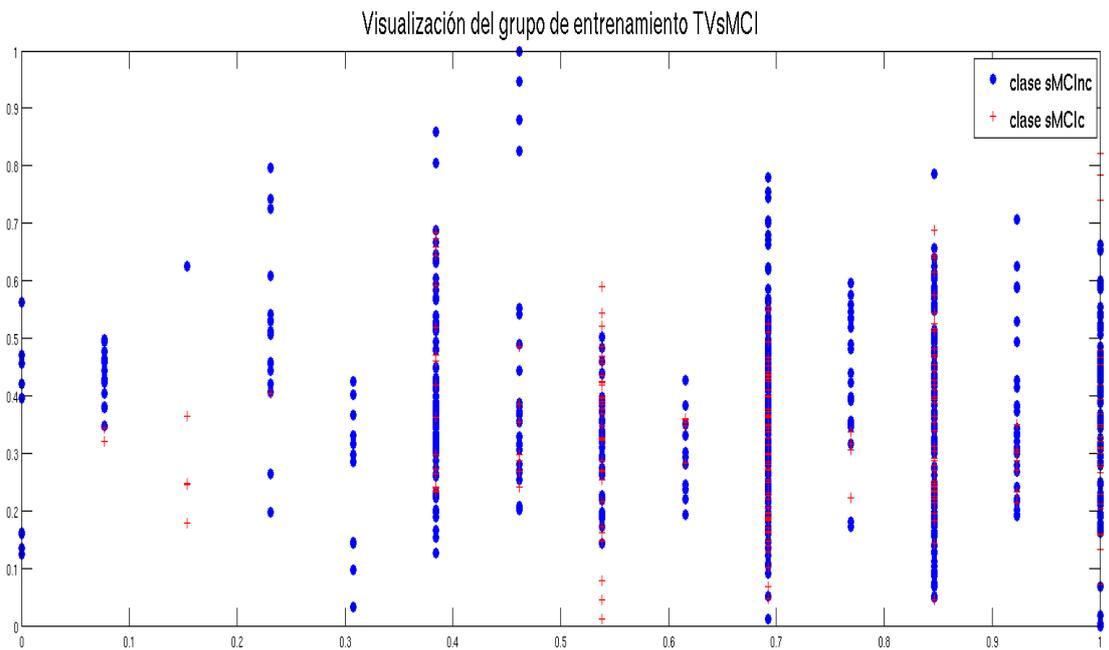
Los datos de los cuatro grupos fueron analizados. Para los grupos de diagnóstico $sCsAD_{ADNI-train}$ y $sCsAD_{ADNI-test}$, fueron consideradas las características del problema de clasificación (no linealidad de los datos y la superioridad del número de instancias con respecto al número de variables), descritas en la sección 4.4.3.1 para la etapa de selección de variables. De igual manera, los grupos de predicción $sMCI_{ADNI-train}$ y $sMCI_{ADNI-test}$ fueron analizados, encontrándose las mismas características: separación no lineal de las clases, ver tabla 5.2 y un número de instancias mucho mayor que el número de variables (894 instancias vs. 29 variables); siendo establecida la misma configuración que los grupos de diagnóstico. De este modo, el kernel RBF fue utilizado para la configuración de los modelos SVM. Adicionalmente existen otras razones que justifican nuestra selección de este kernel con respecto a los otros tipos existentes. Por un lado, RBF mapea no-linealmente las muestras sobre un espacio de dimensiones grandes (recomendado para un número de variables pequeño), lo cual, a diferencia del kernel de tipo lineal, permite manejar satisfactoramente los datos que poseen una relación no-lineal entre los *inputs* (variables de entrada) y los *targets*. Por otro lado, investigaciones han demostrado que el kernel RBF abarca (y en algunos casos supera) las funcionalidades de los otros kernels, por lo que es recomendable su selección como primera opción. Al respecto, Keerthi and Lin [2003] ha demostrado que RBF es un caso especial del kernel lineal con un parámetro adicional, Lin and Lin [2003] demostró que RBF alcanza el mismo rendimiento que el kernel sigmoid cuando el rango de ciertos parámetros son modificados, por otro lado, RBF presenta menos dificultades que el kernel polinomial al tratar el grado de las dimensiones de los valores de las variables. En Chih-Wei et al. [2003] se presenta una guía completa de los criterios a considerar para la selección de la función kernel, los cuales fueron aplicados en este estudio.

5.2.2 Búsqueda de los mejores parámetros de configuración de los modelos SVM.

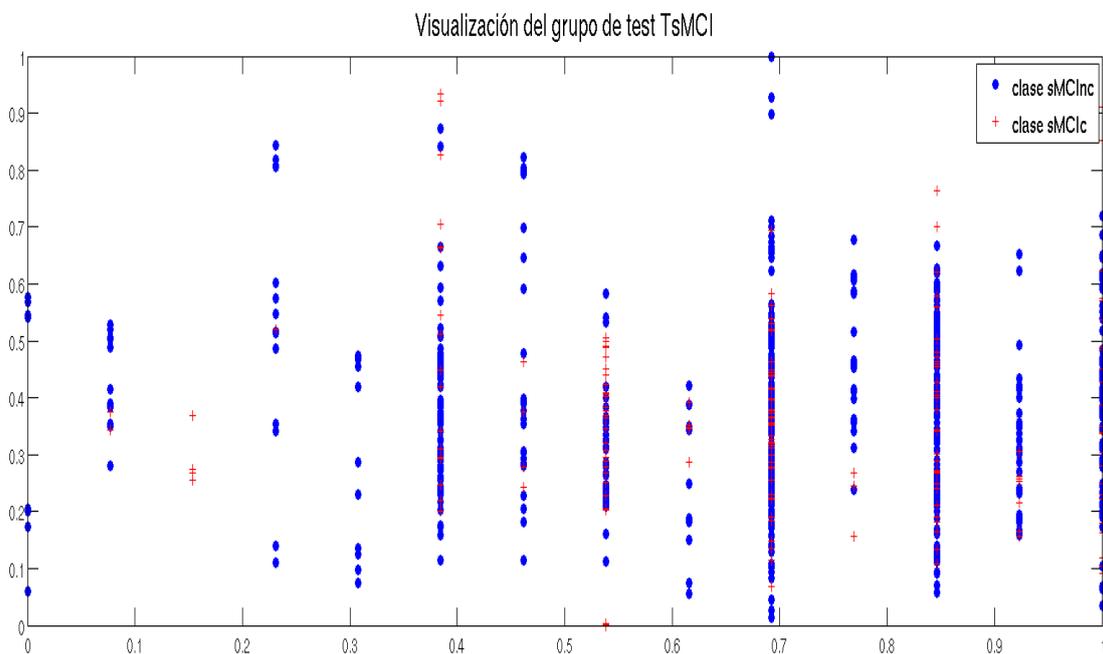
En la sección 3.3.2.1 se mencionó que los parámetros necesarios para un clasificador SVM basado en el kernel RBF son C y γ , ver ecuación 3.4. C es un parámetro de costo del clasificador SVM, el cual permite mayor flexibilidad en la separación de las clases, por lo que es importante asignarle un margen adecuado; y γ (*Gaussian kernel width*) es un parámetro propio del kernel RBF. Debido a que inicialmente, los valores para ambos parámetros no son conocidos, las asignación de los valores más óptimos es realizada con una red de búsqueda (*grid-research*).

En este trabajo, cada combinación de estos parámetros fue comprobada utilizando el algoritmo de validación cruzada (*cross-validation*). Este algoritmo fue implementado en MATLAB [MathWorks, 2010] y se utilizó la librería LIBSVM para calcular la exactitud de clasificación de los modelos basados en SVM. El algoritmo fue iniciado asignado la mejor exactitud de validación (*bestc*) igual a cero y con el establecimiento de los límites o posibles intervalos de (C y γ) en el *grid space*. Entonces, la combinación de todos los posibles *grid points* de

5. ENTRENAMIENTO, TEST Y EVALUACIÓN DE LOS MODELOS



(a)



(b)

Figura 5.2: Distribución no lineal de los grupos de entrenamiento y test para el modelo de predicción: (a) grupo $sMCI_{ADNI-train}$. (b) Grupo $sMCI_{ADNI-test}$

(C, γ) fueron probados a fin de encontrar los primeros mejores valores ($bestc$ y $bestg$), con los cuales se mantuviera o incrementara el mejor valor para $bestcv$ en la clasificación de datos desconocidos (validación), es decir datos no utilizados durante el entrenamiento del modelo. La búsqueda de los mejores parámetros continuó hasta recorrer por completo los límites definidos para $(C$ y $\gamma)$. El algoritmo de búsqueda utilizado es presentado a continuación:

Código de MATLAB del algoritmo de validación cruzada utilizada para la búsqueda de los mejores parámetros (C y γ)

```
bestcv = 0; %inicia mejor valor de exactitud como cero
for log2c = -5:5,$ % define grid space para par\`ametro C
    for log2g = -10:10,%define grid space para par\`ametro gamma
        %%% Comando para una clasificaci\`on no lineal
        cmd = ['-t 2 -v 7 -c ', num2str(2^log2c), ' -g ', num2str(2^log2g)];
        % donde: t= 2 define tipo de kernel igual a Radial Basis Function: exp(-gamma*|u-v|^2)
        %          v=2 define 2-fold cross validation para el clasificador SVM
        %%% Entrenamiento del modelo con los par\`ametros actuales
        cv = svmtrain(data_scale_label, data_scale_inst, cmd);
        % el modelo SVM es entrenado y se devuelve su exactitud de clasificaci\`on
        % en variable cv
        %%% Verifica la nueva exactitud cv alcanzada con los valores actuales de C y gamma
        if (cv >= bestcv), % Si cv es mejor que actual valor de bestcv,
            % entonces los par\`ametros bestc, bestg y bestcv son reemplazados
            bestcv = cv; bestc = 2^log2c; bestg = 2^log2g;
        end
    end
end
end
```

Como resultado, para el modelo de diagnóstico se encontraron como mejores parámetros: ($C = 4.0$ y $\gamma = 2.0$, con los cuales el modelo obtuvo como mejor $bestcv = 99.2893\%$. Asimismo, para el modelo de predicción, los mejores valores fueron ($C = 32.0$ y $\gamma = 4.0$, y el mejor valor de $bestcv$ fue de 88.7025%).

5.2.3 Entrenamiento y validación

Finalmente, basándonos en la configuración establecida en los pasos anteriores, el modelo SVM destinado al diagnóstico de sujetos sC y sAD, el cual denominamos SVM-Diag₂₉, fue entrenado y validado con las 1265 instancias y 29 variables predictoras del grupo sCsAD_{ADNI-train}. Asimismo, el modelo SVM de predicción destinado a la clasificación de los casos como sMCInc y sMCIC, el cual denominamos SVM-Pred₂₉, fue entrenado con las 894 instancias y 29 variables predictoras del grupo sMCI_{ADNI-train}. La construcción de ambos modelos fue realizada mediante la herramienta prototipo desarrollada en MATLAB [MathWorks, 2010], la cual incluye librerías de LIBSVM [Chang and Lin, 2011] para el entrenamiento, validación y test de modelos SVM.

En el siguiente capítulo son presentados los resultados de la etapa de test de ambos modelos, con los casos clínicos de ADNI y nuevos casos obtenidos de la base de datos *Open Access Series of Imaging Studies* (OASIS) [Marcus et al., 2007].

5.3 Test de los modelos SVM.

5.3.1 Casos de Estudio ADNI: Diagnóstico y predicción de AD

El modelo de diagnóstico SVM-Diag₂₉ fue utilizado para clasificar las 1145 instancias del grupo $sCsAD_{ADNI-test}$, obteniéndose una matriz de confusión conformada por: TP=554, es decir, 554 instancias originalmente clasificadas como sAD fueron clasificadas correctamente como sAD; TN=582, es decir 582 instancias originalmente diagnosticadas como sC fueron clasificadas correctamente por el modelo como sC; FP=9; es decir 9 instancias diagnosticada como sC fueron incorrectamente clasificadas por el modelo como sAD; y FN=0, es decir es decir ninguna instancia diagnosticada como sAD fue incorrectamente clasificada como sC. Asimismo, el modelo de predicción fue utilizado para clasificar las 814 instancias del grupo $sMCi_{ADNI-test}$, obteniéndose como matriz de confusión: TP=171, es decir, 171 instancias originalmente clasificadas como sMCiC fueron clasificadas correctamente por el modelo como sMCiC; TN=598, es decir 598 instancias originalmente clasificadas como sMCInc fueron clasificadas correctamente como sMCInc; FP=35; es decir 35 instancias diagnosticadas como sMCInc fueron incorrectamente clasificadas como sMCiC; y FN=10, es decir es decir 10 instancias diagnosticadas como sMCiC fue incorrectamente clasificadas como sMCInc.

Las matrices de confusión y y sus respectivas curvas ROC obtenidas para ambos modelos, son presentadas en las figuras 5.3 y 5.4.

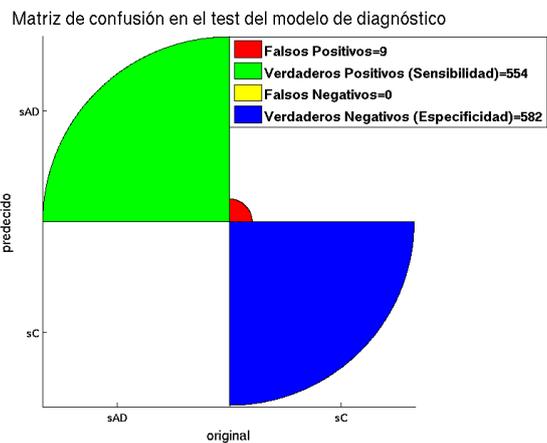
5.3.2 Casos de estudio OASIS: Diagnóstico de AD

OASIS es una base de datos libre dedicada a la investigación en demencia, principalmente del tipo Alzheimer, la cual proporciona información muy similar a la proporcionada por la base de datos ADNI. Por cada participante, OASIS proporciona información demográfica y clínica, así como las diferentes sesiones de Imágenes de Resonancia Magnética (MRIs) obtenidas en cada una de sus visitas. Estas imágenes han sido procesadas con el software *Fresurfer* [Dale et al., 1999; Fischl et al., 2001; Fischl and Dale, 2000; Fischl et al., 1999a,b, 2002, 2004a,b; Han et al., 2006; Jovicich et al., 2006; Segonne et al., 2004] para reconstruir las zonas corticales y el volumen de ciertas regiones de interés (ROIs) del cerebro, a fin de proporcionar un conjunto de biomarcadores morfométricos.

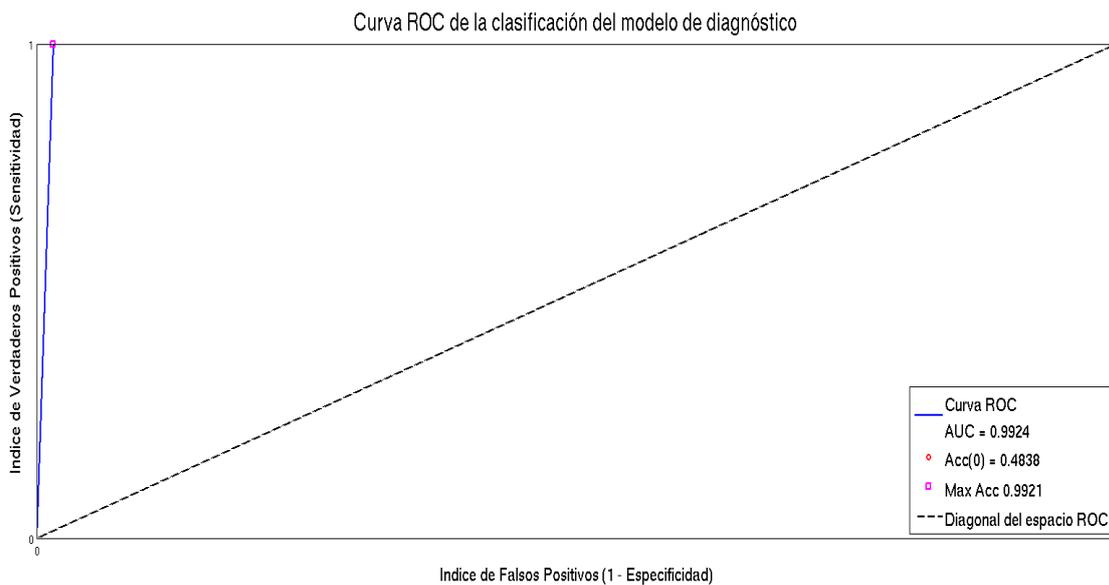
5.3.2.1 Obtención de los casos clínicos

La información sociodemográfica, neuropsicológica y morfométrica de OASIS corresponde a participantes clasificados como “Dementes” o “no-dementes”. Asimismo, dicha información está clasificada como estudio transversal y estudio longitudinal, determinado por la orientación de las MRIs obtenidas para cada participante. El estudio transversal incluyó inicialmente información de 416 participantes adultos, con edades entre 18-96. Por cada sujeto, 3 o 4 MRIs individuales del tipo *T1-weighted* fueron obtenidas por sesión, de acuerdo a cada visita del paciente. 100 de los sujetos incluidos, con edad sobre los 60 años, fueron diagnosticados con muy leve a moderado AD. Por otro lado, el estudio longitudinal incluyó información de 150 participantes con edades de 60-96 años, incluyendo 64 individuos con muy leve a moderado estado de AD y 14 individuos originalmente diagnosticados como no dementes, lo cuales en alguna posterior visita fueron diagnosticados con AD. En ambos estudios, el diagnóstico clínico de los participantes estuvo basado en la escala *Clinical Dementia Rating* (CDR) [Morris, 1993; Morris et al., 2001], clasificados como “No-dementes” a los participantes con CDR igual a 0 y como “dementes” a los participantes con CDR mayor o igual a 0.5. Los estudio completos puede ser descargado libremente desde el sitio Web <http://www.oasis-brains.org/>.

