



INSTITUTO TECNOLÓGICO DE BUENOS AIRES

ESCUELA DE INGENIERÍA Y TECNOLOGÍA

BIOINGENIERÍA - Proyecto Final de Carrera

Desarrollo de algoritmo para asistencia al diagnóstico diferencial de Alzheimer y Demencia fronto-temporal

Autoras:

Camila GUARDIA, 58722

María Antonia SARACCO, 58473

Tutores:

Cecilia GONZALEZ CAMPO, Médica

Sebastián MOGUILNER, PhD en Biología

*TRABAJO FINAL PRESENTADO PARA LA OBTENCIÓN DEL TÍTULO DE
BIOINGENIERAS*

Buenos Aires, 7 de Septiembre 2022

Índice general

Lista de figuras	6
Lista de cuadros	11
Lista de acrónimos	12
1. Resumen	16
2. Planteo del tema y justificación	18
3. Hipótesis	20
3.1. Objetivos generales	20
3.2. Objetivos específicos	20
4. Marco teórico	22
4.1. Demencias	22
4.2. Enfermedad de Alzheimer	23
4.2.1. Epidemiología y factores de riesgo	23
4.2.2. Cuadro clínico	24

4.3.	Demencia fronto-temporal	24
4.3.1.	Epidemiología y factores de riesgo	24
4.3.2.	Cuadro clínico	25
4.4.	Resonancia magnética	26
4.4.1.	Secuencias FLAIR	28
4.4.2.	RM en EA y DFT	31
4.5.	Tests cognitivos	33
4.5.1.	INECO Frontal Screening	34
4.5.2.	Montreal Cognitive Assessment	35
4.6.	Inteligencia Artificial y Machine Learning	36
4.6.1.	Python	37
4.6.2.	Algoritmos	38
4.6.3.	Entrenamiento del modelo	43
4.6.4.	Validación y testeo	44
4.6.5.	Evaluación del modelo	46
5.	Metodologías y técnicas a utilizar	49
5.1.	Base de datos	49
5.2.	Estadística y análisis de datos	51
5.3.	Machine Learning	53
5.3.1.	Preprocesamiento	53
5.3.2.	Algoritmos	54

6. Resultados	56
6.1. Estadística	56
6.2. EA-Control	63
6.2.1. Datos combinados: SVM	63
6.2.2. Datos neurorradiológicos: RF	64
6.2.3. Datos cognitivos: XGB	65
6.2.4. Métricas y comparación	66
6.3. EA-DFT	67
6.3.1. Datos combinados: GB	67
6.3.2. Datos neurorradiológicos: GB	68
6.3.3. Datos cognitivos: GB	69
6.3.4. Métricas y comparación	70
6.4. DFT-Control	71
6.4.1. Datos combinados: SVM	71
6.4.2. Datos neurorradiológicos: RF	72
6.4.3. Datos cognitivos: RF	73
6.4.4. Métricas y comparación	74
6.5. Métricas y comparación general	75
7. Conclusión y discusión	76
8. Tutores	80

Bibliografía	87
A. Anexo	87
A.1. EA-Control	87
A.1.1. Datos combinados	87
A.1.2. Datos neurorradiológicos	89
A.1.3. Datos cognitivos	91
A.2. EA-DFT	93
A.2.1. Datos combinados	93
A.2.2. Datos neurorradiológicos	95
A.2.3. Datos cognitivos	96
A.3. DFT-Control	98
A.3.1. Datos combinados	98
A.3.2. Datos neurorradiológicos	100
A.3.3. Datos cognitivos	102

Índice de figuras

2.1. Número de personas con demencia (millones) en países de bajo y mediano ingreso comparado con países de alto ingreso[3]	19
4.1. Efectos TR en el contraste de la imagen. Un TR más largo permite más tiempo para la relajación de T1 y produce mayor señal de los tejidos con valores de T1 largos. (a) Imagen T1 con TR de 500 ms; (b) Imagen FLAIR con TR prolongado (2000 ms). [9]	27
4.2. Imágenes obtenidas utilizando (a) Secuencia ponderada T1; (b) Secuencia ponderada T2; (c) Secuencia FLAIR (imagen ponderada en T2 con supresión de líquido). [10]	28
4.3. Secuencia FLAIR normal. [11]	29
4.4. Imágenes ponderadas en secuencia FLAIR. Figura proporcionada por Prins et al., 2015.	30
4.5. Resonancias magnéticas de una presentación típica de la enfermedad de Alzheimer (EA) de inicio tardío. RM potenciadas en T1 axial (a, c) y coronal (b) que muestran atrofia de las regiones temporoparietales, incluidos los hipocampos (flechas). La RM FLAIR (d) muestra hiperintensidades de la sustancia blanca compatibles con una enfermedad cerebrovascular, a menudo coexistente con la EA. Cortesía: FILIPPI et al., 2022 [1]	31
4.6. Los pacientes con la vcDFT exhiben atrofia prominente frontal, insular y en corteza cingulada anterior. [1]	32

4.7. Representación de SVM	39
4.8. Representación de SVM con un espacio dimensional superior	40
4.9. Representación de Random Forest	41
4.10. Representación de Gradient Boosting	42
4.11. Validación cruzada	45
4.12. Flujo de <i>Grid Search</i>	45
4.13. Ejemplo de matriz de confusión	47
4.14. Ejemplo de curva ROC	48
5.1. Representación gráfica que muestra algunos tractos de interés [47]. . .	50
6.1. Boxplot de edades previo al filtrado	57
6.2. Boxplot de edades luego de filtrar	58
6.3. Histograma de la distribución por género para cada clase	59
6.4. HSB según clase	59
6.5. Distribución de las HSB distribuidas por área en la clase Control. . .	61
6.6. Distribución de las HSB distribuidas por área en la clase EA.	61
6.7. Distribución de las HSB distribuidas por área en la clase DFT.	61
6.8. Resultados del algoritmo SVM para los grupos de EA y Control con datos combinados	63
6.9. Resultados del algoritmo RF para los grupos de EA y Control con datos neurorradiológicos	64
6.10. Resultados del algoritmo XGB para los grupos de EA y Control con datos cognitivos	65

6.11. Resultados del algoritmo GB para los grupos de EA y DFT con datos combinados	67
6.12. Resultados del algoritmo GB para los grupos de EA y DFT con datos neurorradiológicos	68
6.13. Resultados del algoritmo GB para los grupos de EA y DFT con datos cognitivos	69
6.14. Resultados del algoritmo SVM para los grupos de DFT y Control con datos combinados	71
6.15. Resultados del algoritmo RF para los grupos de DFT y Control con datos neurorradiológicos	72
6.16. Resultados del algoritmo RF para los grupos de DFT y Control con datos cognitivos	73
A.1. Resultados del algoritmo SVM para los grupos de EA y Control	87
A.2. Resultados del algoritmo RF para los grupos de EA y Control	88
A.3. Resultados del algoritmo GB para los grupos de EA y Control	88
A.4. Resultados del algoritmo XGB para los grupos de EA y Control	89
A.5. Resultados del algoritmo RF para los grupos de EA y Control	89
A.6. Resultados del algoritmo GB para los grupos de EA y Control	90
A.7. Resultados del algoritmo XGB para los grupos de EA y Control	90
A.8. Resultados del algoritmo SVM para los grupos de EA y Control	91
A.9. Resultados del algoritmo RF para los grupos de EA y Control	91
A.10. Resultados del algoritmo GB para los grupos de EA y Control	92
A.11. Resultados del algoritmo XGB para los grupos de EA y Control	92

A.12.Resultados del algoritmo SVM para los grupos de EA y DFT	93
A.13.Resultados del algoritmo RF para los grupos de EA y DFT	93
A.14.Resultados del algoritmo GB para los grupos de EA y DFT	94
A.15.Resultados del algoritmo XGB para los grupos de EA y DFT	94
A.16.Resultados del algoritmo RF para los grupos de EA y DFT	95
A.17.Resultados del algoritmo GB para los grupos de EA y DFT	95
A.18.Resultados del algoritmo XGB para los grupos de EA y DFT	96
A.19.Resultados del algoritmo SVM para los grupos de EA y DFT	96
A.20.Resultados del algoritmo RF para los grupos de EA y DFT	97
A.21.Resultados del algoritmo GB para los grupos de EA y DFT	97
A.22.Resultados del algoritmo XGB para los grupos de EA y DFT	98
A.23.Resultados del algoritmo SVM para los grupos de DFT y Control . . .	98
A.24.Resultados del algoritmo RF para los grupos de DFT y Control . . .	99
A.25.Resultados del algoritmo GB para los grupos de DFT y Control . . .	99
A.26.Resultados del algoritmo XGB para los grupos de DFT y Control . . .	100
A.27.Resultados del algoritmo RF para los grupos de DFT y Control . . .	100
A.28.Resultados del algoritmo GB para los grupos de DFT y Control . . .	101
A.29.Resultados del algoritmo XGB para los grupos de DFT y Control . . .	101
A.30.Resultados del algoritmo SVM para los grupos de DFT y Control . . .	102
A.31.Resultados del algoritmo RF para los grupos de DFT y Control . . .	102
A.32.Resultados del algoritmo GB para los grupos de DFT y Control . . .	103

A.33.Resultados del algoritmo XGB para los grupos de DFT y Control . . . 103

Índice de cuadros

4.1. Subtest del IFS	35
6.1. Valores del Test de Shapiro Wilk	56
6.2. Cantidad de sujetos	57
6.3. Estadísticas básicas de edad	58
6.4. Resultados del test de Tukey-Kramer-HSD	60
6.5. Resultados del test Spearman-R	60
6.6. Resultados del test Tukey-Kramer para cada área	62

Lista de acrónimos

ADI: Alzheimer's Disease International

ABC: Área bajo la curva

CC_D: Haz cíngulo-hipocampal derecho

CC_I: Haz cíngulo-hipocampal izquierdo

CE_D: Tracto corticoespinal derecho

CE_I: Tracto corticoespinal izquierdo

CH_D: Hipocampo del cíngulo derecho

CH_I: Hipocampo del cíngulo izquierdo

DFT: Demencia fronto-temporal

DSM V: Libro *“Diagnostic and Statistical Manual of Mental Disorders” quinta edición*

EA: Enfermedad de Alzheimer

ECV: Enfermedad cerebrovascular

FIFO_D: Fascículo inferior fronto occipital derecho

FIFO_I: Fascículo inferior fronto occipital izquierdo

FLI_D: Fascículo longitudinal inferior derecho

FLI_I: Fascículo longitudinal inferior izquierdo

FLS_D: Fascículo longitudinal superior derecho

FLS_I: Fascículo longitudinal superior izquierdo

FLST_D: Fascículo longitudinal superior parte temporal derecho

FLST_I: Fascículo longitudinal superior parte temporal izquierdo

FMA: Fórceps mayor

FME: Fórceps menor

FN: Falso negativo

FP: Falso positivo

FPR: Tasa de falsos positivos (false positive rate)

FU_D: Fascículo uncinado derecho

FU_I: Fascículo uncinado izquierdo

GB: Gradient Boosting

HSB: Hiperintensidad de la sustancia blanca

IFS: INECO Frontal Screening

IFS_D: Subtest de repetición de dígitos atrás del IFS

IFS_IC: Subtest de instrucciones conflictivas del IFS

IFS_IM: Subtest de control inhibitorio motor del IFS

IFS_IV: Subtest de control inhibitorio verbal del IFS

IFS_M: Subtest de repetición de meses atrás del IFS

IFS_P: Subtest de refranes del IFS

IFS_SM: Subtest de series motoras del IFS

IFS_V: Subtest de memoria de trabajo visual del IFS

LCR: Líquido cefalorraquídeo

ML: Machine Learning

MoCA: Evaluación Cognitiva de Montreal (Montreal cognitive assessment)

nfvPPA: Variante no fuente/agramática de la afasia primaria progresiva

RF: Random Forest

RM: Resonancia magnética

ROC: Receiver Operating Characteristic

RTA_D: Radiación talámica anterior derecha

RTA_I: Radiación talámica anterior izquierda

TE: Time to Echo

TPR: Tasa de verdaderos positivos(true positive rate)

TR: Repetition Time

SB: Sustancia blanca

SSDC: Sistema de soporte a la decisión clínica

SVM: Support Vector Machine

svPP: Variante semántica de la afasia primaria progresiva

vcDFT: Variante conductual de DFT

VIT: Volumen intracraneal total

VN: Verdadero negativo

VP: Verdadero positivo

XGB: XG Boost

Capítulo 1

Resumen

Se ha sugerido el uso de enfoques basados en datos de imágenes automáticos o semi-automáticos para apoyar el diagnóstico de algunas enfermedades. En este contexto, Machine Learning (ML) aparece como una herramienta útil, permitiendo desde extracción de características hasta clasificación automática.

La Enfermedad de Alzheimer (EA) y la Demencia fronto-temporal (DFT) son dos formas comunes y prevalentes de demencia con síntomas y patrones de atrofia cerebral diferentes, pero parcialmente superpuestos.

La base de datos utilizada fue provista por el Centro de Neurociencias Cognitivas de la Universidad de San Andrés y consistió de información sobre volúmenes de hiperintensidades de sustancia blanca (HSB) a partir de resonancias magnéticas (RM) en secuencia FLAIR, edad, género, años de educación y test cognitivos de 190 sujetos siendo la distribución: 105 sujetos de control, 54 sujetos con EA y 31 sujetos con DFT. Estos datos se utilizaron como entrada de cada modelo tanto de manera conjunta como separando en las siguientes categorías: datos neuroradiológicos, exámenes cognitivos y datos demográficos. Previamente, dichos datos fueron sometidos a evaluaciones estadísticas y de análisis de datos para estudiar la distribución de los mismos, sus relaciones, correlaciones y obtener un entendimiento de las clases a trabajar.

En el análisis realizado se comparó el rendimiento de cuatro modelos de Machine learning: Support Vector Machine (SVM), Random Forest (RF), Gradient Boosting (GB) y XGBoost (XGB). Para realizar la evaluación se utilizó la validación de 5

capas cruzadas.

En cuanto a la evaluación de los modelos propuestos se contó con la matriz de confusión, curva ROC-ABC y la métrica de F1. Cada uno de ellos brinda información que permitió comparar los distintos algoritmos para poder encontrar el mejor de ellos en las categorías planteadas.

El resultado más relevante y novedoso obtenido fue el modelo de GB utilizando datos provenientes de imágenes de resonancia magnética en la clasificación de EA con DFT. Dicho resultado fue $ABC = 0.92$.

El objetivo de este modelo es asistir al diagnóstico médico, por lo que se lo considera un sistema de soporte a las decisiones clínicas (SSDC). El propósito del algoritmo no es tomar decisiones o realizar un diagnóstico formal, sino ayudar a quien esté realizando la evaluación para que logre tomar una decisión informada a partir de toda la información disponible.

Capítulo 2

Planteo del tema y justificación

La EA es la principal causa de demencia en personas mayores. Se caracteriza por afectar en particular la memoria del paciente. En tanto, la DFT es menos frecuente, sin embargo, también es una causa importante de demencia que produce cambios en la personalidad y una disminución en las interacciones personales.

La EA se define principalmente en términos de sus características cognitivas: alteraciones en la memoria, las funciones visuoespaciales y el lenguaje y, son estos cambios en las funciones instrumentales los que fundamentan el diagnóstico clínico.

La DFT se refiere principalmente a un grupo de afecciones cerebrales neurodegenerativas caracterizadas por la atrofia de los lóbulos frontal y temporal anterior. Las características sobresalientes de la DFT son alteraciones profundas en la personalidad y la conducta social.

Si bien la EA y la DFT son trastornos distintos que se caracterizan por afectar diferentes regiones del cerebro y producir síntomas clínicos particulares, no obstante, muchas de sus características se superponen, especialmente en fase tardía. La RM es una de las herramientas más utilizadas en el abordaje diagnóstico de las demencias. Diversos signos radiológicos son utilizados para caracterizar estas patologías.

Sin embargo, al igual que en su expresión clínica, existen importantes superposiciones de las manifestaciones radiológicas entre la DFT y la EA, siendo en muchos casos imposible aproximarse a un diagnóstico clínico-radiológico.

Un abordaje analítico más preciso combinado con el empleo de técnicas de inteligencia artificial, podrían aumentar la precisión y reducir el tiempo del diagnóstico, ampliar la cantidad de personas que pueden ser examinadas y ayudar a las organizaciones de atención médica a optimizar sus operaciones. Esta estrategia podría brindar diversas soluciones a los problemas actuales observados en la mayoría de los sistemas de salud y, además, podría tener un impacto favorable en el diagnóstico y generar nuevos conocimientos sobre la fisiopatología de las principales enfermedades causales de demencia.

Según Alzheimer's Disease International (ADI), en el año 2020 se estimó que más de 55 millones de personas en todo el mundo vivían con demencia. Este número casi se duplicará cada 20 años, llegando a 82 millones en 2030 y 152 millones en 2050.[2]

Gran parte del aumento se producirá en los países en desarrollo. Esto, se debe a que los países más pobres tienen menos recursos económicos y humanos para satisfacer las necesidades de atención sanitaria y social de sus poblaciones de personas mayores que crecen rápidamente. El 60 % de las personas con demencia ya vive en países de ingresos bajos y medios y se espera un aumento del 71 % para 2050.

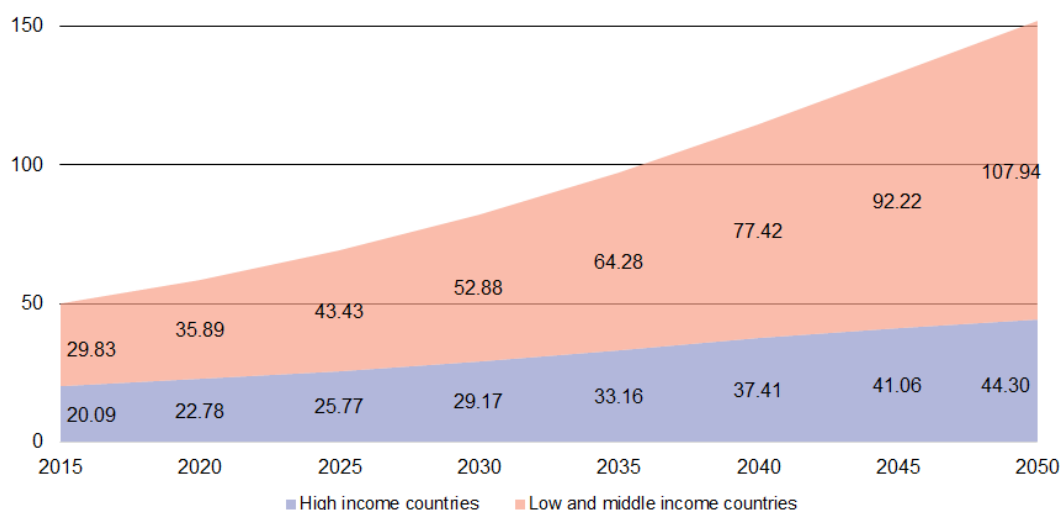


Figura 2.1: Número de personas con demencia (millones) en países de bajo y mediano ingreso comparado con países de alto ingreso[3]

En la figura 2.1 se puede observar que la prevalencia de las demencias se aumentará en los países de altos ingresos, en tanto, el aumento más significativo se verá en los países de bajo y mediano ingreso.

Capítulo 3

Hipótesis

Un algoritmo de ML que emplee datos multivariados de pruebas cognitivas y de imágenes, podrá asistir y mejorar el diagnóstico diferencial de las demencias de tipo Alzheimer y fronto-temporal.

3.1. Objetivos generales

Crear un algoritmo de ML que emplee datos multivariados de pruebas cognitivas y de imágenes, para poder asistir al diagnóstico diferencial de las demencias de tipo Alzheimer y fronto-temporal.

3.2. Objetivos específicos

Comparar la clasificación realizada por los distintos modelos: SVM, RF, GB y XGB bajo distintos escenarios.

Estos escenarios difieren según los datos de entrada que se le proporciona a los modelos, siendo los mismos:

- Datos provenientes únicamente de imágenes de RM: hiperintensidades de la sustancia blanca total y regional por tractos.

- Datos provenientes únicamente de evaluaciones cognitivas: Test de MoCA y IFS.
- Datos provenientes tanto de imágenes como de evaluaciones cognitivas

Además, se analizará la correlación entre los parámetros obtenidos a partir de las imágenes de resonancia magnética con los datos cognitivos y demográficos.