Un primer grupo de 45 casos fue seleccionado del estudio tranversal de OASIS, el cual denominamos $sCsAD_{OASIS-45}$. Todos los casos seleccionados satisfacían la condición de tener información neuropsicológica, demográfica y morfométrica completa, libremente disponible. $sCsAD_{OASIS-45}$ estuvo conformado por 26 casos clasificados como No-dementes y 19 casos clasificados como dementes (sujetos con AD leve), los cuales fueron re-etiquetados en este



(a)



(b)

Figura 5.3: Resultados del modelo de diagnóstico en la clasificación del grupo $sCsAD_{ADNI-test}$: (a) Matriz de confusión. (b) Curva ROC

5. ENTRENAMIENTO, TEST Y EVALUACIÓN DE LOS MODELOS

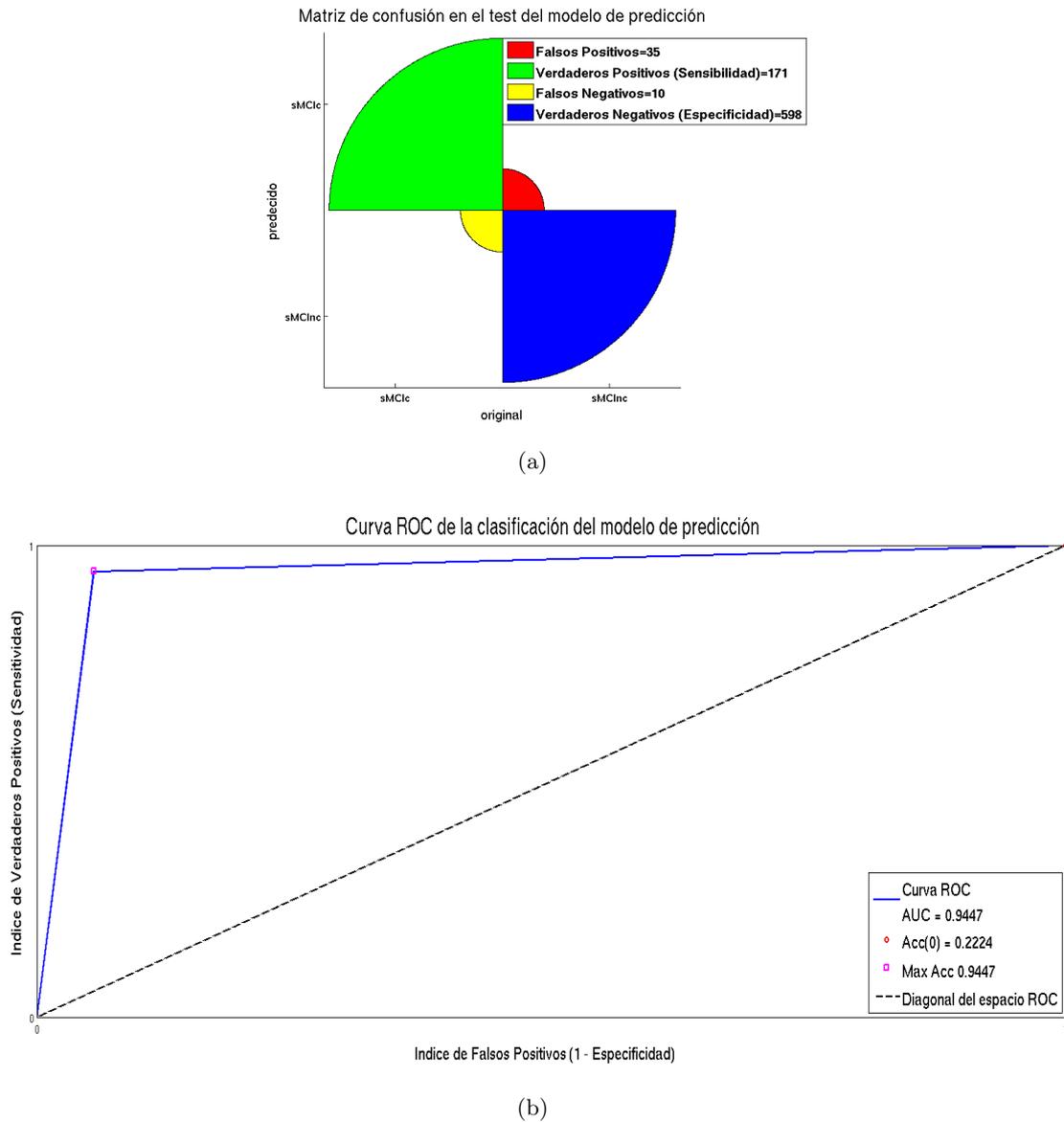


Figura 5.4: Resultados del modelo de predicción en la clasificación del grupo $sMCI_{ADNI-test}$: (a) Matriz de confusión. (b) Curva ROC

estudio como sC (sujetos), Un segundo grupo para el diagnóstico estuvo conformado por 7 casos clínicos precedentes del estudio longitudinal de OASIS, el cual denominamos sCsAD_{OASIS-7}. Para este grupo sólo fue considerada la información demográfica y neuropsicológica a fin de completar las variables clínicas del modelo. Con la finalidad de completar la información de las variables morfométricas implicadas en el modelo de diagnóstico, desarrollamos una herramienta prototipo en MATLAB que procesara y analizara las MRIs longitudinales de los participantes, obtenidas en cada una de sus visitas. Esta herramienta incluyó las principales rutinas de procesamiento de MRIs mencionadas en la sección 4.3.5 y descritas con mayor detalle en el capítulo 2.

5.3.2.2 Análisis de las variables de OASIS

Para la clasificación de ambos grupos de OASIS, sólo debían ser consideradas las 29 variables de entrada, con las cuales el modelo de diagnóstico fue entrenado¹. Sin embargo, durante este proceso, se determinó que la variable PTEDUCATION (llamada "Educación" en OASIS) es asumida de diferente manera a ADNI. En OASIS, esta variable corresponde a los niveles de educación del participante, la cual puede asumir los siguientes valores: 1= *less than high school grad.*, 2: *high school grad.*, 3: *some college*, 4: *college grad.* y 5: *beyond college*. Por otro lado, en los estudios de ADNI, esta variable hace referencia a los años de educación del participante asignados a cinco grados de estudios base. En la tabla 5.1 se presenta una comparación de los valores asumidos por este variable en ambos estudios.

Tabla 5.1: Comparación de la variable "Educación" en los estudios ADNI y OASIS

ADNI		OASIS	
Grado	Valor (Años)	Grado	Valor
GED	12	<i>Less than high school grad.</i>	1
<i>High School Diploma</i>	12	<i>High school grad.</i>	2
AA	14	<i>Some college</i>	3
BA/BS	16	<i>College grad.</i>	4
MA/MS	18	<i>Beyond college</i>	5
<i>Law Degree</i>	19	–	–
MD, PhD, or both	20	–	–

5.3.2.3 Re-entrenamiento del modelo de diagnóstico

Debido a la diferencia encontrada con la variable PTEDUCATION, esta variable fue excluida de los datos de entrada y el modelo de diagnóstico fue re-entrenado y validado considerando sólo 28 variables. Este modelo fue denominado SVM-Diag₂₈. La construcción de SVM-Diag₂₉ siguió el mismo procedimiento descrito en la sección 5.2: la rutina de búsqueda de los parámetros de configuración SVM obtuvo como mejores valores ($C = 8.0$ y $\gamma = 0.5$) y RBF fue seleccionado como kernel. Asimismo, durante el test con las 1145 instancias del grupo sCsAD_{ADNI-test}, SVM-Diag₂₈ obtuvo como matriz de confusión: TP=552, TN=579, FP=12 y FN=2.

5.3.2.4 Clasificación

En la clasificación del grupo sCsAD_{OASIS-45}, el modelo de diagnóstico SVM-Diag₂₈ (modelo SVM de diagnóstico re-entrenado con 28 variables de entrada), obtuvo como matriz de confusión: TP=18; TN=24; FP=2 y FN=1.

¹Los estudios transversal y longitudinal de OASIS incluye inicialmente 197 variables, de las cuales: 190 variables morfométricas correspondientes a mediciones de la zona cortical y volúmenes segmentados del cerebro, obtenidas con Freesurfer; 5 variables demográficas; y 2 variables neuropsicológicas

5. ENTRENAMIENTO, TEST Y EVALUACIÓN DE LOS MODELOS

Asimismo, en la clasificación de los 7 casos longitudinales del grupo $sCsAD_{OASIS-7}$, el modelo SVM-Diag₂₈ obtuvo una matriz de confusión conformada por: TP=3; TN=2; FP=2 y FN=0.

Las matrices de confusión y sus respectivas curvas ROC obtenidas por el modelo de diagnóstico SVM-Diag₂₉ para ambos grupos OASIS son presentadas en las figura 5.5 y 5.6.

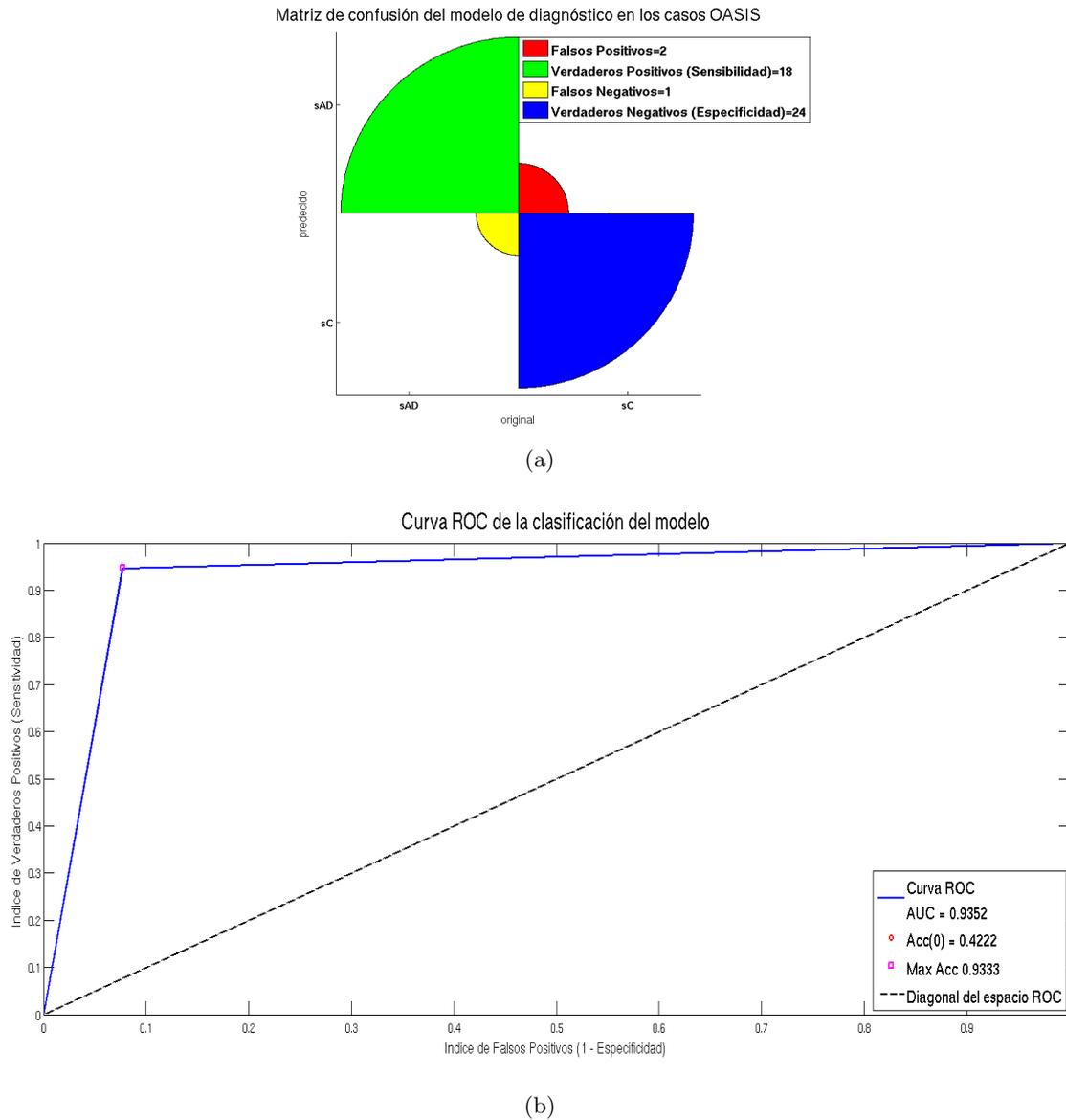


Figura 5.5: Resultados del modelo de diagnóstico en la clasificación del grupo $sCsAD_{OASIS-45}$: (a) matriz de confusión. (b) curva ROC.

5.3.3 Desempeño de los modelos de diagnóstico y predicción de AD

Con la finalidad de analizar el desempeño de los modelos construidos a partir de las variables predictoras seleccionadas, los cinco indicadores estadísticos descritos en la sección 3.3.3: Exactitud(ACC), sensibilidad (TPR),

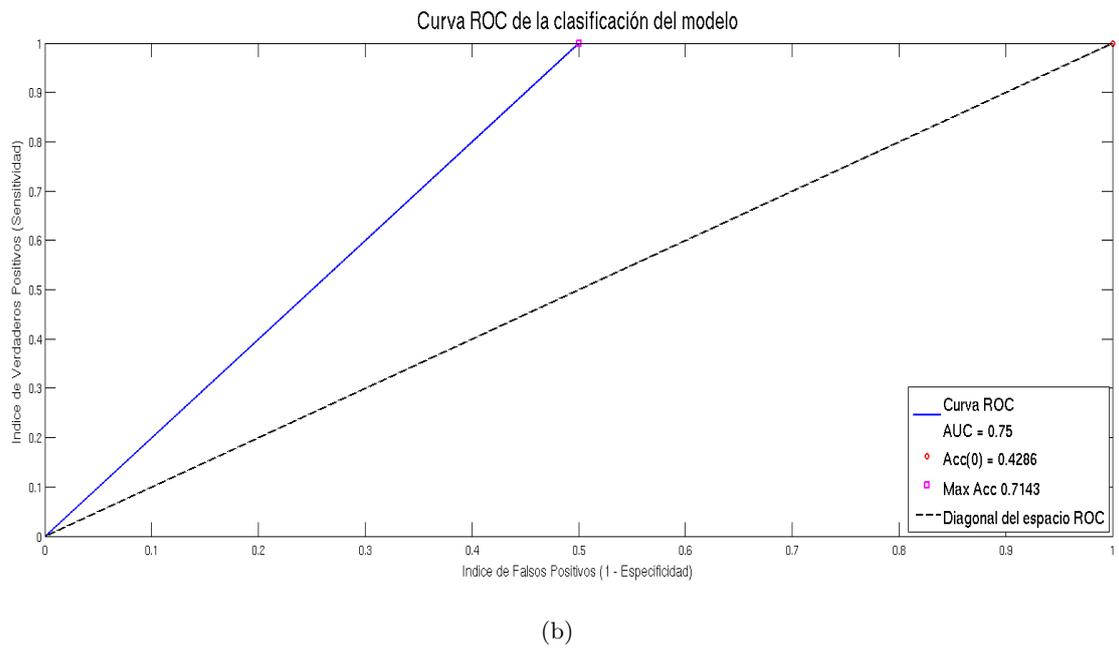
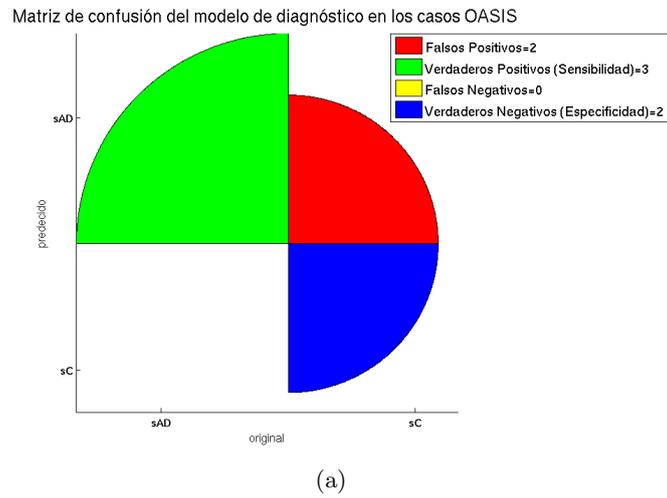


Figura 5.6: Resultados del modelo de diagnóstico en la clasificación del grupo $sCsAD_{OASIS-7}$: (a) matriz de confusión. (b) curva ROC.

5. ENTRENAMIENTO, TEST Y EVALUACIÓN DE LOS MODELOS

especificidad (TNR) y precisión (PRE) calculados sobre las respectivas matrices de confusión de los modelos y los valores AUC; son presentados en la tabla 5.2. En el capítulo 6 son analizados estos resultados.

Tabla 5.2: Desempeño de los modelos SVM en el diagnóstico y predicción de AD.

Grupos	Modelo	Tarea	Variables	ACC	PRE	TPR	TNR	AUC
ADNI	SVM-Diag ₂₉	Diagnóstico	29	99.21	98.40	100	98.48	0.9924
	SVM-Pred ₂₉	Predicción	29	94.47	83.01	94.48	94.47	0.9447
	SVM-Diag ₂₈	Diagnóstico	28	98.78	97.87	99.64	99.64	0.9797
OASIS	SVM-Diag ₂₈	Diagnóstico	28	93.33	90	94.74	92.31	0.9352
	SVM-Diag ₂₈	Diagnóstico	28	71.43	60.00	100.00	50.00	0.7500

Capítulo 6

Análisis de Resultados, Conclusiones y Trabajo Futuro

6.1 Introducción

En los capítulos 4 y 5 se detallaron paso a paso las tres etapas de la metodología propuesta para dar soporte al diagnóstico de AD mediante la construcción de modelos de clasificación basados en SVM, el primer modelo dedicado al diagnóstico de AD y el segundo modelo dedicado a la predicción de la conversión de sujetos con MCI a sujetos con AD. Haciendo un repaso de los principales procesos inmersos en esta metodología, varias técnicas numéricas y estadísticas fueron aplicadas. Partiendo desde la obtención y preparación de los datos de ensayos clínicos obtenidos de la base de datos ADNI, cuatro grupos de casos clínicos con 324 variables fueron construidos inicialmente. Con la finalidad de descartar las variables redundantes y poco influyentes en la clasificación, fueron aplicados tres enfoques de selección de las mejores variables predictoras para el problema de clasificación. Entonces, sólo una vez definidas las variables más útiles, se utilizaron las capacidades del aprendizaje basado en SVM para construir los dos modelos de interés: SVM-Diag₂₉ (modelo de diagnóstico de AD) y SVM-Pred₂₉ (modelo de predicción de AD).

En el capítulo 5 se presentaron los resultados obtenidos por los modelos SVM en la clasificación de dos grupos de casos clínicos provenientes de ADNI. Asimismo, se presentaron los resultados de clasificación del modelo de diagnóstico SVM-Diag₂₈ re-entrenado con 28 de las 29 variables predictoras seleccionadas, el cual fue utilizado para clasificar dos nuevos grupos de clínicos extraídos de las bases de datos transversal y longitudinal OASIS. Adicionalmente, para completar la información de las variables morfométricas de unos de los grupos de OASIS, fue necesario procesar las MRIs de los participantes utilizando una herramienta implementada en MATLAB.

Los resultados presentados en el capítulo 5 son prometedores. Por un lado, el subconjunto de variables seleccionadas tiene correlación con los resultados de investigaciones previas [Cui et al., 2011; Desikan et al., 2009; Heckemann et al., 2011], lo cual es un buen indicio de la confiabilidad de nuestros resultados. Asimismo, la clasificación de los casos clínicos de ADNI y OASIS, realizada por los modelos SVM: SVM-Diag₂₉, SVM-Pred₂₉ y SVM-Diag₂₈; han demostrado la eficiencia de la metodología aplicada, la efectividad de las variables predictoras seleccionadas, así como los procesos establecidos en la construcción de modelos SVM construidos a partir de estas variables. Estos modelos son capaces de discriminar sujetos con AD de sujetos sanos (diagnóstico) y distinguir a los sujetos MCI que se convertirán a AD de los que no se convertirán (predicción) con valores de especificidad y sensibilidad muy aceptables. Por otro lado, los resultados nos permiten hacer otras afirmaciones que confirman investigaciones previas a este trabajo y plantear un abanico de posibilidades para seguir investigando en esta área.

6. ANÁLISIS DE RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO

En este capítulo presentamos un análisis de estos resultados, centrándonos en el desempeño de los modelos a partir de las variables predictoras seleccionadas, la posible influencia de las características sociodemográficas o neuropsicológicas en la clasificación de los casos clínicos y de los costos computacionales implicados en este estudio. Asimismo, son presentados las conclusiones finales y trabajos futuros que han surgido a partir del abanico de posibilidades existentes en esta área.

6.2 Análisis de Resultados

Los resultados en la clasificación obtenidos por el modelo de diagnóstico de AD, ya sea considerando las 29 variables seleccionadas o solamente 28 de ellas, han demostrado la eficiencia de la metodología aplicada en este trabajo, la efectividad de las variables predictoras seleccionadas, así como de los procesos implicados en la construcción de modelos de clasificación SVM. Estos resultados nos permiten hacer afirmaciones que confirman investigaciones previas a este trabajo y abren un abanico de posibilidades para seguir investigando en esta área.

6.2.1 Diagnóstico y predicción de AD. Problema de clasificación no-lineal

Las matrices de confusión obtenidas por los modelos SVM-Diag₂₉ y SVM-Pred₂₉, específicamente en los valores de FP y FN, presentadas en las figuras 5.3(a) y 5.4(a), demuestran que a pesar de haber utilizado para su configuración el kernel RBF y seleccionado los mejores valores para los parámetros (C, γ) , la separación de las clases en el hiperplano SVM no ha sido 100% correcta. Esto se debe a que en ambos casos, el problema es no-lineal para todos los espacios de entrada, por lo que bajo ninguna configuración se conseguirá clasificar correctamente todas las instancias (obtener valores FP=0 y FN=0). Asimismo, comparando ambas matrices, observamos que en el modelo de predicción presentó un error de clasificación mucho más alto (FP y FN mayores) que el modelo de diagnóstico. En este caso, la separación de las clases positivas (sMCIC) de las clases negativas (sMCInc) en el hiperplano fue mucho más complicado por tener un comportamiento no-lineal mayor.

6.2.2 Desempeño de los modelos de diagnóstico y predicción de AD

En la tabla 5.2 se presentaron los indicadores estadísticos: Exactitud, precisión, sensibilidad, especificidad y AUC; calculados sobre las respectivas matrices de confusión de los modelos entrenados, con la finalidad de evaluar su desempeño en la discriminación de las instancias sC de las sAD; y entre sMCInc y sMCInc. Estos indicadores resultaron muy útiles por combinar la sensibilidad y especificidad de los modelos. A continuación se analizan los resultados más interesantes.