La finalidad del presente estudio es encontrar una herramienta de análisis que sea capaz de incorporar información adicional y así asistir a los médicos especialistas en la realización del diagnóstico.

Capítulo 4

Marco teórico

4.1. Demencias

La demencia es un grupo de condiciones caracterizadas por el deterioro de al menos dos funciones cerebrales, como la pérdida de la memoria y el juicio. Los síntomas incluyen olvidos, habilidades sociales limitadas y habilidades de pensamiento tan deterioradas que interfieren con el funcionamiento diario. Existen condiciones potencialmente reversibles que pueden causar o imitar la demencia.

Sin embargo, la mayoría de las demencias están causadas por enfermedades primarias degenerativas, progresivas e irreversibles [1].

Un diagnóstico temprano y preciso de la demencia es esencial para permitir que la enfermedad se aborde con la intervención disponible o experimental, cambios en el estilo de vida y arreglos logísticos, antes de que se desarrolle la discapacidad.

En el diagnóstico temprano y diferencial de enfermedades causales de demencia neurodegenerativa, la historia clínica debe centrarse en los dominios cognitivos afectados, el curso de la enfermedad, el impacto en la actividad de la vida diaria y cualquier síntoma no cognitivo asociado. Los antecedentes médicos, las comorbilidades, los antecedentes familiares y educativos son muy importantes.

Se pueden distinguir distintos tipos de demencia.

Este trabajo final, se enfoca en el estudio de dos demencias neurodegenerativas, la EA y la DFT, dos de las enfermedades más frecuentes causales de demencia.

4.2. Enfermedad de Alzheimer

La EA es un trastorno neurodegenerativo progresivo, irreversible e incapacitante, que lentamente genera déficits en la memoria y en las habilidades cognitivas generales.

4.2.1. Epidemiología y factores de riesgo

La EA es la causa más común de demencia en la gente mayor y representa hasta el 70-75 % de todos los casos de demencia.

Su prevalencia se duplica cada 5 años después de los 65 años, oscilando entre aproximadamente el 2 % y el 3 % en personas de 65 a 69 años hasta el 30 % en personas de mas de 80 años. Su prevalencia es mayor en mujeres que en hombres; en el grupo de edad de 65 a 69 años, el 0,7 % de las mujeres y el 0,6 % de los hombres padecen la enfermedad con frecuencias crecientes del 14,2 % y del 8,8 % en individuos de 85 a 89 años. [1]

Esta enfermedad fue descubierta por el Dr. Alois Alzheimer quien trató una paciente diagnosticada con una enfermedad mental desconocida cuyos síntomas incluyeron la perdida del la memoria, problemas de lenguaje y un comportamiento impredecible. En la autopsia post mortem, el cerebro de la paciente mostró varias anomalías. La corteza cerebral era más delgada de lo normal y se encontró placa senil, que anteriormente solo se encontraba en personas mayores, en el cerebro junto con ovillos neurofibrilares. Con acceso a una nueva tinción, Alzheimer pudo identificar estos ovillos nerviosos que nunca antes se habían descrito [4].

Se han identificado varios factores de riesgo para la EA esporádica. Algunos de estos son: diabetes mellitus, tabaquismo, depresión, sedentarismo, etc. [1]

También se han identificado diferentes factores genéticos que aumentan el riesgo esporádico de EA. El alelo 4 del gen de la apolipoproteína E (APOE) es el principal

factor de riesgo genético único para la EA esporádica. [1]

4.2.2. Cuadro clínico

Las manifestaciones clínicas tempranas de la EA pueden ser cognitivas o neuroconductuales (p. ej., depresión, ansiedad y apatía).

La presentación típica y más común de la EA es la de un individuo de edad mayor a 65 años con problemas insidiosos y progresivos de memoria episódica (es decir, la capacidad de aprender y retener nueva información) solo o en combinación con la participación de otros dominios cognitivos. Típicamente, en esta etapa, se puede observar una preservación de la independencia en las habilidades funcionales, definiendo así el estado amnésico de Deterioro Cognitivo Leve (DCL). A medida que avanza la condición, las dificultades cognitivas se vuelven más profundas y generalizadas, afectando los otros dominios cognitivos y finalmente interfiriendo con las actividades de la vida diaria.

Más raramente, la patología de EA subtiende otras presentaciones clínicas que se incluyen en la definición general de “síndromes de EA atípicos” y se caracterizan típicamente por una edad de inicio más temprana (es decir, <65 años) y una preservación general de la memoria a corto plazo en el etapas iniciales de la enfermedad.

4.3. Demencia fronto-temporal

La DFT es un trastorno neurodegenerativo que se presenta como un conjunto de síntomas comportamentales y cognitivos principalmente secundarios a la neurodegeneración y atrofia en regiones frontales y temporales del cerebro.

4.3.1. Epidemiología y factores de riesgo

La DFT, un trastorno neurodegenerativo heterogéneo, es una causa común de demencia de inicio temprano (menor a 65 años).

La prevalencia puntual estimada es de 15 a 22/100 000 y la incidencia de 2,7 a

4,1/100 000. Alrededor del 25 % son casos de aparición tardía. [5]

Los estudios de población muestran una distribución casi equitativa por género, lo que contrasta con la mirada de informes clínicos y de neuropatologías.

La DFT es frecuentemente familiar y hereditaria. Los datos convergentes de estudios familiares, clínicos y de población muestran que la DFT es familiar en el 30-50 % de los casos, y hasta en el 40 % es un pedigrí autosómico dominante. [5]

4.3.2. Cuadro clínico

La DFT consta de tres variantes clínicas que se distinguen por los síntomas de presentación predominante: variante conductual de la DFT (vcDFT), variante semántica de la afasia primaria progresiva (svPPA) y variante no fuente/agramática de la afasia primaria progresiva (nfvPPA) La vcDFT es la variante más común de DFT y comprende más del 50 % de todos los casos. Por otro lado, se encuentra comprendida en el grupo de las demencias frontotemporales de inicio tardío (generalmente a partir de los 65 años) [6].

Los criterios diagnósticos básicos de la vcDFT propuestos por Rascovsky et al. [7] en 2011 incluyen:

- desinhibición conductual;
- apatía o inercia;
- pérdida de simpatía o empatía;
- perseverancia, comportamiento estereotipado o compulsivo/ritualista;
- hiperoralidad;
- disfunción ejecutiva.

Todos menos uno de estos criterios (disfunción ejecutiva) son de naturaleza conductual y tienen características superpuestas con muchos síndromes psiquiátricos, que incluyen esquizofrenia, trastorno obsesivo-compulsivo, trastorno bipolar y trastorno depresivo mayor. Debido a esto, muchos pacientes con vcDFT son diagnosticados erróneamente con una afección psiquiátrica primaria.

4.4. Resonancia magnética

La RM es una técnica no invasiva de adquisición de imágenes que permite obtener detalles anatómicos con capacidades tridimensionales completas basándose en la resonancia magnética nuclear. Desde el desarrollo del primer escáner y la obtención de las primeras imágenes en la década del '70, esta técnica demostró ser la que mejor permite observar a los tejidos blandos sin el uso de radiación ionizante y de manera no invasiva. Además de ser de gran utilidad en las neurociencias para obtener información anatómica a nivel de la sustancia blanca y sustancia gris, a diferencia de otras metodologías imagenológicas (p.ej., tomografía computada), la RM puede ser utilizada para estudiar el funcionamiento del cerebro.

La RM utiliza el fenómeno de la resonancia magnética nuclear. En esta, ciertos átomos como los de Hidrógeno, poseen una propiedad denominada “spin” (precesar). Estos, en la ausencia de un campo magnético precesan en direcciones aleatorias. En cambio, cuando se introduce un determinado campo, estos spins se alinean o se anti-alinean al mismo. Debido a que más spins se alinean que los que se anti-alinean, se obtiene una magnetización neta (M_0).

En la RM se denomina relajación al proceso por el cual los protones liberan energía después de haber sido excitados con un pulso de radio frecuencia.

Se pueden medir dos tiempos de relajación, conocidos como T1 y T2. Ambos tiempos miden la transferencia espontánea de energía por parte de un protón excitado, pero difieren en la disposición final de la energía.

El tiempo de relajación T1 es el tiempo necesario para que la componente z de M vuelva al 63 % de su valor original después de un pulso de excitación. También se conoce como tiempo de relajación de red de espín o tiempo de relajación longitudinal. Este retorno de la magnetización sigue un proceso de crecimiento exponencial, siendo T1 la constante de tiempo que describe la tasa de crecimiento:

$$M(\tau) = M_0(1 - e^{-\frac{\tau}{T1}}) \quad (4.1)$$

donde τ es el tiempo que sigue al pulso de radiofrecuencia. Después de tres periodos de tiempo T1, M habrá vuelto al 95 % de su valor anterior al pulso de excitación, M_0 .

Por otro lado, la variación en la secuencia de pulsos de radiofrecuencia afecta el tipo de imagen que se obtiene a la salida. “Repetition Time” (TR) es el tiempo que transcurre entre secuencias de pulsos sucesivos sobre el mismo corte. “Time to Echo” (TE) es el tiempo entre la emisión del pulso y la recepción de la señal por la antena. Según los valores en los que se los configure, al tejido se lo puede caracterizar según los tiempos de relajación mencionados: T1 y T2.

Las imágenes ponderadas en T1 requieren tiempos TR y TE cortos (en general 500 ms y 30 ms respectivamente). Largos períodos de TR minimizan los efectos T1 puesto que los tejidos no tienen tiempo suficiente para recuperarse de la excitación.

En el caso de las imágenes ponderadas en T2, los tiempos de TR y TE son más largos (1500 ms y 90 ms respectivamente). Períodos cortos de TE minimizan los efectos T2 dado que no se le dio el tiempo suficiente para que los protones excitados decaigan, y por ende, los tejidos no se logren diferenciar correctamente.

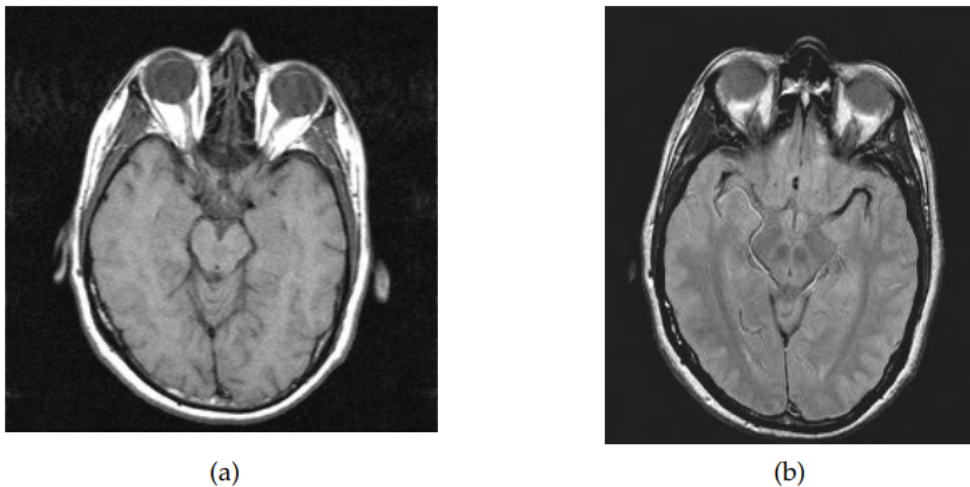


Figura 4.1: Efectos TR en el contraste de la imagen. Un TR más largo permite más tiempo para la relajación de T1 y produce mayor señal de los tejidos con valores de T1 largos. (a) Imagen T1 con TR de 500 ms; (b) Imagen FLAIR con TR prolongado (2000 ms). [9]

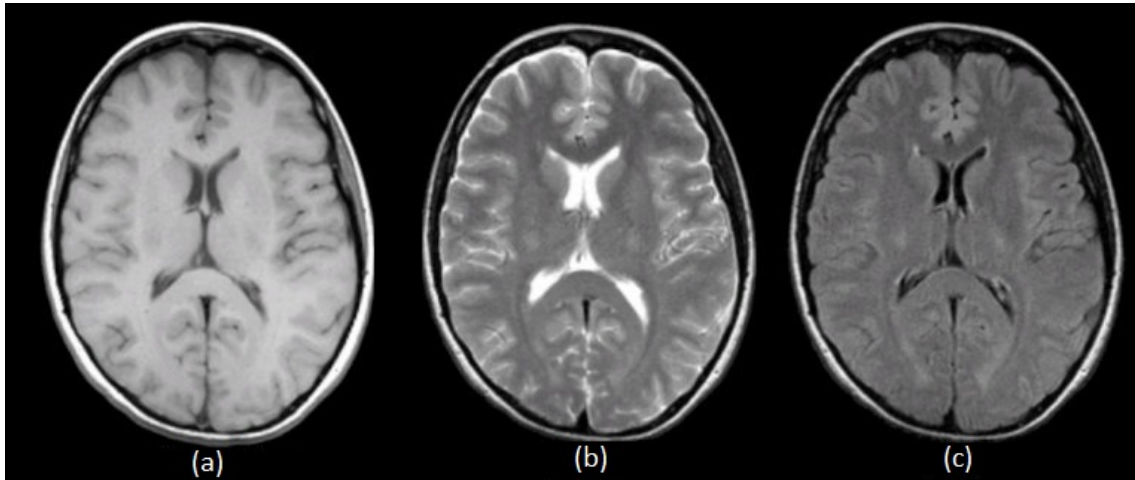


Figura 4.2: Imágenes obtenidas utilizando (a) Secuencia ponderada T1; (b) Secuencia ponderada T2; (c) Secuencia FLAIR (imagen ponderada en T2 con supresión de líquido). [10]

En imágenes cerebrales, T1 y T2 se pueden diferenciar fundamentalmente por el brillo que toma el líquido cefalorraquídeo. Mientras que en T1 se ve más bien oscuro/negro, en las imágenes T2 se ve brillante/blanco. Además, en la secuencia T1 pueden diferenciarse estructuras anatómicas compuestas por sustancia blanca y sustancia gris, lo cual no es posible en secuencias T2 oscuro FLAIR. Estos efectos se pueden apreciar en las figuras 4.1 y 4.2.

Combinando de diversas maneras los gradientes y pulsos, se pueden diseñar numerosas secuencias para adaptarse a diferentes características del tejido. Las secuencias más utilizadas hoy en día en RM son las T1, T2, FLAIR (Fluid-attenuated inversion recovery), y difusión, entre otras.

4.4.1. Secuencias FLAIR

La secuencia FLAIR es una secuencia de inversión de recuperación especial con un tiempo de inversión prolongado. Esto, elimina la señal del líquido cefalorraquídeo en las imágenes resultantes de la figura 4.3. El tejido cerebral en las imágenes FLAIR se ve similar a las imágenes potenciadas, pero el LCR (líquido cefalorraquídeo) es oscuro en lugar de brillante. [11]

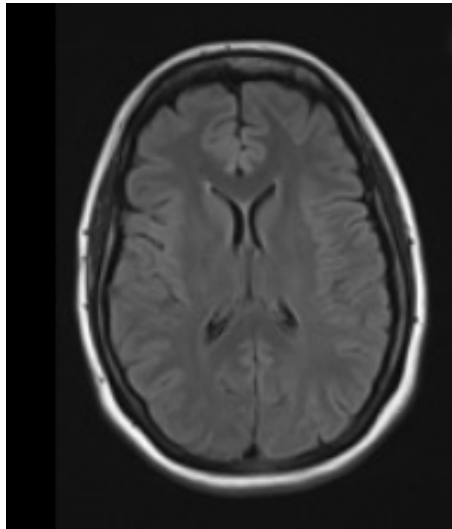
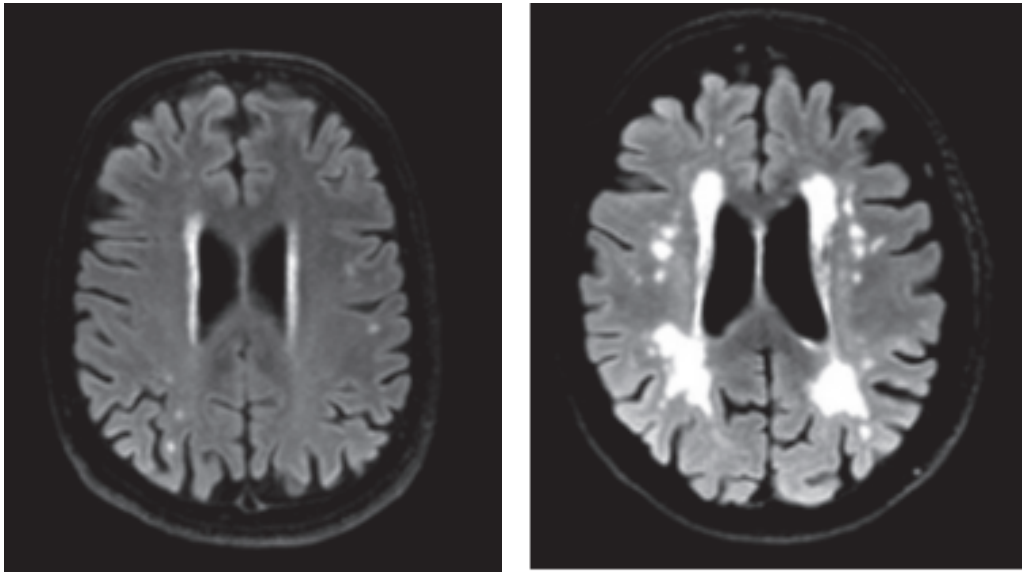


Figura 4.3: Secuencia FLAIR normal. [11]

En cuanto a los principios físicos de esta secuencia, para anular la señal del fluido, el tiempo de inversión (TI) de la secuencia de pulsos FLAIR se ajusta de modo que en el equilibrio no haya magnetización transversal neta del fluido.

La secuencia FLAIR forma parte de casi todos los protocolos de imagen del cerebro, particularmente útil en la detección de cambios sutiles en la periferia cerebral y en la región periventricular cercana al LCR.

Es de remarcar que las HSB (hierintensidades de la sustancia blanca) son uno de los hallazgos más frecuentemente observados en RM asociados a la edad, y suelen ser interpretados, al menos en parte, como de origen vascular.[15]



(a) Corte axial de cerebro en el que se aprecian HSB puntiformes subcorticales. (b) Corte axial de cerebro en el que se aprecian HSB periventriculares.

Figura 4.4: Imágenes ponderadas en secuencia FLAIR. Figura proporcionada por Prins et al., 2015.

Las HSB se encuentran entre los cambios más prominentes relacionados con la edad observados en RM. Las HSB se ven como áreas focales o difusas de alta intensidad de señal (por lo tanto, “hiperintensas”) en secuencias FLAIR/T2 (ejemplos en la Fig. 4.4). Las HSB son, en términos generales, equivalentes a la leucoaraiosis observada en las tomografías computarizadas.

Las HSB suelen interpretarse como un sustituto de la enfermedad cerebral de vasos pequeños. Debido a que diversas patologías pueden provocar un aumento de la intensidad de la señal de SB (sustancia blanca) en RM, como es la esclerosis múltiple y otras enfermedades cerebrales inflamatorias o leucodistrofias metabólicas, las HSB deben ser consideradas en un contexto clínico y etario, y en función de su distribución regional en el cerebro [13].

Amplia evidencia respalda una asociación entre un mayor volumen de HSB y disminuciones en el rendimiento cognitivo global o específico del dominio. Se considera que las HSB están particularmente correlacionadas con reducciones en la velocidad de procesamiento de la información y la función ejecutiva, aunque también se han observado correlaciones con otros dominios cognitivos [13].

Los estudios longitudinales en diversas poblaciones demuestran consistentemente que el aumento del volumen de HSB predice el deterioro cognitivo, el deterioro cognitivo leve, la demencia incidente, el accidente cerebrovascular y la muerte [13].

4.4.2. RM en EA y DFT

Patrón de atrofia cerebral

Las enfermedades degenerativas causan de manera global una reducción del volumen cerebral o atrofia.

La EA típica conduce a una reducción del tejido cerebral, particularmente los lóbulos temporales mediales (LTM). El hipocampo y la corteza entorrinal, se encuentran entre los sitios más tempranamente afectados. Otras regiones gravemente afectadas incluyen la parte posterior de la circunvolución del cíngulo, el precúneo y el esplenio del cuerpo calloso en la superficie medial, y las regiones parietal, temporal superior posterior y frontal en las superficies cerebrales laterales.