Clasificación de los casos ADNI: Diagnóstico de AD (sC vs. sAD). En la clasificación de los datos ADNI, el modelo de diagnóstico SVM-Diag₂₉, construido con las 29 variables, obtuvo las calificaciones más altas (sensibilidad = 100% y especificidad = 98.48%) en comparación al modelo SVM-Diag₂₈ (sensibilidad = 94.74% y especificidad = 92.31%), entrenado sin la variable PTEDUCATION. Sin embargo la diferencia entre ambos resultados fue muy pequeña. Asimismo, el desempeño del modelo SVM-Diag₂₉ superó a la mayoría de los indicadores calculados para los seis modelos auxiliares (ver tabla 4.8), siendo superado mínimamente por el modelo auxiliar SVM_{vMRI-Wrapper}, construido con 159 variables morfométricas. Comparando los subconjuntos de variables de ambos modelos, las 29 variables del modelo SVM-Diag₂₉, además de ser un número más reducido de variables y posiblemente garantizar una menor redundancia, también presentan valores de AUC y F-score más aceptables que el segundo subconjunto. Asimismo, estos resultados superaron el desempeño obtenido por otros modelos de diagnóstico de AD entrenados sobre casos clínicos obtenidos de ADNI: (1) Los modelos de diagnóstico sC vs. sAD construidos por [Cuingnet et al., 2011] (clasificación basada en vóxel y espesor cortical) obtuvieron una sensibilidad de 81% y especificidad de 95%, asimismo, en el mejor de los casos, sus modelos de predicción

sMCIc vs. sMCIc obtuvieron una sensibilidad de 62% y especificidad de 69%. En la tabla 3.3 se presentaron los resultados obtenidos por otras investigaciones, los cuales han sido superados por los resultados obtenidos en este trabajo.

Clasificación de los casos ADNI: Predicción de AD (sMCIc vs. sMCIc). En la clasificación de los datos ADNI, el modelo de predicción SVM-Pred₂₉, construido con las 29 variables obtuvo un gran desempeño: sensibilidad = 94.48% y especificidad = 94.47%, resultados que han superado el desempeño de otros modelos de predicción entrenados sobre datos de ADNI: (1) En [Cuingnet et al., 2011] (clasificación basada en vóxel y espesor cortical), el mejor de los modelos entrenados para predicción obtuvo una sensibilidad de 62% y especificidad de 69%; (2) Cui et al. [2011] construyó un modelo de predicción de AD combinando variables predictoras obtenidas de MRIs, mediciones de CSF y características neuropsicológicas, el cual obtuvo sensibilidad = 96.43%, especificidad = 48.28%. Para una comparación más completa de los resultados de este trabajo con otros modelos de predicción ver tabla 3.3.

Clasificación de los casos OASIS. La utilización del modelo de diagnóstico en la clasificación de los 2 nuevos grupos clínicos de OASIS: sCsAD_{OASIS-45} y sCsAD_{OASIS-7}; implicó la aplicación de tareas adicionales a fin de ajustar correctamente sus variables predictoras a la configuración de entrada del modelo. En primer lugar, la variable predictora PTEDUCATION (Años de Educación) fue excluida y se procedió a re-entrenar el modelo de diagnóstico, siendo creado el modelo SVM-Diag₂₈. Asimismo, a fin de completar la información morfométrica faltante en el grupo sCsAD_{OASIS-7}, las 26 variables morfométricas fueron calculadas sobre las MRIs cerebrales de los participantes, utilizando para ello, una herramienta de procesamiento y análisis de imágenes desarrollada en este estudio.

A pesar que la reducción de 29 a 28 variables de entrada podría implicar la reducción del desempeño del modelo, los indicadores continuaron siendo muy óptimos. El modelo SVM-Diag₂₈ alcanzó una especificidad y sensibilidad de 94.74% y 92.31%, respectivamente, en la clasificación del grupo sCsAD_{OASIS-45}; y especificidad de 100% y sensibilidad de 50% en la clasificación del grupo sCsAD_{OASIS-7}. Una sensibilidad de 50%, resultado muy poco aceptable en el diagnóstico clínico, representa que el modelo ha identificado correctamente sólo la mitad de los casos sAD sobre todos los casos realmente diagnosticados como sAD. sin embargo, a pesar de que este resultado puede ser considerado como una posible deficiencia del modelo, este resultado no puede ser asumido estrictamente debido al número escaso de casos clínicos evaluados (7 casos), por lo que la evaluación del modelo con nuevos casos se hace necesario. Por otro lado, el modelo obtuvo buenos resultados con los otros casos, lo cual, además de demostrar el aceptable desempeño del modelo SVM y la eficiencia y robustez de las variables predictoras utilizadas en su entrenamiento, también demuestran la viabilidad para calcular las variables morfométricas a partir de las MRIs de los pacientes, sin la necesidad de utilizar un software sofisticado.

Valores AUC. Analizando la tabla 4.9, observamos que salvo dos variables con AUC menor de 0.6: PTEDUCAT(Años de educación) y ST113TA (Cortical Thickness Average of RightRostralAnteriorCingulate), la mayoría de estas variables tienen un AUC mayor a 0.6, lo cual las clasifica de "suficientes" a "excelentes" en la distinción de los dos grupos clínicos sC y sAD. Los valores AUC de cada modelo, presentados en la novena columna de la tabla 5.2 y sus respectivas curvas ROC (ver figuras 5.3(b), 5.4(b), 5.5(b) y 5.6(b)) también reflejan la eficiencia de los modelos. Por ejemplo, el AUC del modelo de diagnóstico SVM-Diag₂₉ fue de 0.9924, valor muy cercano al AUC= 1 de una curva ideal, ver figura 5.3(b). En este caso, la curva ROC del modelo SVM-Diag₂₉ se interpreta como que un individuo seleccionado al azar entre el grupo sAD tiene un valor de prueba mayor que para un individuo elegido al azar entre el grupo sC en 99.24% de las veces (Zweig y Campbell, 1993). Si las 29 variables seleccionadas no pudiesen distinguir entre los dos grupos, es decir, cuando no hay diferencia entre las dos distribuciones, el AUC hubiese sido igual a 0,5 (la curva ROC coincidiría con la línea diagonal). En el caso que existiese una separación perfecta de los valores de los dos grupos, es decir, no hay solapamiento de las distribuciones, el AUC hubiese sido

6. ANÁLISIS DE RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO

igual a 1 (la curva ROC se llega a la esquina superior izquierda del gráfico). Diversos investigadores consideran que el área bajo el receptor de funcionamiento característico (ROC) curvas, conocido como AUC debería ser el criterio principal para evaluar la efectividad de los biomarcadores de diagnóstico, es así, que la efectividad de uso de esta medición estadística ha sido demostrada en diversas publicaciones, varias de ellas enfocados en cardiología y neurología [Cook, 2007; Cui et al., 2011; Doniger et al., 2005; Faraggi and Reiser, 2002; Pencina et al., 2008].

6.2.3 Influencia de las características sociodemográficas y neuropsicológicas en la enfermedad

Con la finalidad de analizar la influencia de las características sociodemográficas y neuropsicológicas utilizadas en este estudio sobre AD y la conversión de MCI a AD, se procedió a calcular los promedios de las variables PTEDUCATION (Años de educación), PTAGE (Edad del paciente) y MMSCORE (Mini-Mental Examination Score), así como el número de hombres y mujeres (PTGENDER). Los cálculos se realizaron para cada uno de los casos TP, TN, FP y FN de las matrices de confusión obtenidas durante la etapa de test con los grupos de ADNI y OASIS, ver tabla 6.1. Asimismo, estos resultados han sido generalizados para los totales de los casos positivos (TP + FN) y negativos (TN + FP) por cada grupo, ver tabla 6.2. Obsérvese que a efectos de uniformizar este análisis, sólo fueron considerados los casos de estudio de OASIS correspondientes a sujetos con edades mayor o igual a los 56 años. De este modo, sólo fueron analizados 35 de los 45 casos del grupo sCsAD_{OASIS-45}. A pesar que el número de casos positivos y negativos es pequeño para establecerse como pruebas de significación estadística de la influencia de las variables clínicas en AD, los resultados del análisis confirman afirmaciones hechas en investigaciones previas a este trabajo, además de sugerir que su presencia no se debe a azar. A continuación mencionamos algunas observaciones establecidas a partir de ambas tablas.

1. **Educación:** El promedio del nivel de educación de los casos positivos (TP + FN) del grupo sCsAD_{ADNI-test} fue ligeramente menor a los casos negativos (TN + FP) de este grupo y a los casos negativos del grupo sMCI_{ADNI-test}, es decir los sujetos con sAD presentaron un nivel de educación ligeramente menor a los sujetos sC y a los sujetos sMCI_{Inc}. Asimismo, en el grupo sMCI_{ADNI-test}, el promedio de educación de los casos positivos (sujetos sMCI_C) fue menor que sus casos negativos.

Por otro lado, esta tendencia también fue observada en los dos grupos de OASIS, donde el promedio de educación de los casos positivos de los grupos sCsAD_{OASIS-45} y sCsAD_{OASIS-7} fue menor a sus respectivos casos negativos del mismo grupo

Los resultados de ADNI y OASIS tienen relación con la afirmación que los sujetos con AD tienden a tener un nivel de educación inferior a sus similares sanos o sujetos con MCI, debido a su deterioro cognitivo. Asimismo, la relación encontrada en el nivel de educación entre los sujetos sMCI_{Inc} y los sMCI_C, sugiere que el nivel de educación podría influir también en otros estados de demencia como el estado de MCI.

2. **Mini-Mental Score (MMSCORE).** En los cuatro grupos, se observó que el promedio del índice MMSCORE de sus casos respectivos casos positivos fue considerablemente menor al respectivo valor de sus casos negativos. Incluso, esta tendencia fue observada entre los casos positivos del grupo el grupo sCsAD_{ADNI-test} con respecto a los casos negativos y positivos de sMCI_{ADNI-test}. Estos resultados podrían justificarse con el hecho de que los sujetos con AD presentan un nivel cognitivo mucho más deteriorado en comparación a los adultos mayores sanos. Asimismo, la ligera diferencia en los promedios del índice MMSCORE entre los sujetos sMCI no-convertidos y los sMCI convertidos a AD, refleja que los sujetos con MCI, a pesar de no progresar a AD, podrían presentar un deterioro cognitivo similar a los sujetos convertidos a AD.
3. **Edad y género.** No se encontraron diferencias significativas ni tendencias de edad y género en el diagnóstico de AD o para la predicción de esta enfermedad.

Tabla 6.1: Influencia de las características sociodemográficas y neuropsicológicas en AD. Resultados sobre la matriz de confusión de los grupos

	sCsAD_{ADNI-test}				sMCI_{ADNI-test}				sCsAD_{OASIS-35}				sCsAD_{OASIS-7}			
	TP	TN	FP	FN	TP	TN	FP	FN	TP	TN	FP	FN	TP	TN	FP	FN
Nro. casos	554	582	9	0	171	598	35	10	18	15	1	1	3	2	2	0
Femenino	235	265	6	–	54	215	19	7	10	12	0	1	0	0	2	–
Masculino	319	317	3	–	117	383	16	3	8	3	1	0	3	2	0	–
Edad	76.86	77.58	86.67	–	75.61	77.11	76.03	76.90	75.67	73.80	80.00	69.00	77.00	87.50	89.00	–
Educación	15.17	16.04	15.33	–	16.02	15.81	16.51	15.90	2.56	3.40	4.00	2.00	12.00	14.00	18.00	–
MMSCORE	19.77	25.71	27.22	–	24.07	23.86	25.14	22.60	25.78	29.33	29.00	23.00	24.33	28.50	27.50	–

Tabla 6.2: Influencia de las características sociodemográficas y neuropsicológicas en AD. Resultados sobre todos los casos positivos y negativos de los grupos

	sCsAD_{ADNI-test}		sMCI_{ADNI-test}		sCsAD_{OASIS-35}		sCsAD_{OASIS-7}	
	sC (TN+FP)	sAD (TP+FN)	sMCI_{nc} (TN+FP)	sMCI_c (TP+FN)	sC (TN+FP)	sAD (TP+FN)	sC (TN+FP)	sAD (TP+FN)
Nro. casos	591	554	633	181	16	19	4	3
Femenino	271	235	234	61	12	11	2	0
Masculino	320	319	399	120	4	8	2	3
Edad	82.13	76.86	76.57	76.26	76.90	72.34	88.25	77
Educación	15.69	15.17	16.16	15.96	3.70	2.28	16	12
MMSCORE	26.47	19.77	24.50	23.34	29.17	24.39	28	24.33

6. ANÁLISIS DE RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO

6.2.4 Tiempo de procesamiento

Todos los cálculos necesarios para la construcción y utilización de los modelos de diagnóstico y predicción basados en SVM fueron llevados a cabo en un ordenador de 4 procesadores de 2.4 GHz cada uno y memoria RAM de 3.6 GB. Los tiempos de procesamiento de las tareas implicadas en la construcción de los modelos como la validación cruzada y aprendizaje de los modelos, estuvo en el orden de los minutos. Con respecto a la obtención de las variables predictoras, los tiempos dependieron de la clase de variable. Para las variables clínicas y morfométricas obtenidas directamente de las bases de datos, el único procesamiento realizado fue su conversión a un formato de número real y escalamiento al rango $[0,1]$, en este caso los tiempos de procesamiento estuvieron en el orden de los segundos. Sin embargo, con respecto a la variables morfométricas obtenidas de las MRIs longitudinales de los 7 casos del grupo sCsAD_{OASIS-7}, los tiempos de procesamiento fueron mayores debido al procesamiento y análisis de estas imágenes utilizando la herramienta desarrollada en MATLAB y las librerías de Freesurfer. Por cada uno de los casos de este grupo, el tiempo empleado en procesar las zonas corticales y no-corticales (volumen y espesor de la corteza cerebral) para obtener las mediciones de las 27 variables morfométricas fue de aproximadamente 14 horas.

6.3 Conclusiones

6.3.1 Máquinas de soporte vectorial (SVM) para el diagnóstico y predicción de AD

El porcentaje de casos clasificados incorrectamente (FP y FN) frente a los aciertos (TP y TN) de los modelos construidos confirman la eficiencia y utilidad de SVM como herramienta para entrenar modelos de soporte en el diagnóstico y predicción de AD. Asimismo, es importante resaltar que el entrenamiento de todos los modelos utilizando SVM no sólo resultó eficiente en las tareas de clasificación de los casos clínicos, también resultó rápido y versátil en los procesos implicados en su configuración y aprendizaje.

6.3.2 Variables predictoras seleccionadas

6.3.2.1 Correlación de las variables predictoras con otras investigaciones

Con la finalidad de determinar una interpretación común del diagnóstico de AD y replicar los resultados sobre otros pacientes, este trabajo estuvo orientado a integrar diferentes tipos de datos clínicos con información procedente de biomarcadores morfométricos. Como resultado, 29 variables fueron seleccionadas como más útiles en la discriminación de sujetos con AD de los sujetos de control mayores. Asimismo, estas variables fueron utilizadas para predecir los sujetos con MCI con posibilidad de convertirse a AD. Este subconjunto de variables predictoras tienen una estrecha relación con los resultados de investigaciones previas a este estudio [Cui et al., 2011; Desikan et al., 2009; Heckemann et al., 2011].

6.3.2.2 Viabilidad de la obtención de las variables de entrada

El subconjunto de 29 variables de entradas seleccionadas como más útiles en el diagnóstico y predicción de AD implica la necesidad de obtener información sobre las características demográficas, clínicas y morfométricas de los pacientes bajo estudio. Las variables demográfica y neuropsicológicas pueden ser obtenidas aplicando diversos test clínicos de amplio dominio por lo médicos, por lo que en la mayoría de casos es completamente viable su obtención. Con respecto a la obtención de las variables morfométricas a partir de las sesiones de MRIs del cerebro, su viabilidad está sujeto a la disposición de una herramienta de software capaz de leer, procesar y analizar estas imágenes a fin de cuantificar las variables morfométricas necesarias. En la sección 5.3.2.1 se mencionó brevemente

la utilización de una herramienta prototipo de procesamiento y análisis de MRIs desarrollada en MATLAB para este estudio, lo cual demuestra que es posible implementar los procesos implicados en el cálculo de estas variables sin la necesidad de contar con un software sofisticado o de laboratorios especializados. Sin embargo es importante resaltar la importancia y necesidad de contar con un ordenador con buenas prestaciones para el procesamiento y análisis de las MRIs, debido a que el tiempo de procesamiento y en consecuencia el tiempo de disposición de la información morfométrica de los pacientes, dependerá principalmente de sus características de hardware.

6.3.2.3 Enfoque de variables predictoras singulares vs. enfoque de variables predictoras combinadas

A efectos de obtener modelos con diagnóstico y predicción de AD, la combinación de variables clínicas con variables morfométricas del cerebro de los pacientes, resultó ser más efectiva y robusta en comparación con los modelos entrenados únicamente con alguno de estos dos tipos de variables. Como ejemplo, los modelos SVM entrenados sólo con variables clínicas: $SVM_{vClin-Wrapper}$, $SVM_{vClin-Fscore}$ y $SVM_{vClin-FscoreWrapper}$; obtuvieron los indicadores estadísticos más bajos, ver tabla 4.8. Asimismo, a pesar que el modelo $SVM_{vMRI-Wrapper}$, entrenado con 126 variables morfométricas, obtuvo los mejores resultados en la clasificación, la mayoría de estas variables presentaron valores bajos de AUC (menor a 0.6), por lo que este subconjunto no garantizaba la robustez de la clasificación frente a otros casos clínicos.

Finalmente, es importante resaltar que la eficiencia del enfoque mixto de variables no sólo se ve reflejado en los índices estadísticos obtenidos por los modelos SVM-Diag₂₉ y SVM-Pred₂₉, también se ve reflejada en la robustez de la clasificación de los modelos SVM-Diag₂₈, sCsAD_{OASIS-45} y sCsAD_{OASIS-7}, los cuales fueron entrenados con 28 de estas variables.

6.3.2.4 Importancia de los biomarcadores morfométricos en el diagnóstico y predicción de AD

Durante al etapa de selección de variables, los resultados han demostrado que los modelos SVM entrenados únicamente con variables morfométricas presentan mejores resultados en comparación a los modelos entrenados con variables clínicas (demográficas y neuropsicológicas), ver tabla 4.8. Estos resultados reafirman otras las observaciones realizadas en otras investigaciones [Risacher et al. \[2010\]](#), las cuales defienden la necesidad de considerar biomarcadores morfométricos obtenidos de MRIs para el diagnóstico y predicción de AD. Estos biomarcadores pueden servir como sustituto in vivo para conocer la gravedad de la enfermedad y como un método de diagnóstico no invasivo de AD y la predicción de la conversión de sujetos con deterioro cognitivo leve (MCI) hacia AD.

6.4 Limitaciones de este estudio

La población de AD es heterogénea. Una de las principales limitaciones de este trabajo es que la selección de variables predictoras y el entrenamiento de los principales modelos SVM para el diagnóstico y predicción de AD han sido realizadas únicamente sobre los casos clínicos obtenidos de ADNI. Estos casos de ADNI representan una población “homogénea” de ensayos clínicos, previamente preparados y seleccionados para investigaciones de AD y no una muestra epidemiológica de una población de la vida real. En la vida real, las personas con AD o MCI constituyen una población heterogénea, lo cual se manifiesta en diferentes niveles de los pacientes (niveles pato-bioquímicos, patológico, clínico, etc.)[\[Teunissen and Scheltens, 2004\]](#). Por ejemplo, los pacientes con AD pueden estar afectados por otras enfermedades o por factores clínicos, sociológicos e incluso ambientales; mientras que en ADNI, han sido descartados los participantes afectados por algunos de estos factores.

6. ANÁLISIS DE RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO

AD puede afectar a personas menores de 55 años. Una segunda limitación es que en nuestro estudio, principalmente en la etapa del aprendizaje de los modelos SVM, ha sido realizada sobre casos clínicos de participantes ADNI con un rango de edades de 55 a 90 años. Al respecto hay considerable evidencia que las personas pueden ser afectadas por AD antes de los 55 años. Sin embargo, cabe resaltar que durante la etapa de test con los grupos de OASIS, algunos casos clínicos pertenecieron a sujetos con edades menores y el modelo clasificó correctamente la mayoría de estos casos.

Exclusión de otros tipos de biomarcadores de AD. Asimismo, otra limitación es haber incluido en este estudio sólo dos tipos de variables: variables clínicas y variables morfométricas obtenidas de biomarcadores basados en MRIs. En la actualidad, existe un número superior de características demográficas, neuropsicológicas y biomarcadores de AD que no han sido incluidas en este estudio.

Clasificación enfocada a discriminar entre sujetos sanos y sujetos con AD . Los biomarcadores morfométricos considerados en este estudio y en consecuencia los modelos SVM construidos a partir de estos, han sido diseñados principalmente para comparar los sujetos con AD con los sujetos sanos mayores, sin embargo, en la práctica clínica, además de esta discriminación, se hace necesario discriminar los pacientes AD con pacientes afectados con otras enfermedades neurodegenerativas [Teunissen and Scheltens, 2004].