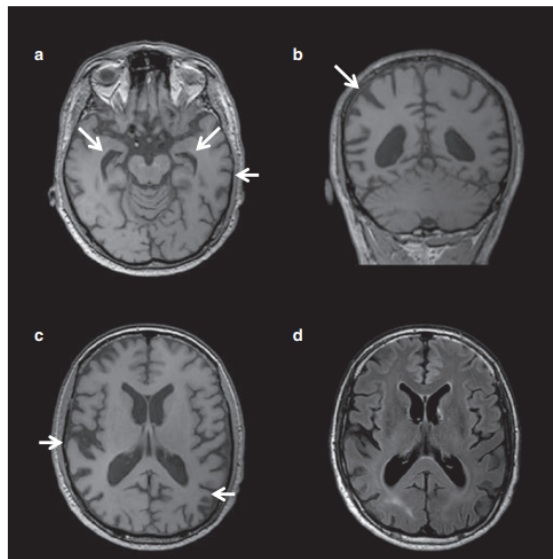


Figura 4.5: Resonancias magnéticas de una presentación típica de la enfermedad de Alzheimer (EA) de inicio tardío. RM potenciadas en T1 axial (a, c) y coronal (b) que muestran atrofia de las regiones temporoparietales, incluidos los hipocampos (flechas). La RM FLAIR (d) muestra hiperintensidades de la sustancia blanca compatibles con una enfermedad cerebrovascular, a menudo coexistente con la EA. Cortesía: FILIPPI et al., 2022 [1]

Los pacientes con vcDFT suelen presentar una combinación de atrofia cortical temporal anterior, insular, cingulada anterior y orbitofrontal medial con un gradiente anteroposterior y preservación relativa de los lóbulos parietal y occipital.

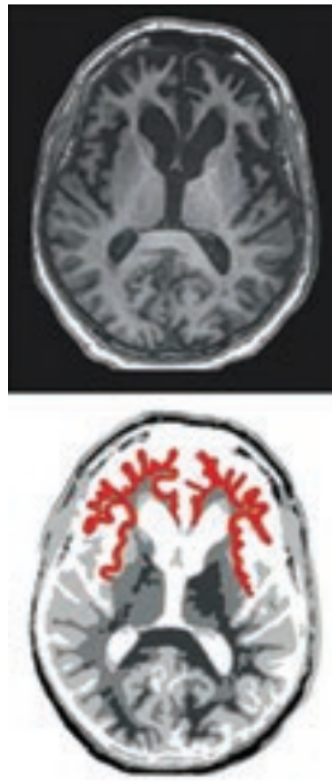


Figura 4.6: Los pacientes con la vcDFT exhiben atrofia prominente frontal, insular y en corteza cingulada anterior. [1]

El lóbulo temporal medial también está afectado, con mayor daño en la amígdala y el hipocampo anterior que en las estructuras posteriores. La atrofia cerebral en la vcDFT también involucra varias estructuras subcorticales, incluido el cuerpo estriado, el tálamo y el tronco encefálico.

Si bien cada patología suele presentar un patrón específico de atrofia, muchas áreas suelen superponerse, y en particular, en etapas avanzadas, la afectación suele ser tan generalizada que las diferencias no suelen ser evidentes.

Hiperintensidades de la sustancia blanca en EA y DFT

Dada la alta prevalencia de patología mixta de EA y ECV (enfermedad cerebrovascular), las HSB han sido más estudiadas en sujetos con EA y su carga se ha asociado a riesgo de desarrollo de demencia, incluida la EA suelen verse con frecuencia en sujetos con EA [12] [14].

Actualmente, es posible encontrar una cantidad significativamente mayor de estudios para la afección de HSB en EA que en DFT.

En DFT, solo se han realizado algunos estudios de cohortes, y solo en casos genéticos de DFT. En particular, se ha mostrado un aumento del volumen de HSB en portadores de mutaciones de progranulina (GRN) (Sudre CH et al, Neuroimage Clin 2017)

También en un estudio reciente se estudiaron las correlaciones clínicas y biológicas de las HSB en pacientes con vcDFT y EA, los volúmenes totales de HSB fueron mayores en los pacientes con vcDFT que en los pacientes con EA y controles. En vcDFT, los volúmenes de HSB fueron asociados con la severidad de la enfermedad y no así con riesgo vascular. Los pacientes con vcDFT y los pacientes con EA mostraron distintos patrones espaciales de HSB que reflejaban patrones característicos de atrofia cortical. [15]

4.5. Tests cognitivos

El DSM V (Libro *“Diagnostic and Statistical Manual of Mental Disorders” quinta edición*) define la demencia como un trastorno neurocognitivo mayor, en el que existe evidencia de un declive cognitivo significativo comparado con el nivel previo de rendimiento en uno o más dominios cognitivos (atención compleja, función ejecutiva, aprendizaje y memoria, lenguaje, habilidad perceptual motora o cognición social) que puede estar basado en la preocupación del propio individuo, en un informante que le conoce o en el clínico. Este declive en la función cognitiva debe estar preferentemente documentado por un test neuropsicológico estandarizado o, en su defecto, por otra evaluación clínica cuantitativa. Dos de los tests más utilizados son el Ineco Frontal Screening y el Montreal cognitive Assesment que se presentan a

continuación.

4.5.1. INECO Frontal Screening

La prueba INECO Frontal Screening (IFS) fue diseñada para proporcionar a los profesionales de la salud una prueba de detección ejecutiva sensible y específica para determinar la disfunción frontal en pacientes con demencia.

El diseño del IFS se concibió principalmente para representar tres grupos de tareas, de la siguiente manera:

- a) inhibición de respuesta y cambio de conjunto: evalúa la capacidad de cambiar de un conjunto cognitivo a otro e inhibir la respuesta inapropiada de forma verbal y motora;
- b) capacidad de abstracción: obtenida a partir de la interpretación de proverbios, siendo la interpretación concreta propia de los pacientes con lesión del lóbulo frontal;
- c) memoria de trabajo: se refiere a un sistema cerebral que proporciona almacenamiento temporal y manipulación de la información necesaria para otras tareas cognitivas complejas.

Para poder realizar la evaluación de estas capacidades, se diseñaron ocho subtests. Estos cuentan cada uno con una valoración en particular para que, al final, el puntaje total obtenido en el examen pueda ayudar a determinar la disfunción frontal de los pacientes.

Función ejecutiva	IFS subtest
<i>Inhibición de respuesta y cambio de set</i>	<i>Programación de motores</i> <i>Instrucciones contradictorias</i> <i>Ir-No ir</i> <i>Control inhibitorio verbal</i>
<i>Capacidad de abstracción</i>	<i>Interpretación de proverbio</i>
<i>Memoria de trabajo</i>	<i>Intervalo de dígitos hacia atrás</i> <i>Memoria de trabajo verbal</i> <i>Memoria de trabajo espacial</i>

Cuadro 4.1: Subtest del IFS

4.5.2. Montreal Cognitive Assessment

El examen Montreal Cognitive Assessment (MoCA) es un instrumento breve de tamizaje, sensible y específico, utilizado para la detección del deterioro cognitivo leve (DCL) y la demencia, con puntos de corte que varían según la población estudiada.

Sus características psicométricas describen un alto nivel de confiabilidad y validez con una sensibilidad de 87 % y una especificidad en un rango de 90 % para DCL, y una sensibilidad del 87 % para la demencia de tipo Alzheimer. [17]

El MoCA es una prueba de 30 puntos de una página que se administra en 10 minutos.

Los detalles sobre los artículos MoCA específicos son los siguientes:

- La tarea de recuerdo de memoria a corto plazo (5 puntos) implica dos ensayos de aprendizaje de cinco sustantivos y recuerdo diferido después de aproximadamente 5 minutos.
- Las habilidades visuoespaciales se evalúan mediante una tarea de dibujo de un reloj (3 puntos) y una copia de un cubo tridimensional (1 punto).
- Se evalúan múltiples aspectos de las funciones ejecutivas mediante una tarea de alternancia adaptada de la tarea Trail Making B (1 punto), una tarea de

- fluidez fonémica (1 punto) y una tarea de abstracción verbal de dos elementos (2 puntos).
- d) La atención, la concentración y la memoria de trabajo se evalúan mediante una tarea de atención sostenida (detección de objetivos mediante tapping; 1 punto), una tarea de resta en serie (3 puntos) y dígitos hacia adelante y hacia atrás (1 punto cada uno).
 - e) El lenguaje se evalúa mediante una tarea de denominación de confrontación de tres elementos con animales poco familiares (león, camello, rinoceronte; 3 puntos), repetición de dos oraciones sintácticamente complejas (2 puntos) y la tarea de fluidez antes mencionada.
 - f) Finalmente, se evalúa la orientación en tiempo y lugar (6 puntos).

4.6. Inteligencia Artificial y Machine Learning

La inteligencia artificial le permite a los sistemas y máquinas aprender a partir de la experiencia, ajustarse a nuevas entradas y realizar tareas de una manera muy similar a la que un humano la realizaría. Actualmente la inteligencia artificial se encuentra presente en un amplio espectro de aplicaciones, por ejemplo, vehículos autónomos (aquellos que no necesitan un conductor para manejar) o chatbots conversacionales. Por medio de la tecnología, una computadora puede ser entrenada para realizar una tarea específica por medio de la revisión y procesamiento de grandes cantidades de datos para encontrar patrones que se repiten en la misma [18].

ML y Deep Learning son subconjuntos de la inteligencia artificial. La diferencia entre ML e inteligencia artificial reside en que la última hace referencia al método en donde las máquinas simulan actos y comportamientos realizando decisiones inteligentes, mientras que ML corresponde al proceso en donde las máquinas aprenden de manera autónoma sin código explícito. Además el Deep Learning refiere a las redes neuronales con más de una capa que aprenden y realizan decisiones independientes con la ayuda de algoritmos de ML [19].

Lo primero y principal que requiere el ML son datos. Sin datos relevantes directamente no hay ML ya que los algoritmos no podrán funcionar correctamente.