6.5 Trabajo futuro

6.5.1 Inclusión de nuevos biomarcadores de AD

Para mejorar la eficacia de la clasificación, es posible analizar el desempeño de nuevas variables morfométricas obtenidas de otro tipo de imágenes médicas como las MRIs funcionales que incluyen agentes marcadores de contraste o las imágenes PET. Al respecto existen diversas publicaciones que sustentan el uso de estos tipos de imágenes médicas ya sea de manera exclusiva [Herholz et al., 2011], en combinación con los biomarcadores morfométricos obtenidos de MRIs estructurales [Dukart et al., 2011; Jack et al., 2010b; Lo et al., 2011] o en combinación con otro tipo de biomarcadores como los genéticos o CSF [Caroli et al., 2012], a fin de obtener información del grado de atrofia cerebral o el deterioro funcional de los afectados con MCI y AD. Asimismo, en el proyecto ADNI se tiene disponibilidad de biomarcadores adicionales obtenidos a partir de imágenes PET, lo cual, realizando un proceso de análisis y selección de variables descrito en este trabajo, podrían ser incluidas como nuevas variables de entrada.

Por otro lado, existen otros tipos de biomarcadores como los genéticos, químicos y biológicos, entre otros, los cuales no han sido considerados en este estudio y que podrían ser de gran utilidad en la clasificación clínica de sujetos con AD o en un estado de transición a esta enfermedad. Al igual que en el caso de los biomarcadores obtenidos de las imágenes PET, existen una amplia investigación en estos biomarcadores, asimismo, proyectos dedicados a la investigación de AD como ADNI, proporciona información sobre estos y establecen protocolos de cómo obtenerlos y analizarlos.

6.5.2 Desarrollo e integración de un módulo de procesamiento y análisis de imágenes cerebrales.

En este trabajo se desarrolló una herramienta en versión prototipo de procesamiento y análisis de MRIs estructurales basándonos en las técnicas descritas en la sección 2.4.2. Para su diseño y codificación fueron utilizados la GUIDE de MATLAB y el *toolbox* de procesamiento de imágenes de MATLAB [MathWorks, 2010]. Asimismo, para el análisis morfológico de las ROIs del cerebro y el cálculo final de las variables morfométricas, esta herramienta incluyó las librerías de procesamiento del software libre Freesurfer [Fischl, 2012]. Un siguiente paso sería

el desarrollo de una versión más estable de esta herramienta en otro lenguaje de programación, la cual facilitara el rápido procesamiento de varias sesiones de imágenes y la ejecución en paralelo de los procesos implicados, ver figura 6.1. Asimismo, se hace necesario construir una interfaz de usuario amigable que permita a los clínicos manipular libremente y sin complicaciones estos procesos.

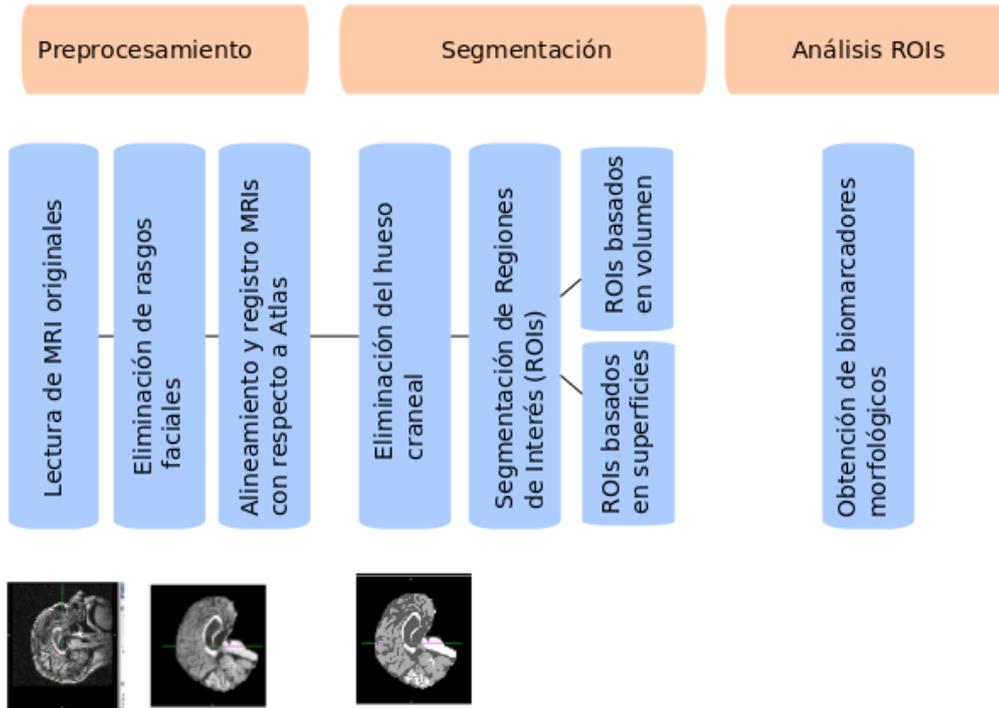


Figura 6.1: Vista general de los principales procesos implicados en la obtención de biomarcadores morfológicos basados en MRIs

6. ANÁLISIS DE RESULTADOS, CONCLUSIONES Y TRABAJO FUTURO

Apéndice A: Tablas y Gráficos

.1 Resultados de los métodos de selección de variables sobre las 318 variables morfológicas originales. Calificación final

Tabla 3: Calificación final de las 318 variables morfológicas originales. **CTAvg**:Cortical Thickness Average, **VCP**:Volume from Cortical Parcellation, **VWM**:Volume from White Matter, **CTStd**:Cortical Thickness Standard Deviation

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
1	ST24TA	CTAvg of LeftEntorhinal	0.8704	0.0322	✓	✓	✓	3
2	ST90TA	CTAvg of RightInferiorParietal	0.7977	0.0143	✗	✗	✗	0
3	ST88SV	VWM of RightHippocampus	0.8641	0.0123	✗	✓	✗	1
4	ST32TA	CTAvg of LeftInferiorTemporal	0.8328	0.0280	✓	✗	✗	1
5	ST83TA	CTAvg of RightEntorhinal	0.8671	0.0403	✓	✓	✗	2
6	ST83CV	VCP of RightEntorhinal	0.8252	0.0241	✓	✓	✗	2
7	ST13TA	CTAvg of LeftBankssts	0.7916	0.0283	✓	✓	✓	3
8	ST29SV	VWM of LeftHippocampus	0.886	0.0204	✗	✓	✗	1
9	ST99TA	CTAvg of RightMiddleTemporal	0.855	0.0018	✗	✓	✗	1
10	ST71SV	VWM of RightAmygdala	0.7978	0.0014	✗	✓	✗	1
11	ST85TA	CTAvg of RightFusiform	0.7998	0.0555	✓	✓	✗	2
12	ST103TA	CTAvg of RightParahippocampal	0.7633	0.7329	✓	✓	✓	3
13	ST40TA	CTAvg of LeftMiddleTemporal	0.8682	0.0090	✗	✓	✗	1
14	ST12SV	VWM of LeftAmygdala	0.8056	0.0768	✓	✓	✓	3
15	ST26TA	CTAvg of LeftFusiform	0.7986	0.0001	✗	✗	✗	0
16	ST44TA	CTAvg of LeftParahippocampal	0.7716	0.0651	✓	✗	✗	1
17	ST91TA	CTAvg of RightInferiorTemporal	0.8215	0.0000	✗	✓	✗	1
18	ST40CV	VCP of LeftMiddleTemporal	0.7897	0.0013	✗	✗	✗	0
19	ST93TA	CTAvg of RightIsthmusCingulate	0.7385	0.0004	✗	✓	✗	1
20	ST72TA	CTAvg of RightBankssts	0.7781	0.0112	✗	✓	✗	1

continúa en la siguiente página

. APÉNDICE A: TABLAS Y GRÁFICOS

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
21	ST44CV	VCP of LeftParahippocampal	0.7574	0.0136	✗	✓	✗	1
22	ST117TA	CTAvg of RightSuperiorTemporal	0.7934	0.0347	✓	✓	✗	2
23	ST114TA	CTAvg of RightRostralMiddleFrontal	0.751	0.1801	✓	✓	✓	3
24	ST31CV	VCP of LeftInferiorParietal	0.7444	0.0096	✗	✓	✗	1
25	ST58TA	CTAvg of LeftSuperiorTemporal	0.804	0.0345	✓	✓	✗	2
26	ST32CV	VCP of LeftInferiorTemporal	0.7537	0.0209	✗	✓	✗	1
27	ST111TA	CTAvg of RightPrecuneus	0.7689	0.1311	✓	✓	✗	2
28	ST83TS	CTStd of RightEntorhinal	0.6407	0.0045	✗	✓	✗	1
29	ST60TA	CTAvg of LeftTemporalPole	0.7767	0.0044	✗	✗	✗	0
30	ST31TA	CTAvg of LeftInferiorParietal	0.7984	0.0034	✗	✗	✗	0
31	ST103CV	VCP of RightParahippocampal	0.7515	0.4283	✓	✓	✓	3
32	ST90CV	VCP of RightInferiorParietal	0.7477	0.0027	✗	✓	✗	1
33	ST99CV	VCP of RightMiddleTemporal	0.7616	0.0000	✗	✓	✗	1
34	ST24CV	VCP of LeftEntorhinal	0.8303	0.0965	✓	✗	✓	2
35	ST25TA	CTAvg of LeftFrontalPole	0.6576	0.0753	✓	✓	✗	2
36	ST34TA	CTAvg of LeftIsthmusCingulate	0.7471	0.0016	✗	✓	✗	1
37	ST85CV	VCP of RightFusiform	0.7465	0.0115	✗	✓	✗	1
38	ST3SV	VWM of CorpusCallosumCentral	0.6493	0.0039	✗	✗	✗	0
39	ST59TA	CTAvg of LeftSupramarginal	0.7884	0.0054	✗	✗	✗	0
40	ST119TA	CTAvg of RightTemporalPole	0.7622	0.1163	✓	✓	✗	2
41	ST109TA	CTAvg of RightPosteriorCingulate	0.6973	0.2066	✓	✗	✗	1
42	ST91CV	VCP of RightInferiorTemporal	0.7585	0.0008	✗	✓	✗	1
43	ST39TA	CTAvg of LeftMedialOrbitofrontal	0.7104	0.0129	✗	✓	✗	1
44	ST55TA	CTAvg of LeftRostralMiddleFrontal	0.7586	0.0013	✗	✓	✗	1
45	ST34TS	CTStd of LeftIsthmusCingulate	0.651	0.0588	✓	✓	✓	3
46	ST129TA	CTAvg of LeftInsula	0.7125	0.0105	✗	✓	✗	1
47	ST57TA	CTAvg of LeftSuperiorParietal	0.704	0.0900	✓	✓	✓	3
48	ST37SV	VWM of LeftLateralVentricle	0.7068	0.0243	✓	✓	✗	2
49	ST111TS	CTStd of RightPrecuneus	0.6909	0.2480	✓	✗	✓	2
50	ST59CV	VCP of LeftSupramarginal	0.7211	0.0274	✓	✓	✗	2
51	ST130TA	CTAvg of RightInsula	0.7263	0.0107	✗	✗	✗	0
52	ST26CV	VCP of LeftFusiform	0.7082	0.0457	✓	✓	✗	2
53	ST128SV	VWM of WMHypoIntensities	0.6464	0.1331	✓	✓	✗	2
54	ST89SV	VWM of RightInferiorLateralVentricle	0.8167	0.0437	✓	✗	✗	1
55	ST118TA	CTAvg of RightSupramarginal	0.7551	0.0951	✓	✓	✓	3
56	ST60CV	VCP of LeftTemporalPole	0.6599	0.0062	✗	✗	✗	0
57	ST60TS	CTStd of LeftTemporalPole	0.6363	0.0018	✗	✗	✗	0

continúa en la siguiente página

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
58	ST5SV	VWM of CorpusCallosumMidPosterior	0.6527	0.0347	✓	✓	✓	3
59	ST98TA	CTAvg of RightMedialOrbitofrontal	0.7141	0.0000	✗	✓	✗	1
60	ST34CV	VCP of LeftIsthmusCingulate	0.6665	0.0775	✓	✗	✗	1
61	ST84TA	CTAvg of RightFrontalPole	0.6644	0.0011	✗	✗	✗	0
62	ST119TS	CTStd of RightTemporalPole	0.6075	0.0194	✗	✓	✗	1
63	ST52TA	CTAvg of LeftPrecuneus	0.7506	0.0187	✗	✓	✗	1
64	ST96SV	VWM of RightLateralVentricle	0.7053	0.0000	✗	✓	✗	1
65	ST68SV	VWM of NonWMHypoIntensities	0.6108	0.0218	✗	✓	✗	1
66	ST97TA	CTAvg of RightLingual	0.6737	0.0002	✗	✗	✗	0
67	ST24TS	CTStd of LeftEntorhinal	0.5866	0.0047	✗	✗	✗	0
68	ST94TA	CTAvg of RightLateralOccipital	0.7031	0.0010	✗	✗	✗	0
69	ST93CV	VCP of RightIsthmusCingulate	0.6756	0.0048	✗	✓	✗	1
70	ST105TA	CTAvg of RightParsOrbitalis	0.6849	0.2852	✓	✓	✓	3
71	ST129TS	CTStd of LeftInsula	0.5875	0.1274	✓	✓	✗	2
72	ST51SA	Surface Area of LeftPrecentral	0.527	0.0031	✗	✗	✗	0
73	ST58CV	VCP of LeftSuperiorTemporal	0.71	0.0055	✗	✗	✗	0
74	ST15TA	CTAvg of LeftCaudalMiddleFrontal	0.7472	0.0153	✓	✓	✗	2
75	ST14TS	CTStd of LeftCaudalAnteriorCingulate	0.5756	0.0076	✗	✗	✗	0
76	ST116TA	CTAvg of RightSuperiorParietal	0.7017	0.0792	✓	✓	✗	2
77	ST75SV	VWM of RightCaudate	0.5028	0.0877	✗	✗	✗	0
78	ST117CV	VCP of RightSuperiorTemporal	0.6924	0.0000	✓	✗	✓	2
79	ST38TA	CTAvg of LeftLingual	0.6721	0.1303	✓	✗	✓	2
80	ST50TA	CTAvg of LeftPosteriorCingulate	0.6962	0.0839	✗	✓	✗	1
81	ST72CV	VCP of RightBankssts	0.6998	0.0098	✗	✓	✗	1
82	ST95TA	CTAvg of RightLateralOrbitofrontal	0.6941	0.0001	✓	✓	✗	2
83	ST73TS	CTStd of RightCaudalAnteriorCingulate	0.5534	0.0372	✓	✓	✓	3
84	ST74TA	CTAvg of RightCaudalMiddleFrontal	0.7437	0.0372	✗	✓	✗	1
85	ST4SV	VWM of CorpusCallosumMidAnterior	0.6536	0.0063	✗	✗	✗	0
86	ST52CV	VCP of LeftPrecuneus	0.7088	0.0018	✗	✓	✗	1
87	ST30SV	VWM of LeftInferiorLateralVentricle	0.8267	0.0722	✓	✓	✓	3
88	ST104TS	CTStd of RightParsOpercularis	0.5585	0.5966	✓	✓	✓	3
89	ST115TA	CTAvg of RightSuperiorFrontal	0.7536	0.2771	✓	✓	✓	3
90	ST7SV	VWM of Csf	0.6975	0.0085	✗	✗	✗	0
91	ST119CV	VCP of RightTemporalPole	0.6494	0.1338	✓	✓	✗	2
92	ST84CV	VCP of RightFrontalPole	0.5927	0.0000	✗	✓	✗	1
93	ST113TS	CTStd of RightRostralAnteriorCingulate	0.5623	0.1414	✓	✓	✗	2
94	ST16SV	VWM of LeftCaudate	0.5024	0.1982	✓	✓	✗	2

continúa en la siguiente página

. APÉNDICE A: TABLAS Y GRÁFICOS

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
95	ST101SV	VWM of RightPallidum	0.57	0.6709	✓	✓	✓	3
96	ST121TS	CTStd of RightTransverseTemporal	0.5588	0.0029	✓	✓	✓	3
97	ST44TS	CTStd of LeftParahippocampal	0.5677	0.1203	✗	✓	✗	1
98	ST115CV	VCP of RightSuperiorFrontal	0.6452	0.1488	✓	✓	✗	2
99	ST52TS	CTStd of LeftPrecuneus	0.6598	0.0297	✓	✓	✗	2
100	ST111CV	VCP of RightPrecuneus	0.7032	0.3266	✓	✗	✗	1
101	ST80SV	VWM of RightChoroidPlexus	0.643	0.0348	✓	✗	✗	1
102	ST56TA	CTAvg of LeftSuperiorFrontal	0.7369	0.0432	✓	✗	✗	1
103	ST45TS	CTStd of LeftParsOpercularis	0.5857	0.0378	✓	✗	✗	1
104	ST36TA	CTAvg of LeftLateralOrbitofrontal	0.6896	0.0701	✓	✓	✗	2
105	ST69SV	VWM of OpticChiasm	0.5767	0.0003	✗	✓	✗	1
106	ST19SV	VWM of LeftCerebralCortex	0.687	0.0155	✗	✗	✗	0
107	ST54TS	CTStd of LeftRostralAnteriorCingulate	0.5745	0.0012	✗	✓	✗	1
108	ST60SA	Surface Area of LeftTemporalPole	0.5108	0.0060	✗	✓	✗	1
109	ST84SA	Surface Area of RightFrontalPole	0.5033	0.0019	✗	✓	✗	1
110	ST11SV	VWM of LeftAccumbensArea	0.6312	0.0043	✗	✓	✗	1
111	ST35TA	CTAvg of LeftLateralOccipital	0.6843	0.0462	✓	✓	✗	2
112	ST72TS	CTStd of RightBankssts	0.5421	0.0000	✗	✓	✗	1
113	ST46TA	CTAvg of LeftParsOrbitalis	0.6685	0.0229	✓	✓	✗	2
114	ST42SV	VWM of LeftPallidum	0.5853	0.0134	✗	✗	✗	0
115	ST106CV	VCP of RightParsTriangularis	0.602	0.3076	✓	✓	✗	2
116	ST23TS	CTStd of LeftCuneus	0.5292	0.0043	✗	✓	✗	1
117	ST124SV	VWM of RightVentralDC	0.6227	0.0022	✗	✓	✗	1
118	ST127SV	VWM of ThirdVentricle	0.6421	0.2154	✓	✓	✓	3
119	ST90TS	CTStd of RightInferiorParietal	0.6574	0.0220	✓	✗	✗	1
120	ST103SA	Surface Area of RightParahippocampal	0.5611	0.3325	✓	✗	✓	2
121	ST91SA	Surface Area of RightInferiorTemporal	0.6229	0.0048	✗	✗	✗	0
122	ST49TA	CTAvg of LeftPostcentral	0.666	0.0006	✗	✗	✗	0
123	ST82TS	CTStd of RightCuneus	0.5199	0.0193	✗	✓	✗	1
124	ST109CV	VCP of RightPosteriorCingulate	0.6395	0.3405	✓	✗	✓	2
125	ST104TA	CTAvg of RightParsOpercularis	0.6987	0.2450	✓	✗	✓	2
126	ST23TA	CTAvg of LeftCuneus	0.5954	0.1149	✓	✗	✗	1
127	ST48TA	CTAvg of LeftPericalcarine	0.6103	0.0000	✗	✗	✗	0
128	ST70SV	VWM of RightAccumbensArea	0.642	0.0111	✗	✗	✗	0
129	ST91TS	CTStd of RightInferiorTemporal	0.5573	0.0064	✗	✗	✗	0
130	ST103TS	CTStd of RightParahippocampal	0.5589	0.0159	✓	✗	✗	1
131	ST48TS	CTStd of LeftPericalcarine	0.5749	0.0012	✗	✗	✗	0