Dichos datos se dividen en dos categorías: Training y Testing (entrenamiento y evaluación). Es importante tener información suficientes en ambas categorías para que el algoritmo se pueda entrenar con datos que representen la totalidad de lo que se desea obtener [20]. Los datos de Training y Testing se suelen dividir en distintas proporciones para asegurar una cantidad específica en cada categoría. Las proporciones más frecuentemente utilizadas son 70/30 o 80/20 [21] [22].

Una vez que se definieron los dos sets de datos, el siguiente paso es definir un modelo y entrenarlo. Actualmente existen muchos algoritmos de ML que se pueden utilizar, cada uno diseñado para suplir soluciones a problemas específicos. Es por eso que se suele escoger un modelo que se ajuste a la cuestión que se quiere resolver.

Además, luego del entrenamiento se debe evaluar al modelo. Para ello se utiliza el set de datos que fue seleccionado para Testing.

Otro concepto importante en el ámbito del ML es el ajuste de hiper-parámetros. En la inteligencia artificial los hiper-parámetros hacen referencia a aquellos parámetros que no pueden ser explicados utilizando el modelo de ML en cuestión, sino aquellos que definen la arquitectura del modelo. Los hiper-parámetros son importantes para mejorar el rendimiento del modelo, y por esta razón, deben ser representados. Estos, son definidos manualmente. El ajuste de hiper-parámetros consiste en una combinación de los mismos tal que mejore el rendimiento del modelo [23] [24] [25] [26].

4.6.1. Python

Python es un lenguaje de programación de alto nivel, que se utiliza para una amplia gama de aplicaciones. Es un lenguaje interpretado en donde no es necesario compilarlo para poder ejecutar los comandos escritos, sino que se puede ejecutar directamente utilizando un interpretador que lo traduce a lenguaje de máquina.

Además debido a la similitud con el lenguaje humano, es relativamente sencillo de leer y cuenta con la ventaja de ser código abierto (o open-source) que permite crear desarrollos sin límites. En los últimos años, Python se ha convertido relevante debido a que permite trabajar con ML, Big Data, inteligencia artificial y data science [27].

Las librerías de Python son líneas de código reusable que se utilizan cuando se escribe un programa o se crea un proyecto. Éstas librerías son, en esencia, un conjunto de módulos. Las principales librerías para el ML son:

- NumPy: ofrece funciones y operaciones matemáticas, generadores de números, vectores, matrices y funcionalidades básicas para trabajar con ellos [28].
- Pandas: es una librería rápida con herramientas de análisis y manipulación. Principalmente permite trabajar con excels y estructuras de datos DataFrames [29].
- Scikit-Learn: es la principal librería relacionada al ML ya que contiene múltiples algoritmos desde regresión, agrupamiento o clustering y clasificación. Funciona en conjunto con otras librerías de Python, entre ellas NumPy, SciPy y Matplotlib. Está principalmente basada en el análisis predictivo de datos [30] [31].

4.6.2. Algoritmos

Para el desarrollo de la presente tesis nos enfocaremos en los algoritmos de clasificación de aprendizaje supervisado. El resultado de dichos algoritmos es una clase entre un número limitado de clases. En la detección diferencial de EA y DFT las clases posibles son: EA positivo, DFT positivo o control positivo (el paciente no presenta ninguna de las dos enfermedades).

SVM: Support Vector Machine

El algoritmo de SVM es un método de clasificación que se considera un referente por su buena clasificación en un amplio número de situaciones. Se basa en que, dado un conjunto de datos de entrenamiento, es posible etiquetar las clases y entrenar una máquina de vectores de soportes para construir un modelo que prediga la clase nueva que se desea predecir.

Para su funcionamiento construye un hiperplano, el cual, es un espacio multi-dimensional diseñado para separar las diferentes clases. Su objetivo consta en encontrar aquel hiperplano que mejor divida el conjunto de datos. Visualmente, en un

espacio 2D el hiperplano será una línea mientras que en un espacio 3D, el hiperplano consistirá de un plano.

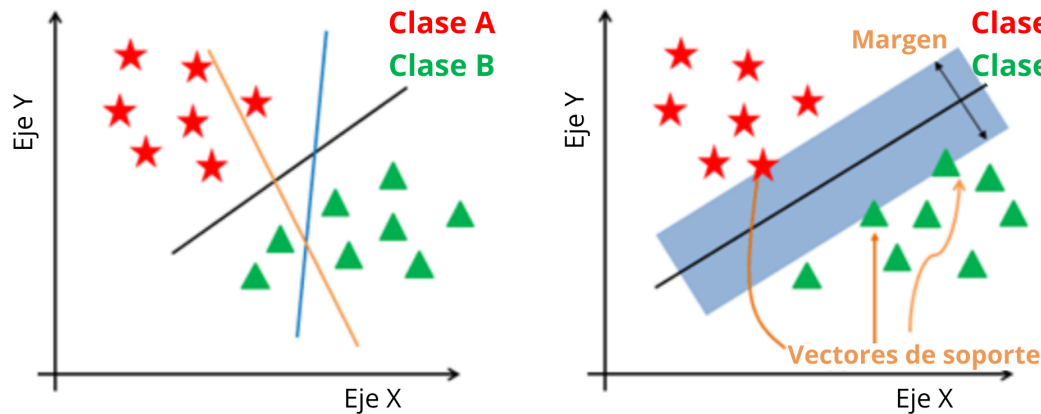


Figura 4.7: Representación de SVM

Los vectores de soporte son los puntos de datos más cercanos al hiperplano.

El algoritmo de SVM se basa en asignar un dato a una categoría por medio de una clasificación lineal binaria no-probabilística. Dicho algoritmo puede ser representado como puntos en el espacio cuyo objetivo es encontrar un espacio divisorio entre las categorías.

En general, si los datos pueden ser separados perfectamente por medio de un hiperplano, entonces existe un número infinito de éstos dado que pueden ser rotados o trasladados levemente sin que entren en contacto con algún dato. Es por esto que se utiliza el hiperplano que optimice la separación, es decir, el hiperplano con máximo margen. Se calculan las distancias perpendiculares de cada dato de Training para cada hiperplano (margen) [32] [33].

Muchas veces la clasificación no es lineal, y una simple línea no puede separar correctamente el espacio de datos, en cuyo caso, se utiliza un kernel para convertir el espacio de entrada en un espacio dimensional superior (figura 4.8). El enfoque del kernel es un enfoque computacional para acomodar una división no lineal entre las clases, es una función que cuantifica la similitud de dos datos. El kernel puede ser

de cualquier grado y utilizar un kernel con mayor grado implica una división más flexible [33].

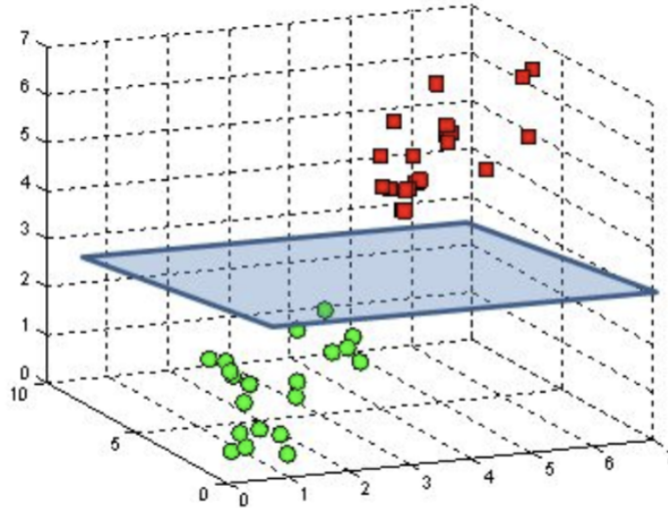


Figura 4.8: Representación de SVM con un espacio dimensional superior

RF: Random Forest

Al igual que SVM, es un algoritmo de aprendizaje supervisado. Un bosque aleatorio está compuesto por árboles y, continuando con la analogía del bosque, cuantos más árboles más robusto y más preciso será el bosque y el modelo.

Los algoritmos de RF o bosques aleatorios se basan en la creación de árboles de decisión sobre los datos seleccionados de manera aleatoria. Cada árbol está compuesto de nodos en donde se evalúa un parámetro o *feature* específico y resulta en una respuesta booleana (sí/no). El nodo siguiente depende de la respuesta del nodo anterior. Como su nombre lo implica, cada árbol toma una decisión sobre el dato que esta evaluando respecto a cada parámetro.

En cada nodo el algoritmo debe tomar la decisión de que parámetro analizar, es por esto que para cada uno de ellos se calcula la ganancia de información que se obtendría si se elige ese parámetro. Dado que se quiere maximizar la ganancia de información, el árbol escoge la pregunta a resolver y el *feature* basándose en aquel que mayor ganancia de información posea.

Los algoritmos de RF pueden contar con decenas o incluso centenas de árboles

de decisión. El número de árboles es un parámetro que debe setearse en la creación del algoritmo.

Al final, cada árbol hace su predicción y el bosque selecciona la solución más votada para luego ponderar todas las salidas (figura 4.9).

Actualmente los bosques aleatorios tienen muchas aplicaciones como: los motores de recomendación, la clasificación de imágenes, los bancos que los utilizan para saber a qué clientes pueden ofrecer o no un producto financiero, para evitar el fraude, entre otras. Es un algoritmo muy popular por su capacidad de combinar los resultados de todos los árboles utilizados para obtener un resultado con mayor nivel de confianza [34].

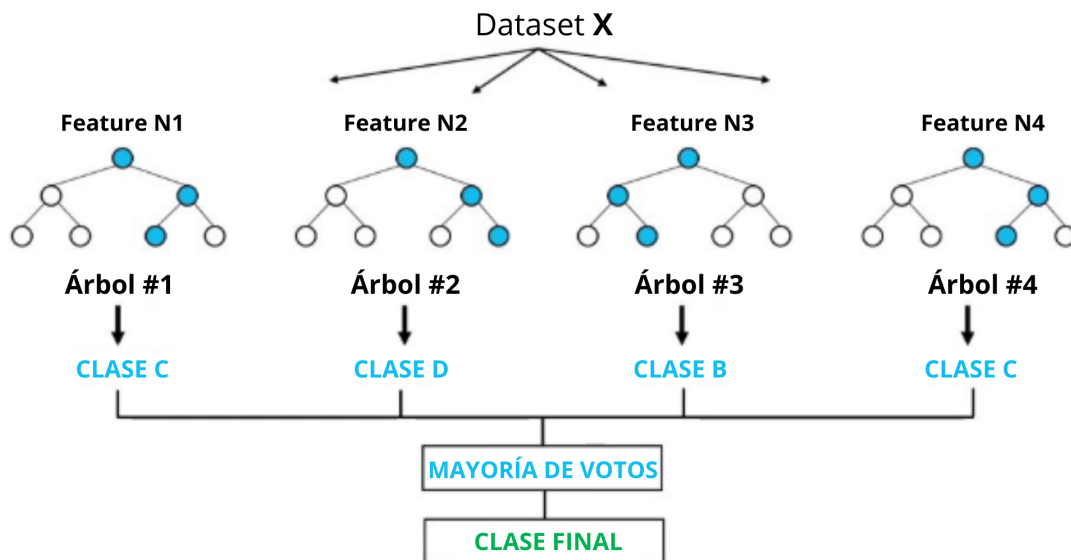


Figura 4.9: Representación de Random Forest

GB: Gradient Boosting

El algoritmo de GB tiene similitudes con RF ya que, al igual que en éste último, la salida final del modelo se obtiene de la combinación de múltiples salidas de árboles individuales.

En este algoritmo lo que se utiliza es el boosting, en donde cada árbol es secuencial al anterior y corrige el error del árbol anterior. Comienza con un solo árbol y utilizando una función denominada loss o pérdida analiza que eficacia del modelo.

La pérdida es alta cuando la predicción y la etiqueta no coinciden [35].

Al agregar un árbol a continuación se busca que la función perdida sea menor de lo que era previamente. Para lograr dicho objetivo, se deriva la función perdida de modo de obtener el gradiente (el cual es la dirección de máximo crecimiento), y se utiliza el negativo (-gradiente) para lograr la dirección de máximo decrecimiento. Luego la salida de ese árbol va a tener en cuenta la salida del anterior, el menos gradiente y un parámetro que es la tasa de aprendizaje. Ésta, se usa para evitar un sobre ajuste y escalar las contribuciones del nuevo árbol. La función de la salida del árbol m se observa en la ecuación 4.2 [36].

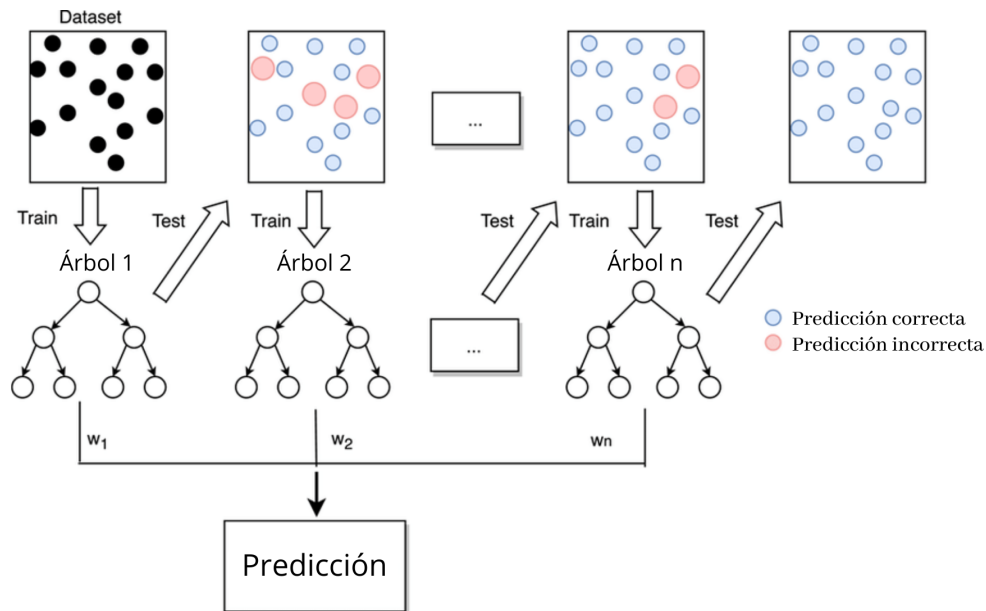


Figura 4.10: Representación de Gradient Boosting

$$F(m) = F(m-1) + \nu * \frac{-\partial L}{\partial F(m-1)} \quad (4.2)$$

XGB: XGBoost

El algoritmo XGB presenta una mejora del modelo anterior ya que utiliza algoritmos de ML bajo el entorno de trabajo del GB. El objetivo es brindar una librería de GB optimizada y diseñada para ser altamente eficaz y flexible [37].

Su principio de funcionamiento se basa en tree boosting en paralelo, de manera

similar al GB. Sin embargo, dicho algoritmo presenta mejoras respecto a este modelo [38], dentro de ellas:

- Regularización: La regularización permite controlar la cantidad de *over-fitting*, por lo que representa una mejora significativa respecto a GB.
- Procesamiento paralelo: Esto logra un algoritmo considerablemente más rápido que un boosting regular.
- Alta flexibilidad: XGB es más flexible que GB ya que le permite al usuario optimizar y personalizar los criterios de evaluación como también objetivos de optimización que se deseen cambiar. Esto, logra quitar las limitaciones y brindarle al algoritmo un mayor espectro de escenarios. Además, XGB puede manejar datos faltantes de mejor manera que GB, el algoritmo intentará realizar distintas tareas mientras localiza los valores faltantes de cada nodo y puede aprender de ello.
- Poda de árboles: GB deja de separar nodos luego de localizar una pérdida negativa. Por ello, es considerado un algoritmo codicioso. En comparación, XGB puede separar nodos hasta que la profundidad máxima del valor se alcance y luego, comienzan los mecanismos de poda de los árboles (donde quita separaciones y vuelve hacia atrás si no encuentra ganancias positivas).
- Cross-Validation: XGB cuenta con un mecanismo de cross-validation interno que se corre en cada repetición. Es por esto que se pueden reducir la cantidad de iteraciones del algoritmo.

4.6.3. Entrenamiento del modelo

Luego de definir un modelo de ML es necesario ajustarlo a los datos que se poseen, en otras palabras, entrenarlo para que pueda aprender de aquellos datos. La calidad de los datos es fundamental ya que, al carecer de ésta, los modelos no serán capaces de extraer patrones y similitudes en los datos para realizar predicciones acertadas.

Para entrenar a un modelo se puede optar por un aprendizaje supervisado o uno no supervisado. La principal diferencia recae en que en el aprendizaje supervisado los datos están previamente etiquetados con lo que debería ser la predicción

final del algoritmo. De esta forma se le enseña al modelo la forma de reconocer y detectar el resultado esperado. Las dos familias de algoritmos que utilizan el aprendizaje supervisado son aquellas que el resultado a predecir es un atributo numérico o categórico.

En cambio, en un aprendizaje no supervisado los datos no fueron previamente etiquetados y para encontrar patrones se basa en inferencias y agrupamiento de los datos (clustering). Del mismo modo, existen formas de entrenamiento híbrido en donde se utiliza una combinación de entrenamiento supervisado y no supervisado [39].

4.6.4. Validación y testeo

El proceso de validación es aquel que permite evaluar el modelo durante el entrenamiento utilizando un conjunto de datos que difiere de aquel que se usa para el testeo y el entrenamiento. La principal ventaja de la validación es que evita el sobreajuste de los datos de entrenamiento, llamado *overfitting*. En un modelo que está sobreajustado se obtienen resultados extremadamente buenos con los datos de entrenamiento, pero al cambiar dichos datos y utilizar los de test, el desempeño decae drásticamente. Esto quiere decir que el modelo no es generalizable a otros datos salvo los utilizados para entrenar [40].

Una técnica de validación muy difundida es la de *K-fold cross validation* o validación cruzada en donde se dividen los datos de entrenamiento en K partes iguales y se repite dicho entrenamiento K veces. En cada una de las repeticiones se separa una partición (fold) y se entrena al modelo con los datos restantes ($K-1$ folds). El fold que fue separado de los datos de entrenamiento se utiliza para realizar una evaluación final de la repetición. Cabe notar que el fold con el que se evalúa el modelo no se utiliza para entrenarlo, logrando reducir el overfitting y junto con un resultado preliminar más acertado y confiable. De esta forma se realizan todas las K repeticiones variando el fold utilizado para la evaluación como se observa en la figura 4.11 [41] [42].

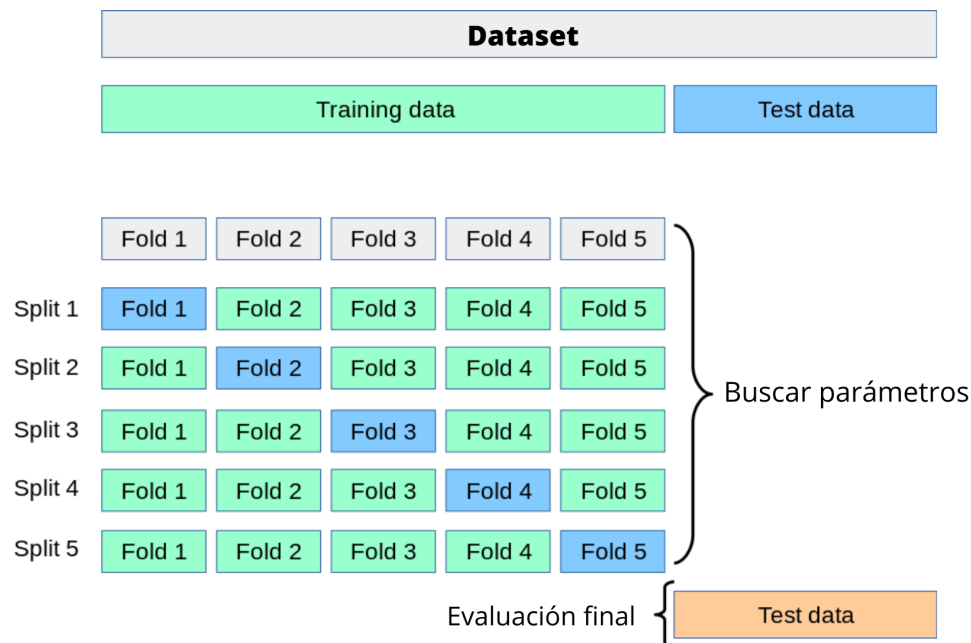
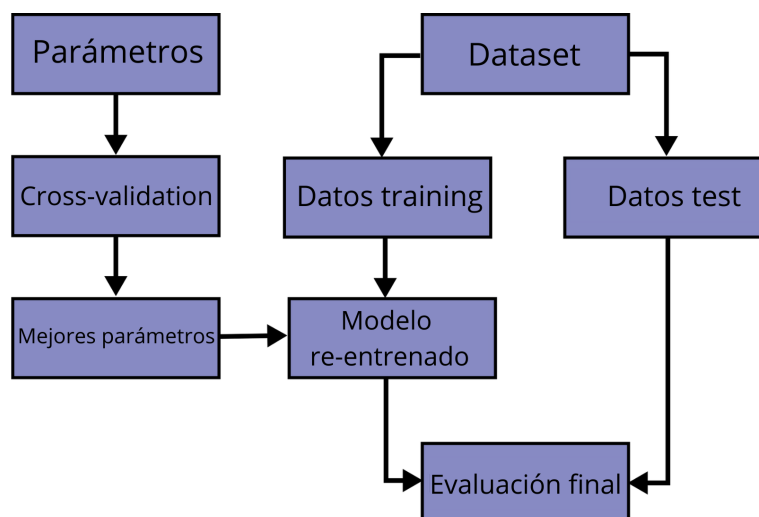