continúa en la siguiente página

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selett.
132	ST39TS	CTStd of LeftMedialOrbitofrontal	0.5611	0.5345	✗	✗	✗	0
133	ST97TS	CTStd of RightLingual	0.5769	0.0004	✗	✗	✗	0
134	ST45TA	CTAvg of LeftParsOpercularis	0.6841	0.0000	✗	✗	✗	0
135	ST40SA	Surface Area of LeftMiddleTemporal	0.6035	0.0453	✓	✗	✗	1
136	ST13SA	Surface Area of LeftBankssts	0.5742	0.0676	✓	✗	✗	1
137	ST31TS	CTStd of LeftInferiorParietal	0.6614	0.0119	✗	✓	✗	1
138	ST108SA	Surface Area of RightPostcentral	0.5197	0.0643	✓	✗	✗	1
139	ST106TA	CTAvg of RightParsTriangularis	0.6836	0.2665	✓	✗	✓	2
140	ST98CV	VCP of RightMedialOrbitofrontal	0.6365	0.0019	✗	✗	✗	0
141	ST108TA	CTAvg of RightPostcentral	0.655	0.3015	✓	✗	✓	2
142	ST50CV	VCP of LeftPosteriorCingulate	0.6607	0.0680	✓	✗	✗	1
143	ST21SV	VWM of LeftChoroidPlexus	0.6319	0.0021	✗	✗	✗	0
144	ST130TS	CTStd of RightInsula	0.5708	0.2193	✓	✗	✓	2
145	ST113TA	CTAvg of RightRostralAnteriorCingulate	0.5986	0.1577	✓	✗	✓	2
146	ST25CV	VCP of LeftFrontalPole	0.5791	0.0473	✓	✗	✗	1
147	ST130CV	VCP of RightInsula	0.6478	0.1456	✓	✗	✓	2
148	ST102SA	Surface Area of RightParacentral	0.5491	0.6057	✓	✗	✗	1
149	ST104CV	VCP of RightParsOpercularis	0.587	0.3316	✓	✗	✓	2
150	ST47TA	CTAvg of LeftParsTriangularis	0.665	0.0010	✗	✗	✗	0
151	ST108CV	VCP of RightPostcentral	0.5823	0.2358	✓	✗	✓	2
152	ST15CV	VCP of LeftCaudalMiddleFrontal	0.6559	0.0722	✓	✗	✓	2
153	ST113SA	Surface Area of RightRostralAnteriorCingulate	0.551	0.0678	✓	✗	✗	1
154	ST110TA	CTAvg of RightPrecentral	0.67	0.2118	✓	✗	✓	2
155	ST118CV	VCP of RightSupramarginal	0.6689	0.1234	✓	✗	✓	2
156	ST83SA	Surface Area of RightEntorhinal	0.5859	0.0009	✗	✗	✗	0
157	ST45CV	VCP of LeftParsOpercularis	0.5674	0.0050	✗	✗	✗	0
158	ST107TS	CTStd of RightPericalcarine	0.5377	0.3568	✓	✗	✓	2
159	ST25TS	CTStd of LeftFrontalPole	0.5164	0.0070	✗	✗	✗	0
160	ST105CV	VCP of RightParsOrbitalis	0.6218	0.3674	✓	✗	✓	2
161	ST14TA	CTAvg of LeftCaudalAnteriorCingulate	0.554	0.0159	✗	✗	✗	0
162	ST97SA	Surface Area of RightLingual	0.5216	0.0212	✗	✗	✗	0
163	ST115TS	CTStd of RightSuperiorFrontal	0.5671	0.2344	✓	✗	✓	2
164	ST25SA	Surface Area of LeftFrontalPole	0.5188	0.0660	✓	✗	✗	1
165	ST43TS	CTStd of LeftParacentral	0.5261	0.0276	✓	✗	✗	1
166	ST58TS	CTStd of LeftSuperiorTemporal	0.5629	0.0170	✗	✗	✗	0
167	ST32SA	Surface Area of LeftInferiorTemporal	0.5969	0.1217	✓	✗	✓	2
168	ST94TS	CTStd of RightLateralOccipital	0.5668	0.0000	✗	✗	✗	0

continúa en la siguiente página

. APÉNDICE A: TABLAS Y GRÁFICOS

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selett.
169	ST9SV	VWM of FourthVentricle	0.5344	0.0037	✗	✗	✗	0
170	ST56CV	VCP of LeftSuperiorFrontal	0.6476	0.0029	✗	✓	✗	1
171	ST13TS	CTStd of LeftBankssts	0.5047	0.0002	✗	✗	✗	0
172	ST109TS	CTStd of RightPosteriorCingulate	0.5097	0.2516	✓	✓	✓	3
173	ST13CV	VCP of LeftBankssts	0.7055	0.1079	✓	✓	✓	3
174	ST46TS	CTStd of LeftParsOrbitalis	0.5004	0.0014	✗	✓	✗	1
175	ST36CV	VCP of LeftLateralOrbitofrontal	0.6139	0.0335	✓	✓	✗	2
176	ST84TS	CTStd of RightFrontalPole	0.5239	0.1128	✗	✗	✗	0
177	ST45SA	Surface Area of LeftParsOpercularis	0.53	0.1664	✓	✗	✓	2
178	ST107TA	CTAvg of RightPericalcarine	0.5875	0.1988	✓	✓	✗	2
179	ST114TS	CTStd of RightRostralMiddleFrontal	0.5847	0.0052	✓	✗	✗	1
180	ST55CV	VCP of LeftRostralMiddleFrontal	0.6404	0.0376	✓	✗	✗	1
181	ST6SV	VWM of CorpusCallosumPosterior	0.5976	0.0019	✗	✓	✗	1
182	ST15TS	CTStd of LeftCaudalMiddleFrontal	0.5006	0.0637	✓	✓	✓	3
183	ST104SA	Surface Area of RightParsOpercularis	0.5039	0.0717	✓	✗	✓	2
184	ST26TS	CTStd of LeftFusiform	0.5545	0.3356	✓	✗	✗	1
185	ST1SV	VWM of Brainstem	0.556	0.1172	✓	✗	✓	2
186	ST119SA	Surface Area of RightTemporalPole	0.5121	0.1051	✓	✗	✗	1
187	ST105TS	CTStd of RightParsOrbitalis	0.5126	0.4054	✓	✗	✓	2
188	ST93TS	CTStd of RightIsthmusCingulate	0.6142	0.0001	✗	✓	✗	1
189	ST38CV	VCP of LeftLingual	0.5764	0.0924	✓	✓	✓	3
190	ST106SA	Surface Area of RightParsTriangularis	0.5097	0.1975	✓	✗	✗	1
191	ST43SA	Surface Area of LeftParacentral	0.5453	0.4565	✗	✗	✗	0
192	ST62TA	CTAvg of LeftTransverseTemporal	0.6197	0.0034	✗	✗	✗	0
193	ST129CV	VCP of LeftInsula	0.6357	0.1292	✓	✗	✗	1
194	ST102TA	CTAvg of RightParacentral	0.6248	0.0133	✓	✗	✓	2
195	ST38TS	CTStd of LeftLingual	0.5567	0.0122	✗	✗	✗	0
196	ST54TA	CTAvg of LeftRostralAnteriorCingulate	0.6085	0.0087	✗	✗	✗	0
197	ST73SA	Surface Area of RightCaudalAnteriorCingulate	0.5663	0.0020	✗	✓	✗	1
198	ST61SV	VWM of LeftThalamus	0.5503	0.0013	✗	✓	✗	1
199	ST47CV	VCP of LeftParsTriangularis	0.6336	0.0670	✓	✗	✗	1
200	ST24SA	Surface Area of LeftEntorhinal	0.601	0.0218	✗	✗	✗	0
201	ST49SA	Surface Area of LeftPostcentral	0.52	0.0175	✗	✗	✗	0
202	ST40TS	CTStd of LeftMiddleTemporal	0.555	0.0006	✗	✓	✗	1
203	ST72SA	Surface Area of RightBankssts	0.5851	0.0226	✓	✗	✗	1
204	ST46CV	VCP of LeftParsOrbitalis	0.5964	0.0022	✗	✗	✗	0
205	ST56TS	CTStd of LeftSuperiorFrontal	0.5556	0.0007	✗	✓	✗	1

continúa en la siguiente página

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
206	ST65SV	VWM of LeftVentralDC	0.6008	0.0002	✗	✓	✗	1
207	ST62SA	Surface Area of LeftTransverseTemporal	0.5301	0.0044	✗	✗	✗	0
208	ST49TS	CTStd of LeftPostcentral	0.5751	0.0569	✓	✗	✓	2
209	ST53SV	VWM of LeftPutamen	0.6173	0.0872	✓	✗	✓	2
210	ST15SA	Surface Area of LeftCaudalMiddleFrontal	0.5143	0.0319	✓	✗	✗	1
211	ST48CV	VCP of LeftPericalcarine	0.537	0.0000	✗	✗	✗	0
212	ST117TS	CTStd of RightSuperiorTemporal	0.5224	0.2251	✓	✗	✓	2
213	ST50TS	CTStd of LeftPosteriorCingulate	0.5184	0.0445	✓	✗	✗	1
214	ST14SA	Surface Area of LeftCaudalAnteriorCingulate	0.5283	0.0178	✗	✗	✗	0
215	ST62TS	CTStd of LeftTransverseTemporal	0.5053	0.0368	✓	✗	✗	1
216	ST79SV	VWM of RightCerebralWM	0.6244	0.0604	✓	✗	✗	1
217	ST102TS	CTStd of RightParacentral	0.5109	0.6394	✓	✗	✓	2
218	ST97CV	VCP of RightLingual	0.583	0.0016	✗	✗	✗	0
219	ST51TA	CTAvg of LeftPrecentral	0.6605	0.0008	✓	✗	✓	2
220	ST95CV	VCP of RightLateralOrbitofrontal	0.6136	0.0463	✗	✗	✗	0
221	ST85TS	CTStd of RightFusiform	0.5267	0.0001	✗	✗	✗	0
222	ST57SA	Surface Area of LeftSuperiorParietal	0.5356	0.0374	✓	✓	✓	3
223	ST47TS	CTStd of LeftParsTriangularis	0.5032	0.0019	✗	✗	✗	0
224	ST82TA	CTAvg of RightCuneus	0.5912	0.0966	✓	✗	✗	1
225	ST39CV	VCP of LeftMedialOrbitofrontal	0.6012	0.0085	✓	✗	✓	2
226	ST57TS	CTStd of LeftSuperiorParietal	0.5791	0.0362	✗	✗	✗	0
227	ST95TS	CTStd of RightLateralOrbitofrontal	0.5357	0.0007	✗	✗	✗	0
228	ST98TS	CTStd of RightMedialOrbitofrontal	0.5493	0.0006	✗	✗	✗	0
229	ST18SV	VWM of LeftCerebellumWM	0.5167	0.1087	✓	✗	✓	2
230	ST74CV	VCP of RightCaudalMiddleFrontal	0.6088	0.0003	✗	✗	✗	0
231	ST26SA	Surface Area of LeftFusiform	0.5577	0.0890	✓	✗	✗	1
232	ST47SA	Surface Area of LeftParsTriangularis	0.5578	0.0306	✓	✗	✗	1
233	ST107CV	VCP of RightPericalcarine	0.5502	0.3207	✓	✗	✓	2
234	ST46SA	Surface Area of LeftParsOrbitalis	0.5223	0.0210	✗	✗	✗	0
235	ST36TS	CTStd of LeftLateralOrbitofrontal	0.52	0.0104	✗	✗	✗	0
236	ST78SV	VWM of RightCerebralCortex	0.685	0.0003	✗	✗	✗	0
237	ST118TS	CTStd of RightSupramarginal	0.5299	0.0220	✓	✗	✗	1
238	ST121CV	VCP of RightTransverseTemporal	0.567	0.2128	✓	✓	✗	2
239	ST112SV	VWM of RightPutamen	0.6165	0.1374	✓	✗	✓	2
240	ST44SA	Surface Area of LeftParahippocampal	0.5171	0.0160	✗	✗	✗	0
241	ST2SV	VWM of CorpusCallosumAnterior	0.5773	0.0219	✗	✓	✗	1
242	ST111SA	Surface Area of RightPrecuneus	0.5259	0.0033	✓	✗	✓	2

continúa en la siguiente página

. APÉNDICE A: TABLAS Y GRÁFICOS

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
243	ST74TS	CTStd of RightCaudalMiddleFrontal	0.5063	0.2396	✗	✗	✗	0
244	ST35TS	CTStd of LeftLateralOccipital	0.5549	0.0159	✗	✗	✗	0
245	ST74SA	Surface Area of RightCaudalMiddleFrontal	0.5205	0.0040	✗	✗	✗	0
246	ST99TS	CTStd of RightMiddleTemporal	0.5071	0.0016	✗	✓	✗	1
247	ST94CV	VCP of RightLateralOccipital	0.5907	0.0013	✗	✗	✗	0
248	ST116SA	Surface Area of RightSuperiorParietal	0.5081	0.0000	✓	✗	✗	1
249	ST73TA	CTAvg of RightCaudalAnteriorCingulate	0.5592	0.0529	✗	✓	✗	1
250	ST32TS	CTStd of LeftInferiorTemporal	0.5337	0.0078	✗	✗	✗	0
251	ST14CV	VCP of LeftCaudalAnteriorCingulate	0.5246	0.0029	✗	✗	✗	0
252	ST93SA	Surface Area of RightIsthmusCingulate	0.5006	0.0041	✗	✓	✗	1
253	ST43TA	CTAvg of LeftParacentral	0.6069	0.0000	✗	✗	✗	0
254	ST73CV	VCP of RightCaudalAnteriorCingulate	0.509	0.0050	✗	✗	✗	0
255	ST106TS	CTStd of RightParsTriangularis	0.5342	0.2395	✓	✗	✓	2
256	ST59TS	CTStd of LeftSupramarginal	0.5673	0.0008	✗	✗	✗	0
257	ST121SA	Surface Area of RightTransverseTemporal	0.5325	0.0220	✗	✗	✗	0
258	ST117SA	Surface Area of RightSuperiorTemporal	0.5037	0.0785	✓	✗	✓	2
259	ST118SA	Surface Area of RightSupramarginal	0.5202	0.0374	✓	✗	✗	1
260	ST49CV	VCP of LeftPostcentral	0.6002	0.0258	✓	✗	✗	1
261	ST113CV	VCP of RightRostralAnteriorCingulate	0.5256	0.2416	✓	✗	✓	2
262	ST116CV	VCP of RightSuperiorParietal	0.6527	0.0782	✓	✗	✓	2
263	ST110TS	CTStd of RightPrecentral	0.5506	0.0767	✓	✓	✓	3
264	ST51CV	VCP of LeftPrecentral	0.6142	0.0118	✗	✗	✗	0
265	ST55TS	CTStd of LeftRostralMiddleFrontal	0.583	0.0097	✗	✗	✗	0
266	ST31SA	Surface Area of LeftInferiorParietal	0.5756	0.0011	✗	✗	✗	0
267	ST121TA	CTAvg of RightTransverseTemporal	0.5995	0.1027	✓	✗	✓	2
268	ST54CV	VCP of LeftRostralAnteriorCingulate	0.5706	0.0094	✗	✗	✗	0
269	ST116TS	CTStd of RightSuperiorParietal	0.6006	0.1474	✓	✓	✗	2
270	ST23SA	Surface Area of LeftCuneus	0.5347	0.0522	✓	✗	✗	1
271	ST110CV	VCP of RightPrecentral	0.6195	0.0737	✓	✗	✓	2
272	ST39SA	Surface Area of LeftMedialOrbitofrontal	0.539	0.0862	✓	✗	✓	2
273	ST48SA	Surface Area of LeftPericalcarine	0.5039	0.0078	✗	✗	✗	0
274	ST43CV	VCP of LeftParacentral	0.5205	0.0013	✗	✗	✗	0
275	ST77SV	VWM of RightCerebellumWM	0.5166	0.0266	✓	✗	✗	1
276	ST120SV	VWM of RightThalamus	0.5549	0.1509	✓	✗	✓	2
277	ST90SA	Surface Area of RightInferiorParietal	0.5617	0.0018	✗	✗	✗	0
278	ST62CV	VCP of LeftTransverseTemporal	0.5541	0.1146	✓	✗	✗	1
279	ST114CV	VCP of RightRostralMiddleFrontal	0.6319	0.1246	✓	✓	✓	3

continúa en la siguiente página

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selett.
280	ST23CV	VCP of LeftCuneus	0.5023	0.0002	✗	✓	✗	1
281	ST34SA	Surface Area of LeftIsthmusCingulate	0.5296	0.0729	✓	✗	✗	1
282	ST108TS	CTStd of RightPostcentral	0.5658	0.2758	✓	✗	✓	2
283	ST20SV	VWM of LeftCerebralWM	0.6229	0.0167	✗	✓	✗	1
284	ST82SA	Surface Area of RightCuneus	0.5175	0.0264	✓	✓	✗	2
285	ST51TS	CTStd of LeftPrecentral	0.5136	0.0173	✗	✗	✗	0
286	ST99SA	Surface Area of RightMiddleTemporal	0.5725	0.0052	✗	✓	✗	1
287	ST85SA	Surface Area of RightFusiform	0.5894	0.0098	✗	✓	✗	1
288	ST107SA	Surface Area of RightPericalcarine	0.527	0.2189	✓	✓	✗	2
289	ST110SA	Surface Area of RightPrecentral	0.5316	0.2282	✓	✗	✓	2
290	ST105SA	Surface Area of RightParsOrbitalis	0.5011	0.3465	✓	✗	✓	2
291	ST59SA	Surface Area of LeftSupramarginal	0.5404	0.0045	✗	✓	✗	1
292	ST52SA	Surface Area of LeftPrecuneus	0.5477	0.0128	✗	✓	✗	1
293	ST102CV	VCP of RightParacentral	0.5371	0.3523	✓	✗	✓	2
294	ST82CV	VCP of RightCuneus	0.518	0.0188	✗	✗	✗	0
295	ST54SA	Surface Area of LeftRostralAnteriorCingulate	0.5081	0.0028	✗	✓	✗	1
296	ST94SA	Surface Area of RightLateralOccipital	0.5127	0.0011	✗	✓	✗	1
297	ST98SA	Surface Area of RightMedialOrbitofrontal	0.5098	0.0010	✗	✗	✗	0
298	ST114SA	Surface Area of RightRostralMiddleFrontal	0.5206	0.1251	✓	✗	✗	1
299	ST76SV	VWM of RightCerebellumCortex	0.543	0.0004	✗	✗	✗	0
300	ST50SA	Surface Area of LeftPosteriorCingulate	0.5143	0.0406	✓	✗	✗	1
301	ST38SA	Surface Area of LeftLingual	0.5137	0.0135	✗	✗	✗	0
302	ST57CV	VCP of LeftSuperiorParietal	0.6124	0.0193	✗	✗	✗	0
303	ST35SA	Surface Area of LeftLateralOccipital	0.5109	0.0822	✓	✗	✗	1
304	ST129SA	Surface Area of LeftInsula	0.5062	0.1134	✓	✗	✗	1
305	ST58SA	Surface Area of LeftSuperiorTemporal	0.5312	0.0175	✗	✗	✗	0
306	ST35CV	VCP of LeftLateralOccipital	0.5843	0.1187	✓	✓	✗	2
307	ST109SA	Surface Area of RightPosteriorCingulate	0.5121	0.2183	✓	✗	✓	2
308	ST36SA	Surface Area of LeftLateralOrbitofrontal	0.5149	0.0192	✗	✓	✗	1
309	ST17SV	VWM of LeftCerebellumCortex	0.5296	0.0230	✓	✗	✗	1
310	ST55SA	Surface Area of LeftRostralMiddleFrontal	0.5158	0.0044	✗	✓	✗	1
311	ST95SA	Surface Area of RightLateralOrbitofrontal	0.5077	0.0007	✗	✓	✗	1
312	ST115SA	Surface Area of RightSuperiorFrontal	0.5161	0.0547	✓	✗	✗	1
313	ST130SA	Surface Area of RightInsula	0.5121	0.2980	✓	✗	✓	2
314	ST10CV	VCP of Icv	0.5298	0.4283	✓	✓	✓	3
315	ST56SA	Surface Area of LeftSuperiorFrontal	0.5203	0.0002	✗	✓	✗	1
316	EICV	Estimated Intracranial Volume	0.5264	0.0001	✗	✗	✗	0

continúa en la siguiente página

Tabla 3 – continuación de página anterior

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
317	ST86SA	Surface Area of RightHemisphere	0.5136	0.0026	✗	✗	✗	0
318	ST27SA	Surface Area of LeftHemisphere	0.5119	0.1310	✓	✓	✗	2

.2 Resultados de los métodos de selección de variables sobre las 6 variables clínicas originales. Calificación final

Tabla 4: Calificación final de las 6 variables clínicas originales.

mRMR	Código	Nombre	AUC	valor Fscore	Fscore	Wrapper	Fscore+Wrapper	Nº selecc.
1	CDGLOBAL	Global CDR	0.9777	0.1254	✓	✓	✓	3
2	PTAGE	Edad del paciente	0.5236	0.0022	✗	✓	✗	1
3	PTEDUCAT	Años de educacion	0.5811	0.0205	✓	✓	✓	3
4	PTGENDER	Genero	0.5232	1.4259	✓	✓	✓	3
5	PTHAND	Handedness	0.5081	0.0052	✗	✓	✗	1
6	MMSCORE	MMSE Total Score	0.8681	0.0009	✗	✓	✗	1

Referencias

- Oscar Acosta, Jurgen Fripp, Vincent Doré, Pierrick Bourgeat, Jean-Marie Favreau, Gaël Chételat, Andrea Rueda, Victor L Villemagne, Cassandra Szoeker, David Ames, Kathryn A Ellis, Ralph N Martins, Colin L Masters, Christopher C Rowe, Erik Bonner, Florence Gris, Di Xiao, Parnesh Raniga, Vincent Barra, and Olivier Salvado. Cortical surface mapping using topology correction, partial flattening and 3d shape context-based non-rigid registration for use in quantifying atrophy in alzheimer's disease. *J Neurosci Methods*, 205(1):96–109, Mar 2012. doi: 10.1016/j.jneumeth.2011.12.011. URL <http://dx.doi.org/10.1016/j.jneumeth.2011.12.011>. 37
- Marilyn S Albert, Steven T DeKosky, Dennis Dickson, Bruno Dubois, Howard H Feldman, Nick C Fox, Anthony Gamst, David M Holtzman, William J Jagust, Ronald C Petersen, Peter J Snyder, Maria C Carrillo, Bill Thies, and Creighton H Phelps. The diagnosis of mild cognitive impairment due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimers Dement*, 7(3):270–279, May 2011. doi: 10.1016/j.jalz.2011.03.008. URL <http://dx.doi.org/10.1016/j.jalz.2011.03.008>. ii, 10, 11, 13, 16, 25, 30
- Todd A Alonzo and Margaret Sullivan Pepe. Development and evaluation of classifiers. *Methods Mol Biol*, 404: 89–116, 2007. 61, 64
- Tracy M Anderson, Perminder S Sachdev, Henry Brodaty, Julian N Trollor, and Gavin Andrews. Effects of sociodemographic and health variables on mini-mental state exam scores in older australians. *Am J Geriatr Psychiatry*, 15(6):467–476, Jun 2007. doi: 10.1097/JGP.0b013e3180547053. URL <http://dx.doi.org/10.1097/JGP.0b013e3180547053>. 89
- Hidetaka Arimura, Takashi Yoshiura, Seiji Kumazawa, Kazuhiro Tanaka, Hiroshi Koga, Futoshi Mihara, Hiroshi Honda, Shuji Sakai, Fukai Toyofuku, and Yoshiharu Higashida. Computer-aided diagnostic method for classification of alzheimer's disease with atrophic image features on mr images. volume 6915, page 69151P. SPIE, 2008. doi: 10.1117/12.769753. URL <http://link.aip.org/link/?PSI/6915/69151P/1>. 73
- J. Ashburner and K. J. Friston. Voxel-based morphometry—the methods. *Neuroimage*, 11(6 Pt 1):805–821, Jun 2000. doi: 10.1006/nimg.2000.0582. URL <http://dx.doi.org/10.1006/nimg.2000.0582>. 34
- Alzheimer's Association. 2012 alzheimer's disease facts and figures. *Alzheimers Dement*, 8(2):131–168, Mar 2012. xi, 3, 4, 11, 12
- Mark Beale, Martin Hagan, and Howard Demuth. *Neural Network Toolbox™ User's Guide R2011b documentation, Neural Network Toolbox*. The MathWorks, Inc, optional, optional edition, optional 2011. URL <http://www.mathworks.com>. optional. 60
- A. L. Bergem, K. Engedal, and E. Kringlen. The role of heredity in late-onset alzheimer disease and vascular dementia. a twin study. *Arch Gen Psychiatry*, 54(3):264–270, Mar 1997. 25

REFERENCIAS

- Lars Bertram and Rudolph E Tanzi. Thirty years of alzheimer's disease genetics: the implications of systematic meta-analyses. *Nat Rev Neurosci*, 9(10):768–778, Oct 2008. doi: 10.1038/nrn2494. URL <http://dx.doi.org/10.1038/nrn2494>. 25
- Christopher M. Bishop. *Pattern Recognition and Machine Learning*. 2006. 53, 55, 58
- Courtney A Bishop, Mark Jenkinson, Jesper Andersson, Jerome Declerck, and Dorit Merhof. Novel fast marching for automated segmentation of the hippocampus (fmash): method and validation on clinical data. *Neuroimage*, 55(3):1009–1019, Apr 2011. doi: 10.1016/j.neuroimage.2010.12.071. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.12.071>. 37
- Avrim L. Blum and Pat Langley. Selection of relevant features and examples in machine learning. *ARTIFICIAL INTELLIGENCE*, 97:245–271, 1997. 76
- Ron Brookmeyer, Elizabeth Johnson, Kathryn Ziegler-Graham, and H. Michael Arrighi. Forecasting the global burden of alzheimer's disease. *Alzheimers Dement*, 3(3):186–191, Jul 2007. doi: 10.1016/j.jalz.2007.04.381. URL <http://dx.doi.org/10.1016/j.jalz.2007.04.381>. 3
- RL Buckner. A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *Neuroimage*, 23:724–738, 2004. 35, 85
- Christopher J. C. Burges. A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2:121–167, 1998. 57
- Anna Caroli, Annapaola Prestia, Kewei Chen, Napatkamon Ayutyanont, Susan M Landau, Cindee M Madison, Cathleen Haense, Karl Herholz, Flavio Nobili, Eric M Reiman, William J Jagust, Giovanni B Frisoni, , N. E. S. T-D. D. the EADC-PET Consortium, and Alzheimer's Disease Neuroimaging Initiative. Summary metrics to assess alzheimer disease-related hypometabolic pattern with 18f-fdg pet: Head-to-head comparison. *J Nucl Med*, 53(4):592–600, Apr 2012. 126
- Chih-Chung Chang and Chih-Jen Lin. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. 94, 109, 111
- Rong Chen and Edward H Herskovits. Machine-learning techniques for building a diagnostic model for very mild dementia. *Neuroimage*, 52(1):234–244, Aug 2010. doi: 10.1016/j.neuroimage.2010.03.084. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.03.084>. 52, 77
- Yi-Wei Chen and Chih-Jen Lin. Combining svms with various feature selection strategies. In Isabelle Guyon, Masoud Nikravesh, Steve Gunn, and Lotfi Zadeh, editors, *Feature Extraction*, volume 207 of *Studies in Fuzziness and Soft Computing*, pages 315–324. Springer Berlin / Heidelberg, 2006. ISBN 978-3-540-35487-1. URL http://dx.doi.org/10.1007/978-3-540-35488-8_13. 70, 71, 72, 94
- G. Chételat, B. Landeau, F. Eustache, F. Mézenge, F. Viader, V. de la Sayette, B. Desgranges, and J.-C. Baron. Using voxel-based morphometry to map the structural changes associated with rapid conversion in mci: A longitudinal mri study. *NeuroImage*, 27(4):934–946, 2005. ISSN 1053-8119. URL <http://www.sciencedirect.com/science/article/pii/S1053811905003277>. 34, 35, 37
- Hsu Chih-Wei, Chang Chih-Chung, and Lin Chih-Jen. *A Practical Guide to Support Vector Classification*. Department of Computer Science National Taiwan University, Taipei 106, Taiwan, 2003. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>. 57, 109

REFERENCIAS

- Youngsang Cho, Joon-Kyung Seong, Yong Jeong, Sung Yong Shin, and Alzheimer's Disease Neuroimaging Initiative. 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, Feb 2012. doi: 10.1016/j.neuroimage.2011.09.085. URL <http://dx.doi.org/10.1016/j.neuroimage.2011.09.085>. 37, 83, 87
- Marie Chupin, Emilie Gérardin, Rémi Cuingnet, Claire Boutet, Louis Lemieux, Stéphane Lehéricy, Habib Benali, Line Garnero, Olivier Colliot, and Alzheimer's Disease Neuroimaging Initiative. Fully automatic hippocampus segmentation and classification in alzheimer's disease and mild cognitive impairment applied on data from adni. *Hippocampus*, 19(6):579–587, Jun 2009. doi: 10.1002/hipo.20626. URL <http://dx.doi.org/10.1002/hipo.20626>. 34, 37, 45, 47, 87
- C. M. Clark, C. Davatzikos, A. Borthakur, A. Newberg, S. Leight, V. M-Y Lee, and J. Q. Trojanowski. Biomarkers for early detection of alzheimer pathology. *Neurosignals*, 16(1):11–18, 2008. doi: 10.1159/000109754. URL <http://dx.doi.org/10.1159/000109754>. 30, 32
- NR Cook. Use and misuse of the receiver operating characteristic curve in risk prediction. *Circulation*, 115(2): 928–935, 2007. doi: 10.1161/CIRCULATIONAHA.106.672402. 91, 122
- Yue Cui, Bing Liu, Suhuai Luo, Xiantong Zhen, Ming Fan, Tao Liu, Wanlin Zhu, Mira Park, Tianzi Jiang, and Jesse Jin. Identification of conversion from mild cognitive impairment to alzheimer's disease using multivariate predictors. *PLoS One*, 6(7):e21896, 2011. 3, 48, 70, 71, 73, 87, 89, 91, 119, 121, 122, 124
- Remi Cuingnet, Emilie Gerardin, Jérôme Tessieras, Guillaume Auzias, Stéphane Lehéricy, Marie-Odile Habert, Marie Chupin, Habib Benali, and Olivier Colliot. Automatic classification of patients with alzheimer's disease from structural mri: A comparison of ten methods using the adni database. *Multivariate Decoding and Brain Reading*, 56(2):766–781, 2011. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2010.06.013. URL <http://www.sciencedirect.com/science/article/pii/S1053811910008578>. 37, 43, 44, 48, 49, 71, 73, 83, 87, 89, 120, 121
- Jeffrey Cummings. Alzheimer's disease: clinical trials and the amyloid hypothesis. *Ann Acad Med Singapore*, 40(7):304–303, Jul 2011. 5
- Jeffrey L Cummings. Alzheimer's disease. *N Engl J Med*, 351(1):56–67, Jul 2004. doi: 10.1056/NEJMra040223. URL <http://dx.doi.org/10.1056/NEJMra040223>. xi, 5, 6, 7
- Anders Dale, Bruce Fischl, and Martin I. Sereno. Cortical surface-based analysis: I. segmentation and surface reconstruction. *NeuroImage*, 9(2):179–194, 1999. 37, 112
- Sandhitsu R Das, Brian B Avants, John Pluta, Hongzhi Wang, Jung W Suh, Michael W Weiner, Susanne G Mueller, and Paul A Yushkevich. Measuring longitudinal change in the hippocampal formation from in vivo high-resolution t2-weighted mri. *Neuroimage*, 60(2):1266–1279, Apr 2012. doi: 10.1016/j.neuroimage.2012.01.098. URL <http://dx.doi.org/10.1016/j.neuroimage.2012.01.098>. 45, 47, 48
- C. Davatzikos, K. Ruparel, Y. Fan, D. G. Shen, M. Acharyya, J. W. Loughead, R. C. Gur, and D. D. Langleben. Classifying spatial patterns of brain activity with machine learning methods: application to lie detection. *Neuroimage*, 28(3):663–668, Nov 2005. doi: 10.1016/j.neuroimage.2005.08.009. URL <http://dx.doi.org/10.1016/j.neuroimage.2005.08.009>. 72
- Christos Davatzikos, Yong Fan, Xiaoying Wu, Dinggang Shen, and Susan M Resnick. Detection of prodromal alzheimer's disease via pattern classification of magnetic resonance imaging. *Neurobiol Aging*, 29(4):514–523, Apr 2008. doi: 10.1016/j.neurobiolaging.2006.11.010. URL <http://dx.doi.org/10.1016/j.neurobiolaging.2006.11.010>. 52, 71, 72, 73

REFERENCIAS

- Sergio Davinelli, Mariano Intriери, Claudio Russo, Alfonso Di Costanzo, Davide Zella, Paolo Bosco, and Giovanni Scapagnini. The "alzheimer's disease signature": potential perspectives for novel biomarkers. *Immun Ageing*, 8:7, 2011. doi: 10.1186/1742-4933-8-7. URL <http://dx.doi.org/10.1186/1742-4933-8-7>. 25
- Steven T DeKosky, Maria C Carrillo, Creighton Phelps, David Knopman, Ronald C Petersen, Richard Frank, Dale Schenk, Donna Masterman, Eric R Siemers, Jesse M Cedarbaum, Michael Gold, David S Miller, Bruce H Morimoto, Ara S Khachaturian, and Richard C Mohs. Revision of the criteria for alzheimer's disease: A symposium. *Alzheimers Dement*, 7(1):e1–12, Jan 2011. xv, 9, 10
- Rahul Desikan, Florent Sigonne, Bruce Fischl, Brian T. Quinn, Bradford C. Dickerson, Deborah Blacker, Randy L. Buckner, Anders M. Dale, R. Paul Maguire, Bradley T. Hyman, Marilyn S. Albert, and Ronald J. Killiany. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980, 2006. ISSN 1053-8119. URL <http://www.sciencedirect.com/science/article/pii/S1053811906000437>. xii, 35, 36, 37, 45, 46, 85, 88
- Rahul Desikan, Howard Cabral, Christopher Hess, William Dillon, Christine Glastonbury, Michael Weiner, Nicholas Schmansky, Douglas Greve, David Salat, Randy Buckner, and Bruce Fischl. Automated mri measures identify individuals with mild cognitive impairment and alzheimer's disease. *Brain*, 132(8):2048–2057, 2009. xii, 3, 29, 31, 32, 37, 45, 46, 87, 119, 124
- Rahul S Desikan, Howard J Cabral, Fabio Settecase, Christopher P Hess, William P Dillon, Christine M Glastonbury, Michael W Weiner, Nicholas J Schmansky, David H Salat, Bruce Fischl, and Alzheimer's Disease Neuroimaging Initiative. Automated mri measures predict progression to alzheimer's disease. *Neurobiol Aging*, 31(8):1364–1374, Aug 2010. 45
- Christophe Destrieux, Bruce Fischl, Anders Dale, and Eric Halgren. Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature. *Neuroimage*, 53(1):1–15, Oct 2010. doi: 10.1016/j.neuroimage.2010.06.010. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.06.010>. 35, 37, 88
- Bradford C Dickerson, Akram Bakkour, David H Salat, Eric Feczko, Jenni Pacheco, Douglas N Greve, Fran Grodstein, Christopher I Wright, Deborah Blacker, H. Diana Rosas, Reisa A Sperling, Alireza Atri, John H Growdon, Bradley T Hyman, John C Morris, Bruce Fischl, and Randy L Buckner. The cortical signature of alzheimer's disease: regionally specific cortical thinning relates to symptom severity in very mild to mild ad dementia and is detectable in asymptomatic amyloid-positive individuals. *Cereb Cortex*, 19(3):497–510, Mar 2009. doi: 10.1093/cercor/bhn113. URL <http://dx.doi.org/10.1093/cercor/bhn113>. 35, 37, 44
- Chris Ding and Hanchuan Peng. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology*, 3(2):185–2005, apr. 2005. 69, 70, 92
- GM Doniger, DM Zucker, A Schweiger, T Dwolatzky, H Chertkow, H Crystal, and ES Simon. Towards practical cognitive assessment for detection of early dementia: a 30-minute computerized battery discriminates as well as longer testing. *Current Alzheimer Research*, 2(2):117–124, 2005. 91, 122
- Bruno Dubois, Howard H Feldman, Claudia Jacova, Steven T Dekosky, Pascale Barberger-Gateau, Jeffrey Cummings, André Delacourte, Douglas Galasko, Serge Gauthier, Gregory Jicha, Kenichi Meguro, John O'brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Yaakov Stern, Pieter J Visser, and Philip Scheltens. Research criteria for the diagnosis of alzheimer's disease: revising the nincls-adrda criteria. *Lancet Neurol*, 6(8):734–746, Aug 2007. doi: 10.1016/S1474-4422(07)70178-3. URL [http://dx.doi.org/10.1016/S1474-4422\(07\)70178-3](http://dx.doi.org/10.1016/S1474-4422(07)70178-3). 10, 16, 25, 45

REFERENCIAS

- Bruno Dubois, Howard H Feldman, Claudia Jacova, Jeffrey L Cummings, Steven T Dekosky, Pascale Barberger-Gateau, André Delacourte, Giovanni Frisoni, Nick C Fox, Douglas Galasko, Serge Gauthier, Harald Hampel, Gregory A Jicha, Kenichi Meguro, John O'Brien, Florence Pasquier, Philippe Robert, Martin Rossor, Steven Salloway, Marie Sarazin, Leonardo C de Souza, Yaakov Stern, Pieter J Visser, and Philip Scheltens. Revising the definition of alzheimer's disease: a new lexicon. *Lancet Neurol*, 9(11):1118–1127, Nov 2010. doi: 10.1016/S1474-4422(10)70223-4. URL [http://dx.doi.org/10.1016/S1474-4422\(10\)70223-4](http://dx.doi.org/10.1016/S1474-4422(10)70223-4). 10
- Juergen Dukart, Karsten Mueller, Annette Horstmann, Henryk Barthel, Harald E Möller, Arno Villringer, Osama Sabri, and Matthias L Schroeter. Combined evaluation of fdg-pet and mri improves detection and differentiation of dementia. *PLoS One*, 6(3):e18111, 2011. doi: 10.1371/journal.pone.0018111. URL <http://dx.doi.org/10.1371/journal.pone.0018111>. 73, 126
- Catherine Edwards and Bhavani Raskutti. The effect of attribute scaling on the performance of support vector machines. In Geoffrey Webb and Xinghuo Yu, editors, *AI 2004: Advances in Artificial Intelligence*, volume 3339 of *Lecture Notes in Computer Science*, pages 211–238. Springer Berlin / Heidelberg, 2005. ISBN 978-3-540-24059-4. URL http://dx.doi.org/10.1007/978-3-540-30549-1_44. 60
- Michael Ewers, Reisa A Sperling, William E Klunk, Michael W Weiner, and Harald Hampel. Neuroimaging markers for the prediction and early diagnosis of alzheimer's disease dementia. *Trends Neurosci*, 34(8):430–442, Aug 2011. doi: 10.1016/j.tins.2011.05.005. URL <http://dx.doi.org/10.1016/j.tins.2011.05.005>. xi, xii, 7, 8, 29, 30, 31, 33, 35, 44
- D Faraggi and B Reiser. Estimation of the area under the roc curve. *Statistics in Medicine*, 21:3093–3106, 2002. 91, 122
- B. Fischl, A. Liu, and A. M. Dale. Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex. *IEEE Medical Imaging*, 20(1):70–80, Jan 2001. 112
- Bruce Fischl. Freesurfer. *Neuroimage*, Jan 2012. doi: 10.1016/j.neuroimage.2012.01.021. URL <http://dx.doi.org/10.1016/j.neuroimage.2012.01.021>. 37, 88, 126
- Bruce Fischl and Anders M. Dale. Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America*, 97(20):11050–11055, 2000. 37, 112
- Bruce Fischl, Martin I. Sereno, and Anders Dale. Cortical surface-based analysis: Ii: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, 9(2):195–207, 1999a. 37, 112
- Bruce Fischl, Martin I. Sereno, Roger B.H. Tootell, and Anders M. Dale. High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, 8(4):272–284, 1999b. ISSN 1097-0193. doi: 10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4. URL [http://dx.doi.org/10.1002/\(SICI\)1097-0193\(1999\)8:4<272::AID-HBM10>3.0.CO;2-4](http://dx.doi.org/10.1002/(SICI)1097-0193(1999)8:4<272::AID-HBM10>3.0.CO;2-4). 112
- Bruce Fischl, David H. Salat, Evelina Busa, Marilyn Albert, Megan Dieterich, Christian Haselgrove, Andre Kouwe, Ron Killiany, David Kennedy, Shuna Klaveness, Albert Montillo, Nikos Makris, Bruce Rosen, and Anders M. Dale. Whole brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron*, 33(3):341–355, 2002. ISSN 0896-6273. 35, 36, 37, 42, 88, 89, 112
- Bruce Fischl, David H. Salat, André J.W. van der Kouwe, Nikos Makris, Florent Ségonne, Brian T. Quinn, and Anders M. Dale. Sequence-independent segmentation of magnetic resonance images. *NeuroImage*, 23(Supplement

REFERENCIAS

- 1):S69–S84, 2004a. ISSN 1053-8119. doi: DOI:10.1016/j.neuroimage.2004.07.016. URL <http://www.sciencedirect.com/science/article/B6WNP-4DCMGVT-2/2/7eee26326dc63f931b826eac33becc8b>. Mathematics in Brain Imaging. 88, 112
- Bruce Fischl, André van der Kouwe, Christophe Destrieux, Eric Halgren, Florent Ségonne, David H. Salat, Evelina Busa, Larry J. Seidman, Jill Goldstein, David Kennedy, Verne Caviness, Nikos Makris, Bruce Rosen, and Anders M. Dale. Automatically Parcellating the Human Cerebral Cortex. *Cerebral Cortex*, 14(1):11–22, 2004b. doi: 10.1093/cercor/bhg087. URL <http://cercor.oxfordjournals.org/content/14/1/11.abstract>. 37, 112
- Anders M Fjell and Kristine B Walhovd. New tools for the study of alzheimer’s disease: what are biomarkers and morphometric markers teaching us? *Neuroscientist*, 17(5):592–605, Oct 2011. doi: 10.1177/1073858410392586. URL <http://dx.doi.org/10.1177/1073858410392586>. 24, 25, 30, 32
- MF Folstein, SE Folstein, and PR McHugh. Mini-mental state. a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–198, 1975. 22, 82, 89
- Giovanni B Frisoni. Alzheimer’s disease neuroimaging initiative in europe. *Alzheimers Dement*, 6(3):280–285, May 2010. doi: 10.1016/j.jalz.2010.03.005. URL <http://dx.doi.org/10.1016/j.jalz.2010.03.005>. 80, 81
- Giovanni B Frisoni and Michael W Weiner. Alzheimer’s disease neuroimaging initiative special issue. *Neurobiol Aging*, 31(8):1259–1262, Aug 2010. doi: 10.1016/j.neurobiolaging.2010.05.006. URL <http://dx.doi.org/10.1016/j.neurobiolaging.2010.05.006>. 80
- Giovanni B Frisoni, Wouter J P Henneman, Michael W Weiner, Philip Scheltens, Bruno Vellas, Emma Reynish, Jaroslava Hudecova, Harald Hampel, Katharina Burger, Kaj Blennow, Gunhild Waldemar, Peter Johannsen, Lars-Olof Wahlund, Giancarlo Zito, Paolo M Rossini, Bengt Winblad, Frederik Barkhof, and Alzheimer’s Disease Neuroimaging Initiative. The pilot european alzheimer’s disease neuroimaging initiative of the european alzheimer’s disease consortium. *Alzheimers Dement*, 4(4):255–264, Jul 2008. 80
- I Guyon and A Elisseeff. An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3:1157–1182, 2003. 64, 68, 69, 91
- H. Hampel, A. Mitchell, K. Blennow, R. A. Frank, S. Brettschneider, L. Weller, and H-J. Möller. Core biological marker candidates of alzheimer’s disease - perspectives for diagnosis, prediction of outcome and reflection of biological activity. *J Neural Transm*, 111(3):247–272, Mar 2004. doi: 10.1007/s00702-003-0065-z. URL <http://dx.doi.org/10.1007/s00702-003-0065-z>. 20, 25
- Harald Hampel, Katharina Bürger, Stefan J Teipel, Arun L W Bokde, Henrik Zetterberg, and Kaj Blennow. Core candidate neurochemical and imaging biomarkers of alzheimer’s disease. *Alzheimers Dement*, 4(1):38–48, Jan 2008. doi: 10.1016/j.jalz.2007.08.006. URL <http://dx.doi.org/10.1016/j.jalz.2007.08.006>. 25, 30
- Harald Hampel, Richard Frank, Karl Broich, Stefan J Teipel, Russell G Katz, John Hardy, Karl Herholz, Arun L W Bokde, Frank Jessen, Yvonne C Hoessler, Wendy R Sanhai, Henrik Zetterberg, Janet Woodcock, and Kaj Blennow. Biomarkers for alzheimer’s disease: academic, industry and regulatory perspectives. *Nat Rev Drug Discov*, 9(7):560–574, Jul 2010a. doi: 10.1038/nrd3115. URL <http://dx.doi.org/10.1038/nrd3115>. 30, 31
- Harald Hampel, Yong Shen, Dominic M Walsh, Paul Aisen, Les M Shaw, Henrik Zetterberg, John Q Trojanowski, and Kaj Blennow. Biological markers of amyloid beta-related mechanisms in alzheimer’s disease. *Exp Neurol*, 223(2):334–346, Jun 2010b. doi: 10.1016/j.expneurol.2009.09.024. URL <http://dx.doi.org/10.1016/j.expneurol.2009.09.024>. 25

REFERENCIAS

- Harald Hampel, Gordon Wilcock, Sandrine Andrieu, Paul Aisen, Kaj Blennow, K. Broich, Maria Carrillo, Nick C Fox, Giovanni B Frisoni, Maria Isaac, Simon Lovestone, Agneta Nordberg, David Prvulovic, Christina Sampaio, Philip Scheltens, Michael Weiner, Bengt Winblad, Nicola Coley, Bruno Vellas, and Oxford Task Force Group. Biomarkers for alzheimer's disease therapeutic trials. *Prog Neurobiol*, 95(4):579–593, Dec 2011. doi: 10.1016/j.pneurobio.2010.11.005. URL <http://dx.doi.org/10.1016/j.pneurobio.2010.11.005>. 21, 24, 29, 31
- Xiao Han, Jorge Jovicich, David Salat, Andre van der Kouwe, Brian Quinn, Silvester Czanner, Evelina Busa, Jenni Pacheco, Marilyn Albert, Ronald Killiany, Paul Maguire, Diana Rosas, Nikos Makris, Anders Dale, Bradford Dickerson, and Bruce Fischl. Reliability of MRI-derived measurements of human cerebral cortical thickness: The effects of field strength, scanner upgrade and manufacturer. *NeuroImage*, 32(1):180–194, 2006. 37, 112
- John Hardy and Dennis J Selkoe. The amyloid hypothesis of alzheimer's disease: progress and problems on the road to therapeutics. *Science*, 297(5580):353–356, Jul 2002. doi: 10.1126/science.1072994. URL <http://dx.doi.org/10.1126/science.1072994>. xi, 5, 6
- Rolf A. Heckemann, Shiva Keihaninejad, Paul Aljabar, Katherine R. Gray, Casper Nielsen, Daniel Rueckert, Joseph V. Hajnal, and Alexander Hammers. Automatic morphometry in alzheimer's disease and mild cognitive impairment. *NeuroImage*, 56(4):2024–2037, 2011. ISSN 1053-8119. URL <http://www.sciencedirect.com/science/article/pii/S1053811911002734>. 3, 34, 35, 37, 87, 119, 124
- Karl Herholz, Sarah Westwood, Cathleen Haense, and Graham Dunn. Evaluation of a calibrated (18)f-fdg pet score as a biomarker for progression in alzheimer disease and mild cognitive impairment. *J Nucl Med*, 52(8):1218–1226, Aug 2011. doi: 10.2967/jnumed.111.090902. URL <http://dx.doi.org/10.2967/jnumed.111.090902>. 126
- Lap Ho, Hayley Fivecoat, Jun Wang, and Giulio Maria Pasinetti. Alzheimer's disease biomarker discovery in symptomatic and asymptomatic patients: experimental approaches and future clinical applications. *Exp Gerontol*, 45(1):15–22, Jan 2010. doi: 10.1016/j.exger.2009.09.007. URL <http://dx.doi.org/10.1016/j.exger.2009.09.007>. 23
- Xue Hua, Alex D Leow, Suh Lee, Andrea D Klunder, Arthur W Toga, Natasha Lepore, Yi-Yu Chou, Caroline Brun, Ming-Chang Chiang, Marina Barysheva, Clifford R Jack, Matt A Bernstein, Paula J Britson, Chadwick P Ward, Jennifer L Whitwell, Bret Borowski, Adam S Fleisher, Nick C Fox, Richard G Boyes, Josephine Barnes, Daniella Harvey, John Kornak, Norbert Schuff, Lauren Boreta, Gene E Alexander, Michael W Weiner, Paul M Thompson, and Alzheimer's Disease Neuroimaging Initiative. 3d characterization of brain atrophy in alzheimer's disease and mild cognitive impairment using tensor-based morphometry. *Neuroimage*, 41(1):19–34, May 2008. doi: 10.1016/j.neuroimage.2008.02.010. URL <http://dx.doi.org/10.1016/j.neuroimage.2008.02.010>. 35, 44, 87
- Michael C Irizarry. Biomarkers of alzheimer disease in plasma. *NeuroRx*, 1(2):226–234, Apr 2004. doi: 10.1602/neurorx.1.2.226. URL <http://dx.doi.org/10.1602/neurorx.1.2.226>. 25
- Clifford R Jack, David S Knopman, William J Jagust, Leslie M Shaw, Paul S Aisen, Michael W Weiner, Ronald C Petersen, and John Q Trojanowski. Hypothetical model of dynamic biomarkers of the alzheimer's pathological cascade. *Lancet Neurol*, 9(1):119–128, Jan 2010a. doi: 10.1016/S1474-4422(09)70299-6. URL [http://dx.doi.org/10.1016/S1474-4422\(09\)70299-6](http://dx.doi.org/10.1016/S1474-4422(09)70299-6). xi, 18, 19
- Clifford R Jack, Heather J Wiste, Prashanthi Vemuri, Stephen D Weigand, Matthew L Senjem, Guang Zeng, Matt A Bernstein, Jeffrey L Gunter, Vernon S Pankratz, Paul S Aisen, Michael W Weiner, Ronald C Petersen, Leslie M Shaw, John Q Trojanowski, David S Knopman, and Alzheimer's Disease Neuroimaging Initiative.

REFERENCIAS

- Brain beta-amyloid measures and magnetic resonance imaging atrophy both predict time-to-progression from mild cognitive impairment to alzheimer’s disease. *Brain*, 133(11):3336–3348, Nov 2010b. 126
- CR Jack, MA Bernstein, and NC Fox. The alzheimer’s disease neuroimaging initiative (adni): Mri methods. *J Magn Reson Imaging*, 27(4):685–691, 2008. 80
- Mark Jenkinson, Christian F Beckmann, Timothy E J Behrens, Mark W Woolrich, and Stephen M Smith. Fsl. *Neuroimage*, 62(2):782–790, Aug 2012. doi: 10.1016/j.neuroimage.2011.09.015. URL <http://dx.doi.org/10.1016/j.neuroimage.2011.09.015>. 36
- Price JL and Morris JC. Tangles and plaques in nondemented aging and “preclinical” alzheimer’s disease. *Annals of Neurology*, 45(3):358–368, 1999. 20
- Jorge Jovicich, Silvester Czanner, Douglas Greve, Elizabeth Haley, Andre van der Kouwe, Randy Gollub, David Kennedy, Franz Schmitt, Gregory Brown, James MacFall, Bruce Fischl, and Anders Dale. Reliability in multi-site structural mri studies: Effects of gradient non-linearity correction on phantom and human data. *NeuroImage*, 30(2):436–443, 2006. ISSN 1053-8119. doi: DOI:10.1016/j.neuroimage.2005.09.046. URL <http://www.sciencedirect.com/science/article/B6WNP-4HM7S0B-2/2/4fa5ff26cad90ba3c9ed12b7e12ce3b6>. 37, 112
- Linus Jönsson, Maria Eriksdotter Jönhagen, Lena Kilander, Hilka Soininen, Merja Hallikainen, Gunhild Walde-
mar, Harald Nygaard, Niels Andreasen, Bengt Winblad, and Anders Wimo. Determinants of costs of care for patients with alzheimer’s disease. *Int J Geriatr Psychiatry*, 21(5):449–459, May 2006. doi: 10.1002/gps.1489. URL <http://dx.doi.org/10.1002/gps.1489>. 3
- Kai kai Shen, Jurgen Fripp, Fabrice Mériaudeau, Gaël Chételat, Olivier Salvado, Pierrick Bourgeat, and Alzheimer’s Disease Neuroimaging Initiative. Detecting global and local hippocampal shape changes in alzheimer’s disease using statistical shape models. *Neuroimage*, 59(3):2155–2166, Feb 2012. doi: 10.1016/j.neuroimage.2011.10.014. URL <http://dx.doi.org/10.1016/j.neuroimage.2011.10.014>. 52, 71, 72, 73
- G.B. Karas, P. Scheltens, S.A.R.B. Rombouts, P.J. Visser, R.A. van Schijndel, N.C. Fox, and F. Barkhof. Global and local gray matter loss in mild cognitive impairment and alzheimer’s disease. *NeuroImage*, 23(2):708–716, 2004. 32, 34, 35
- Vojislav Kecman. Support vector machines basics. Technical Report 616, The University of Auckland, School of Engineering, April 2004. 57
- Vojislav Kecman. *Support Vector Machines: Theory and Applications*, volume 177, chapter Support Vector Machines – An Introduction, pages 1–47. Springer-Verlag, 2005. 57
- S. Sathya Keerthi and Chih-Jen Lin. Asymptotic behaviors of support vector machines with gaussian kernel. *Neural Computation*, 15(7):1667–1689, 2003. 109
- R. J. Killiany, T. Gomez-Isla, M. Moss, R. Kikinis, T. Sandor, F. Jolesz, R. Tanzi, K. Jones, B. T. Hyman, and M. S. Albert. Use of structural magnetic resonance imaging to predict who will get alzheimer’s disease. *Ann Neurol*, 47(4):430–439, Apr 2000. 29, 32, 35
- Stefan Klöppel, Cynthia M Stonnington, Carlton Chu, Bogdan Draganski, Rachael I Scahill, Jonathan D Rohrer, Nick C Fox, Clifford R Jack, John Ashburner, and Richard S J Frackowiak. Automatic classification of mr scans in alzheimer’s disease. *Brain*, 131(Pt 3):681–689, Mar 2008. doi: 10.1093/brain/awm319. URL <http://dx.doi.org/10.1093/brain/awm319>. 35, 43, 44, 52, 71, 72, 73

REFERENCIAS

- D.S. Knopman, S.T. DeKosky, J.L. Cummings, H. Chui, J. Corey-Bloom, N. Relkin, G.W. Small, B. Miller, and J.C. 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. 1, 76
- Ron Kohavi and Foster Provost, editors. *Machine Learning*, volume 30, chapter Glossary of Terms, pages 271–274. Springer Netherlands, February 1998. URL <http://dx.doi.org/10.1023/A:1017181826899>. 10.1023/A:1017181826899. 55, 63, 64, 94, 95
- S. M. Landau, D. Harvey, C. M. Madison, E. M. Reiman, N. L. Foster, P. S. Aisen, R. C. Petersen, L. M. Shaw, J. Q. Trojanowski, C. R. Jack, M. W. Weiner, W. J. Jagust, and Alzheimer’s Disease Neuroimaging Initiative. Comparing predictors of conversion and decline in mild cognitive impairment. *Neurology*, 75(3):230–238, Jul 2010. 89
- Thomas A Lasko, Jui G Bhagwat, Kelly H Zou, and Lucila Ohno-Machado. The use of receiver operating characteristic curves in biomedical informatics. *J Biomed Inform*, 38(5):404–415, Oct 2005. doi: 10.1016/j.jbi.2005.02.008. URL <http://dx.doi.org/10.1016/j.jbi.2005.02.008>. 92
- Xingfeng Li, Arnaud Messé, Guillaume Marrelec, Mélanie Péligrini-Issac, and Habib Benali. An enhanced voxel-based morphometry method to investigate structural changes: application to alzheimer’s disease. *Neuroradiology*, 52(3):203–213, Mar 2010. doi: 10.1007/s00234-009-0600-1. URL <http://dx.doi.org/10.1007/s00234-009-0600-1>. 37
- H.-T. Lin and C.-J. Lin. A study on sigmoid kernels for svm and the training of non- psd kernels by smo-type methods. Technical report, Department of Computer Science, National Taiwan University, 2003. URL <http://www.csie.ntu.edu.tw/~cjlin/papers/tanh.pdf>. 109
- Yawu Liu, Teemu Paaajanen, Yi Zhang, Eric Westman, Lars-Olof Wahlund, Andrew Simmons, Catherine Tunnard, Tomasz Sobow, Patrizia Mecocci, Magda Tsolaki, Bruno Vellas, Sebastian Muehlboeck, Alan Evans, Christian Spenger, Simon Lovestone, and Hilikka Soininen. Combination analysis of neuropsychological tests and structural mri measures in differentiating ad, mci and control groups—the addneuromed study. *Neurobiology of Aging*, 32(7):1198–1206, 2011. 32
- Raymond Y Lo, Alan E Hubbard, Leslie M Shaw, John Q Trojanowski, Ronald C Petersen, Paul S Aisen, Michael W Weiner, William J Jagust, and Alzheimer’s Disease Neuroimaging Initiative. Longitudinal change of biomarkers in cognitive decline. *Arch Neurol*, 68(10):1257–1266, Oct 2011. doi: 10.1001/archneurol.2011.123. URL <http://dx.doi.org/10.1001/archneurol.2011.123>. 126
- Daniel S Marcus, Tracy H Wang, Jamie Parker, John G Csernansky, John C Morris, and Randy L Buckner. Open access series of imaging studies (oasis): cross-sectional mri data in young, middle aged, nondemented, and demented older adults. *Journal of Cognitive Neuroscience*, 19(9):1498–1507, 2007. URL http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17714011. 107, 111
- The MathWorks. *MATLAB. The Language of Thechnical Computing. v. 7.10.0499 (R2010a)*. The MathWorks, February 2010. 109, 111, 126
- N. Mattsson, E. Rosén, O. Hansson, N. Andreasen, L. Parnetti, M. Jonsson, S-K. Herukka, W. M. van der Flier, M. A. Blankenstein, M. Ewers, K. Rich, E. Kaiser, M. M. Verbeek, M. Olde Rikkert, M. Tsolaki, E. Mulugeta, D. Aarsland, P. J. Visser, J. Schröder, J. Marcusson, M. de Leon, H. Hampel, P. Scheltens, A. Wallin, M. Eriksdotter-Jönhagen, L. Minthon, B. Winblad, K. Blennow, and H. Zetterberg. Age and diagnostic performance of alzheimer disease csf biomarkers. *Neurology*, 78(7):468–476, Feb 2012. doi: 10.1212/WNL.0b013e3182477eed. URL <http://dx.doi.org/10.1212/WNL.0b013e3182477eed>. 25

REFERENCIAS

- Niklas Mattsson, Ulf Andreasson, Staffan Persson, Hiroyuki Arai, Sat Dev Batish, Sergio Bernardini, Luisella Bocchio-Chiavetto, Marinus A Blankenstein, Maria C Carrillo, Sonia Chalbot, Els Coart, Davide Chiasserini, Neal Cutler, Gunilla Dahlfors, Stefan Duller, Anne M Fagan, Orestes Forlenza, Giovanni B Frisoni, Douglas Galasko, Daniela Galimberti, Harald Hampel, Aase Handberg, Michael T Heneka, Adrianna Z Herskovits, Sanna-Kaisa Herukka, David M Holtzman, Christian Humpel, Bradley T Hyman, Khalid Iqbal, Mathias Jucker, Stephan A Kaeser, Elmar Kaiser, Elisabeth Kapaki, Daniel Kidd, Peter Klivenyi, Cindy S Knudsen, Markus P Kummer, James Lui, Albert Lladó, Piotr Lewczuk, Qiao-Xin Li, Ralph Martins, Colin Masters, John McAuliffe, Marc Mercken, Abhay Moghekar, José Luis Molinuevo, Thomas J Montine, William Nowatzke, Richard O'Brien, Markus Otto, George P Paraskevas, Lucilla Parnetti, Ronald C Petersen, David Prvulovic, Herman P M de Reus, Robert A Rissman, Elio Scarpini, Alessandro Stefani, Hilkkka Soininen, Johannes Schröder, Leslie M Shaw, Anders Skinningrud, Brith Skrogstad, Annette Spreer, Leda Talib, Charlotte Teunissen, John Q Trojanowski, Hayretin Tumani, Robert M Umek, Bianca Van Broeck, Hugo Vanderstichele, Laszlo Vecsei, Marcel M Verbeek, Manfred Windisch, Jing Zhang, Henrik Zetterberg, and Kaj Blennow. The alzheimer's association external quality control program for cerebrospinal fluid biomarkers. *Alzheimers Dement*, 7(4):386–395.e6, Jul 2011. doi: 10.1016/j.jalz.2011.05.2243. URL <http://dx.doi.org/10.1016/j.jalz.2011.05.2243>. 25
- Konrad Maurer, Stephan Volk, and Hector Gerbaldo. Auguste d and alzheimer's disease. *The Lancet*, 349(9064): 1546–1549, 1997. doi: 10.1016/S0140-6736(96)10203-8. 4
- G McKhann, D Drachman, M Folstein, R Katzman, D Price, and EM Stadlan. Clinical diagnosis of alzheimer's disease: report of the nincls-adrda work group under the auspices of department of health and human services task force on alzheimer's disease. *Neurology*, 34(7):939–944, 1984. 1, 9, 15, 82
- Guy M McKhann, David S Knopman, Howard Chertkow, Bradley T Hyman, Clifford R Jack, Claudia H Kawas, William E Klunk, Walter J Koroshetz, Jennifer J Manly, Richard Mayeux, Richard C Mohs, John C Morris, Martin N Rossor, Philip Scheltens, Maria C Carrillo, Bill Thies, Sandra Weintraub, and Creighton H Phelps. The diagnosis of dementia due to alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimers Dement*, 7(3): 263–269, May 2011. doi: 10.1016/j.jalz.2011.03.005. URL <http://dx.doi.org/10.1016/j.jalz.2011.03.005>. ii, 10, 11, 15, 25, 30
- Lilia Mesrob, Benoit Magnin, Olivier Colliot, Marie Sarazin, Valérie Hahn-Barma, Bruno Dubois, Patrick Gallinari, Stéphane Lehericy, Serge Kinkingnéhun, and Habib Benali. Identification of atrophy patterns in alzheimer's disease based on svm feature selection and anatomical parcellation. In *Proceedings of the 4th international workshop on Medical Imaging and Augmented Reality*, MIAR '08, pages 124–132, Berlin, Heidelberg, 2008. Springer-Verlag. ISBN 978-3-540-79981-8. doi: 10.1007/978-3-540-79982-5_14. URL http://dx.doi.org/10.1007/978-3-540-79982-5_14. 71, 73
- Camila Moraes, Jony Pinto, Marcos Lopes, Julio Litvoc, and Cassio Bottino. Impact of sociodemographic and health variables on mini-mental state examination in a community-based sample of older people. *European Archives of Psychiatry and Clinical Neuroscience*, 260:535–542, 2010. ISSN 0940-1334. URL <http://dx.doi.org/10.1007/s00406-010-0104-3>. 10.1007/s00406-010-0104-3. 89
- Jonathan H Morra, Zhuowen Tu, Liana G Apostolova, Amity E Green, Christina Avedissian, Sarah K Madsen, Neelroop Parikhshak, Arthur W Toga, Clifford R Jack, Norbert Schuff, Michael W Weiner, Paul M Thompson, and Alzheimer's Disease Neuroimaging Initiative. Automated mapping of hippocampal atrophy in 1-year repeat mri data from 490 subjects with alzheimer's disease, mild cognitive impairment, and elderly controls. *Neuroimage*, 45(1 Suppl):S3–15, Mar 2009. 32, 37, 45

REFERENCIAS

- J. C. Morris, D. W. McKeel, M. Storandt, E. H. Rubin, J. L. Price, E. A. Grant, M. J. Ball, and L. Berg. Very mild alzheimer's disease: informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology*, 41(4):469–478, 1991. 5
- JC Morris. The clinical dementia rating (cdr): current version and scoring rules. *Neurology*, 43:2412–2414, 1993. 22, 82, 89, 112
- JC Morris, M Storandt, JP Miller, DW McKeel, JL Price, EH Rubin, and L Berg. Mild cognitive impairment represents early-stage alzheimer disease. *Arch Neurol*, 58:397–405, 2001. 2, 12, 13, 20, 23, 112
- Susanne Mueller, Michael Weiner, Leon Thal, Ronald Petersen, Jack Clifford, William Jagust, John Trojanowski, Arthur Toga, and Laurel Beckett. The alzheimer's disease neuroimaging initiative. *Neuroimaging Clinics of North America*, 15(4):869–877, 2005a. 80
- Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. The alzheimer's disease neuroimaging initiative. *Neuroimaging Clin N Am*, 15(4):869–77, xi–xii, Nov 2005b. doi: 10.1016/j.nic.2005.09.008. URL <http://dx.doi.org/10.1016/j.nic.2005.09.008>. 1
- Susanne G Mueller, Michael W Weiner, Leon J Thal, Ronald C Petersen, Clifford R Jack, William Jagust, John Q Trojanowski, Arthur W Toga, and Laurel Beckett. Ways toward an early diagnosis in alzheimer's disease: the alzheimer's disease neuroimaging initiative (adni). *Alzheimers Dement*, 1(1):55–66, Jul 2005c. doi: 10.1016/j.jalz.2005.06.003. URL <http://dx.doi.org/10.1016/j.jalz.2005.06.003>. 2, 23
- Susanne G Mueller, Norbert Schuff, Kristine Yaffe, Catherine Madison, Bruce Miller, and Michael W Weiner. Hippocampal atrophy patterns in mild cognitive impairment and alzheimer's disease. *Hum Brain Mapp*, 31(9):1339–1347, Sep 2010. doi: 10.1002/hbm.20934. URL <http://dx.doi.org/10.1002/hbm.20934>. xii, 45, 47
- Sean M Nestor, Raul Rupsingh, Michael Borrie, Matthew Smith, Vittorio Accomazzi, Jennie L Wells, Jennifer Fogarty, Robert Bartha, and Alzheimer's Disease Neuroimaging Initiative. Ventricular enlargement as a possible measure of alzheimer's disease progression validated using the alzheimer's disease neuroimaging initiative database. *Brain*, 131(Pt 9):2443–2454, Sep 2008. 37, 87
- OC Okonkwo, ML Alosco, HR Griffith, MM Mielke, LM Shaw, JQ Trojanowski, and G Tremont. Cerebrospinal fluid abnormalities and rate of decline in everyday function across the dementia spectrum:normal aging, mild cognitive impairment, and alzheimer disease. *Archives of Neurology*, 67(6):688–696, 2010. URL <http://archneur.ama-assn.org/cgi/content/full/67/6/688>. 87
- Graziella Orrù, William Pettersson-Yeo, Andre F Marquand, Giuseppe Sartori, and Andrea Mechelli. Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci Biobehav Rev*, 36(4):1140–1152, Apr 2012. doi: 10.1016/j.neubiorev.2012.01.004. URL <http://dx.doi.org/10.1016/j.neubiorev.2012.01.004>. 54, 58
- MJ Pencina, RB D'Agostino Sr, RB D'Agostino Jr, and RS Vasan. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Statistics in medicine*, 27(2):157–172, 2008. ISSN 1053-8119. 91, 122
- Hanchuan Peng, Fuhui Long, and Chris Ding. Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1226–1238, aug. 2005. ISSN 0162-8828. doi: 10.1109/TPAMI.2005.159. 70, 92

REFERENCIAS

- Francisco Pereira, Tom Mitchell, and Matthew Botvinick. Machine learning classifiers and fmri: a tutorial overview. *Neuroimage*, 45(1 Suppl):S199–S209, Mar 2009. doi: 10.1016/j.neuroimage.2008.11.007. URL <http://dx.doi.org/10.1016/j.neuroimage.2008.11.007>. 54, 55
- R. C. Petersen, P. S. Aisen, L. A. Beckett, M. C. Donohue, A. C. Gamst, D. J. Harvey, C. R. Jack, W. J. Jagust, L. M. Shaw, A. W. Toga, J. Q. Trojanowski, and M. W. Weiner. Alzheimer’s disease neuroimaging initiative (adni). *Neurology*, 74(3):201–209, 2010. doi: 10.1212/WNL.0b013e3181cb3e25. URL <http://www.neurology.org/content/74/3/201.abstract>. 20
- 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, Mar 2010. doi: 10.1016/j.neuroimage.2009.11.046. URL <http://dx.doi.org/10.1016/j.neuroimage.2009.11.046>. 52, 71, 72, 73
- Joseph Price, Andy Ko, Marcus Wade, Sarah Tsou, Daniel McKeel, and John Morris. Neuron number in the entorhinal cortex and ca1 in preclinical alzheimer disease. *Arch Neurol*, 58(9):1395–402, 2001. 4, 13
- B. Reisberg, S. H. Ferris, M. J. de Leon, and T. Crook. The global deterioration scale for assessment of primary degenerative dementia. *Am J Psychiatry*, 139(9):1136–1139, Sep 1982. 22
- Martin Reuter, Herminia Diana Rosas, and Bruce Fischl. Highly accurate inverse consistent registration: A robust approach. *NeuroImage*, 53(4):1181–1196, 2010. doi: 10.1016/j.neuroimage.2010.07.020. URL <http://dx.doi.org/10.1016/j.neuroimage.2010.07.020>. 88
- Gerard R Ridgway, Susie M D Henley, Jonathan D Rohrer, Rachael I Scahill, Jason D Warren, and Nick C Fox. Ten simple rules for reporting voxel-based morphometry studies. *Neuroimage*, 40(4):1429–1435, May 2008. doi: 10.1016/j.neuroimage.2008.01.003. URL <http://dx.doi.org/10.1016/j.neuroimage.2008.01.003>. 34
- Shannon L. Risacher, Li Shen, John D. West, Sungeun Kim, Brenna C. McDonald, Laurel A. Beckett, Danielle J. Harvey, Clifford R. Jack Jr., Michael W. Weiner, and Andrew J. Saykin. Longitudinal mri atrophy biomarkers: Relationship to conversion in the adni cohort. *Neurobiology of Aging*, 31(8):1401–1418, 2010. 32, 83, 87, 125
- A. Rocchi, S. Pellegrini, G. Siciliano, and L. Murri. Causative and susceptibility genes for alzheimer’s disease: a review. *Brain Res Bull*, 61(1):1–24, Jun 2003. 25
- Kohavi Ron and John George H. Wrappers for feature subset selection. *ARTIFICIAL INTELLIGENCE*, 97(1): 273–324, 1997. 69, 70, 76, 92
- D. H. Salat, S. Y. Lee, A. J. van der Kouwe, D. N. Greve, B. Fischl, and H. D. Rosas. Age-associated alterations in cortical gray and white matter signal intensity and gray to white matter contrast. *Neuroimage*, 48(1):21–28, Oct 2009. doi: 10.1016/j.neuroimage.2009.06.074. URL <http://dx.doi.org/10.1016/j.neuroimage.2009.06.074>. 37
- Barbara Schneider, David Prvulovic, Viola Oertel-Knöchel, Christian Knöchel, Britta Reinke, Martin Grexa, Bernhard Weber, and Harald Hampel. Biomarkers for major depression and its delineation from neurodegenerative disorders. *Prog Neurobiol*, 95(4):703–717, Dec 2011. doi: 10.1016/j.pneurobio.2011.08.001. URL <http://dx.doi.org/10.1016/j.pneurobio.2011.08.001>. 2
- F. Segonne, A. M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl. A hybrid approach to the skull stripping problem in mri. *NeuroImage*, 22(3):1060–1075, 2004. ISSN 1053-8119. doi: DOI:10.1016/j.neuroimage.2004.03.032. URL <http://www.sciencedirect.com/science/article/B6WNP-4CF5CNY-1/2/33cc73136f06f019b2c11023e7a95341>. 37, 112

REFERENCIAS

- Reisa A Sperling, Paul S Aisen, Laurel A Beckett, David A Bennett, Suzanne Craft, Anne M Fagan, Takeshi Iwatsubo, Clifford R Jack, Jeffrey Kaye, Thomas J Montine, Denise C Park, Eric M Reiman, Christopher C Rowe, Eric Siemers, Yaakov Stern, Kristine Yaffe, Maria C Carrillo, Bill Thies, Marcelle Morrison-Bogorad, Molly V Wagster, and Creighton H Phelps. Toward defining the preclinical stages of alzheimer's disease: recommendations from the national institute on aging-alzheimer's association workgroups on diagnostic guidelines for alzheimer's disease. *Alzheimers Dement*, 7(3):280–292, May 2011. doi: 10.1016/j.jalz.2011.03.003. URL <http://dx.doi.org/10.1016/j.jalz.2011.03.003>. ii, xi, xv, 2, 4, 10, 11, 12, 17, 19, 20, 25, 30
- Rainulf A. Stelzmann, H. Norman Schnitzlein, and F. Reed Murtagh. An english translation of alzheimer's 1907 paper, "über eine eigenartige erkankung der hirnrinde". *Clinical Anatomy*, 8(6):429–431, 1995. ISSN 1098-2353. doi: 10.1002/ca.980080612. URL <http://dx.doi.org/10.1002/ca.980080612>. 3
- Y. Stern, B. Gurland, T. K. Tatemichi, M. X. Tang, D. Wilder, and R. Mayeux. Influence of education and occupation on the incidence of alzheimer's disease. *JAMA*, 271(13):1004–1010, Apr 1994. 89
- J Talairach and P Tournoux. Co-planar stereotaxic atlas of the human brain. *New York: Thieme*, 1988. 40
- Rudolph E Tanzi and Lars Bertram. Twenty years of the alzheimer's disease amyloid hypothesis: a genetic perspective. *Cell*, 120(4):545–555, Feb 2005. doi: 10.1016/j.cell.2005.02.008. URL <http://dx.doi.org/10.1016/j.cell.2005.02.008>. 25
- C. E. Teunissen, J. de Vente, H. W M Steinbusch, and C. De Bruijn. Biochemical markers related to alzheimer's dementia in serum and cerebrospinal fluid. *Neurobiol Aging*, 23(4):485–508, 2002. 25
- Charlotte E Teunissen and Philip Scheltens. Biomarkers for alzheimer's disease. which way to go? *Neurobiol Aging*, 25(6):695–696, Jul 2004. doi: 10.1016/j.neurobiolaging.2003.08.013. URL <http://dx.doi.org/10.1016/j.neurobiolaging.2003.08.013>. 125, 126
- Dietmar R Thal, Udo Rüb, Mario Orantes, and Heiko Braak. Phases of a beta-deposition in the human brain and its relevance for the development of ad. *Neurology*, 58(12):1791–1800, Jun 2002. xi, 7, 8
- Duygu Tosun, Norbert Schuff, Diana Truran-Sacrey, Leslie M Shaw, John Q Trojanowski, Paul Aisen, Ronald Peterson, Michael W Weiner, and Alzheimer's Disease Neuroimaging Initiative. Relations between brain tissue loss, csf biomarkers, and the apoe genetic profile: a longitudinal mri study. *Neurobiol Aging*, 31(8):1340–1354, Aug 2010. 25
- Cynthia Tremblay, Mireille Pilote, Alix Phivilay, Vincent Emond, David A Bennett, and Frédéric Calon. Biochemical characterization of abeta and tau pathologies in mild cognitive impairment and alzheimer's disease. *J Alzheimers Dis*, 12(4):377–390, Dec 2007. 20
- S. A H van Hooren, A. M. Valentijn, H. Bosma, R. W H M Ponds, M. P J van Boxtel, and J. Jolles. Cognitive functioning in healthy older adults aged 64-81: a cohort study into the effects of age, sex, and education. *Neuropsychol Dev Cogn B Aging Neuropsychol Cogn*, 14(1):40–54, Jan 2007. doi: 10.1080/138255890969483. URL <http://dx.doi.org/10.1080/138255890969483>. 89
- Vladimir N. Vapnik. *The Nature of Statistical Learning Theory*. Springer, 1995. 56, 57
- Bruno Vellas, Paul S Aisen, Cristina Sampaio, Maria Carrillo, Philip Scheltens, Bruno Scherrer, Giovanni B Frisoni, Michael Weiner, Lon Schneider, Serge Gauthier, Christine C Gispén de Wied, Suzanne Hendrix, Howard

REFERENCIAS

- Feldman, Jesse Cedarbaum, Ronald Petersen, Eric Siemers, Sandrine Andrieu, David Prvulovic, Jacques Touchon, and Harald Hampel. Prevention trials in alzheimer's disease: an eu-us task force report. *Prog Neurobiol*, 95(4):594–600, Dec 2011. doi: 10.1016/j.pneurobio.2011.08.014. URL <http://dx.doi.org/10.1016/j.pneurobio.2011.08.014>. 1, 2
- Prashanthi Vemuri, Jeffrey L Gunter, Matthew L Senjem, Jennifer L Whitwell, Kejal Kantarci, David S Knopman, Bradley F Boeve, Ronald C Petersen, and Clifford R Jack. Alzheimer's disease diagnosis in individual subjects using structural mr images: validation studies. *Neuroimage*, 39(3):1186–1197, Feb 2008. doi: 10.1016/j.neuroimage.2007.09.073. URL <http://dx.doi.org/10.1016/j.neuroimage.2007.09.073>. 43, 52, 71, 72, 73
- Stephanie Vos, Ineke van Rossum, Leah Burns, Dirk Knol, Philip Scheltens, Hilka Soinen, Lars-Olof Wahlund, Harald Hampel, Magda Tsolaki, Lennart Minthon, Ron Handels, Gilbert L'italien, Wiesje van der Flier, Pauline Aalten, Charlotte Teunissen, Frederik Barkhof, Kaj Blennow, Robin Wolz, Daniel Rueckert, Frans Verhey, and Pieter Jelle Visser. Test sequence of csf and mri biomarkers for prediction of ad in subjects with mci. *Neurobiol Aging*, Jan 2012. doi: 10.1016/j.neurobiolaging.2011.12.017. URL <http://dx.doi.org/10.1016/j.neurobiolaging.2011.12.017>. 25
- K.B. Walhovd, A.M. Fjell, A.M. Dale, L.K. McEvoy, J. Brewer, D.S. Karow, D.P. Salmon, and C. Fennema-Notestine. Multi-modal imaging predicts memory performance in normal aging and cognitive decline. *Neurobiology of Aging*, 31(7):1107–1121, 2010. 32
- Shijun Wang and Ronald M Summers. Machine learning and radiology. *Med Image Anal*, Feb 2012. doi: 10.1016/j.media.2012.02.005. URL <http://dx.doi.org/10.1016/j.media.2012.02.005>. xii, 52, 53
- Michael W. Weiner, Dallas P. Veitch, Paul S. Aisen, Laurel A. Beckett, Nigel J. Cairns, Robert C. Green, Danielle Harvey, Clifford R. Jack, William Jagust, Enchi Liu, John C. Morris, Ronald C. Petersen, Andrew J. Saykin, Mark E. Schmidt, Leslie Shaw, Judith A. Siuciak, Holly Soares, Arthur W. Toga, and John Q. Trojanowski. The alzheimer's disease neuroimaging initiative: A review of papers published since its inception. *Alzheimer's and Dementia*, 8(1, Supplement):S1–S68, 2012. 79, 80
- Eric Westman, Andrew Simmons, Yi Zhang, J-Sebastian Muehlboeck, Catherine Tunnard, Yawu Liu, Louis Collins, Alan Evans, Patrizia Mecocci, Bruno Vellas, Magda Tsolaki, Iwona Kłoszewska, Hilka Soinen, Simon Lovestone, Christian Spenger, and Lars-Olof Wahlund. Multivariate analysis of mri data for alzheimer's disease, mild cognitive impairment and healthy controls. *NeuroImage*, 54(2):1178–1187, 2011. 32, 87
- Ming-Shiang Wu, Tsuo-Hung Lan, Chun-Min Chen, Heng-Chia Chiu, and Tzuo-Yun Lan. Socio-demographic and health-related factors associated with cognitive impairment in the elderly in taiwan. *BMC Public Health*, 11:22, 2011. doi: 10.1186/1471-2458-11-22. URL <http://dx.doi.org/10.1186/1471-2458-11-22>. 89
- Lei Yu and Huan Liu. Efficient feature selection via analysis of relevance and redundancy. *J. Mach. Learn. Res.*, 5: 1205–1224, December 2004. ISSN 1532-4435. URL <http://dl.acm.org/citation.cfm?id=1005332.1044700>. 70
- Henrik Zetterberg. New diagnostic criteria for alzheimer's disease. *Biomark Med*, 5(4):407–409, Aug 2011. doi: 10.2217/bmm.11.46. URL <http://dx.doi.org/10.2217/bmm.11.46>. 11
- Henrik Zetterberg and Kaj Blennow. Biological csf markers of alzheimer's disease. *Handb Clin Neurol*, 89:261–268, 2008. doi: 10.1016/S0072-9752(07)01224-9. URL [http://dx.doi.org/10.1016/S0072-9752\(07\)01224-9](http://dx.doi.org/10.1016/S0072-9752(07)01224-9). 25

REFERENCIAS

Henrik Zetterberg, Kaj Blennow, and Eric Hanse. Amyloid beta and app as biomarkers for alzheimer's disease. *Exp Gerontol*, 45(1):23–29, Jan 2010. doi: 10.1016/j.exger.2009.08.002. URL <http://dx.doi.org/10.1016/j.exger.2009.08.002>. 25