Figura 4.11: Validación cruzada

Además, otro uso que tiene la validación cruzada es iterar sobre el modelo variando los parámetros de cada algoritmo para encontrar la mejor combinación de ellos. Dicho procedimiento se conoce como *Grid Search* y su flujo de funcionamiento se aprecia en la figura 4.12.

Figura 4.12: Flujo de *Grid Search*

Finalmente, una vez que se obtiene un modelo final, se realiza una predicción con datos que no se han utilizado hasta el momento. Dichos datos son los de Tes-

ting. A partir de ellos es posible sacar métricas y resultados para evaluar el modelo final. Al no haberlos utilizado para el entrenamiento y validación permiten obtener información de cuan generalizable y aplicable a datos externos es el modelo.

4.6.5. Evaluación del modelo

Dado que es un modelo de clasificación se suelen evaluar teniendo en cuenta los verdaderos positivos (VP: los valores que el algoritmo clasifica como positivos y que realmente son positivos), verdaderos negativos (VN: los valores que el algoritmo clasifica como positivo cuando realmente son negativos), falsos positivos (FP: los valores que el algoritmo clasifica como negativos y que realmente son negativos) y falsos negativos (FN: los valores que el algoritmo clasifica como negativo cuando realmente son positivos)

Por consiguiente para evaluar los distintos modelos planteados se propusieron las métricas siguientes: [43] [44] [45]

- Accuracy: es calculada como el cociente entre el número total de predicciones correctas sobre el número total de datos.

$$Accuracy = \frac{VP + FP}{VP + VN + FP + FN} \quad (4.3)$$

Sin embargo, la métrica en cuestión brinda información útil cuando las muestras están balanceadas, es decir, cuando se cuenta con igual cantidad de datos en cada clase.

- Recall: también llamada sensibilidad, es el cociente de verdaderos positivos sobre la suma de verdaderos positivos y falsos negativos, es decir, el total de positivos predichos.

$$Recall = \frac{VP}{VP + FN} \quad (4.4)$$

Esta métrica informa sobre el total de valores negativos, cuántos el algoritmo clasificó negativos. En el caso planteado en el informe, es de suma importancia medir el Recall ya que un paciente que presenta la enfermedad pero es predicho

como sano va a llevar a un mal diagnóstico y mal tratamiento del mismo, poniendo en riesgo la vida del paciente.

- Precision: se calcula como el cociente entre los verdaderos positivos sobre la suma de verdaderos positivos y verdaderos negativos

$$Precision = \frac{VP}{VP + FP} \quad (4.5)$$

Suponiendo un modelo con dos clases (por ejemplo positivo y negativo), dicha métrica informa el porcentaje de valores que se han clasificado como positivos que realmente son positivos.

- F1 Score: esta métrica es de gran utilidad cuando los datos están desbalanceados, además se utiliza cuando ambos el Precision y el Recall son importantes ya que los combina para obtener un valor más objetivo.

$$F1score = \frac{2 * Recall * Precision}{Recall + Precision} \quad (4.6)$$

- Matriz de Confusión: también llamada tabla de contingencia, es una herramienta que evidencia el desempeño del algoritmo de clasificación, describiendo cómo se distribuyen los valores reales y las predicciones.

		Predicción	
		Positivos	Negativos
Observación	Positivos	Verdaderos Positivos (VP)	Falsos Negativos (FN)
	Negativos	Falsos Positivos (FP)	Verdaderos Negativos (VN)

Figura 4.13: Ejemplo de matriz de confusión

Como se observa en la figura 4.13, las filas son aquellos elementos que realmente son positivos o negativos mientras que las columnas son aquellos que fueron predichos como positivos o negativos ubicando así, los 4 términos explicados anteriormente (VP, FP, FN, VN).

- Curva ROC-ABC: es una métrica para medir la performance del modelo de clasificación bajo distintos valores umbrales. ROC hace referencia a la curva

(Receiver Operating Characteristic Curve) mientras que ABC corresponde al área bajo la curva. A mayor valor de ABC, el modelo mejor predecirá las clases. Un clasificador con muy buenos resultados tendrá una ABC cercana a 1 mientras que un clasificador de bajo rendimiento tendrá un valor de ABC cercano a 0.

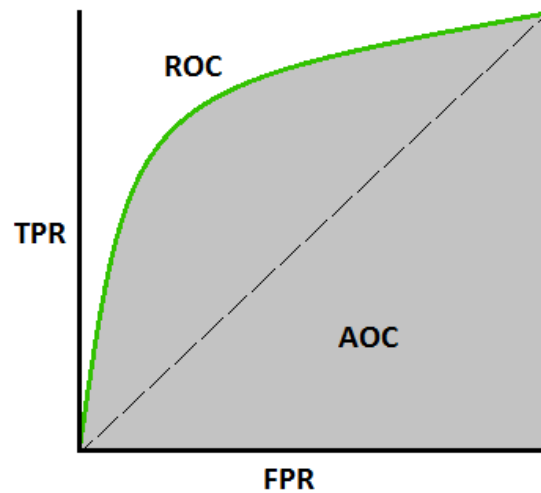


Figura 4.14: Ejemplo de curva ROC

La curva ROC se grafica entre la tasa de verdaderos positivos (TPR: true positive rate) y la tasa de falsos positivos (FPR: false positive rate) (fig. 4.14).

Capítulo 5

Metodologías y técnicas a utilizar

5.1. Base de datos

La base de datos a utilizar en el presente trabajo fue proporcionada por el Centro de Neurociencias Cognitivas de la Universidad de San Andrés. La misma, cuenta con 190 sujetos totales, los cuales se dividen en los siguientes grupos:

- 105 sujetos control,
- 54 sujetos con EA,
- 31 sujetos con DFT.

Para cada sujeto, se realizó una RM de cerebro con el fin de obtener información de las variables descriptas a continuación:

- Volumen total de HSB normalizada por el VIT (volumen intercraneal total),
- Volumen de HSB por tractos normalizado por VIT. Para ello, se utilizó un atlas que divide al cerebro en 20 tractos [\[46\]](#).
 - Radiación talámica anterior izquierda (RTA_I)
 - Radiación talámica anterior derecha (RTA_D)

- Tracto corticoespinal izquierdo (CE_I)
- Tracto corticoespinal derecho (CE_D)
- Haz cíngulo-hipocampal izquierdo (CC_I)
- Haz cíngulo-hipocampal derecho (CC_D)
- Hipocampo del cíngulo izquierdo (CH_I)
- Hipocampo del cíngulo derecho (CH_D)
- Fórceps menor (FME)
- Fórceps mayor (FMA)
- Fascículo fronto occipital inferior izquierdo (FIFO_I)
- Fascículo fronto occipital inferior derecho (FIFO_D)
- Fascículo longitudinal inferior izquierdo (FLI_I)
- Fascículo longitudinal inferior derecho (FLI_D)
- Fascículo longitudinal superior izquierdo (FLS_I)
- Fascículo longitudinal superior derecho (FLS_D)
- Fascículo uncinado izquierdo (FU_I)
- Fascículo uncinado derecho (FU_D)
- Fascículo longitudinal superior parte temporal izquierdo (FLST_I)
- Fascículo longitudinal superior parte temporal derecho (FLST_D)

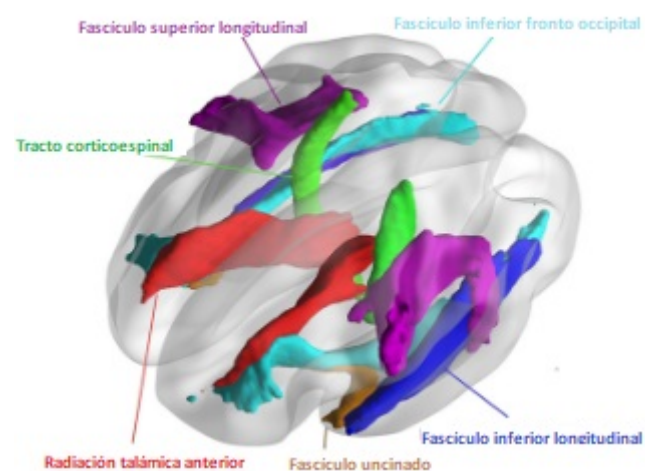


Figura 5.1: Representación gráfica que muestra algunos tractos de interés [47].

Además, se utilizaron variables externas a las resonancias magnéticas, ellas son:

- Edad,
- Género,
- Años de educación,
- Score de MOCA total,
- Score de IFS total,
- Subtest de IFS: series motoras (IFS_SM),
- Subtest de IFS: instrucciones conflictivas (IFS_IC),
- Subtest de IFS: control inhibitorio motor (IFS_IM),
- Subtest de IFS: repetición de dígitos atrás (IFS_D),
- Subtest de IFS: meses atrás (IFS_M),
- Subtest de IFS: memoria de trabajo visual (IFS_V),
- Subtest de IFS: refranes (IFS_P),
- Subtest de IFS: control inhibitorio verbal (IFS_IV).

5.2. Estadística y análisis de datos

En primer lugar, para comenzar a analizar la base de datos, se filtraron aquellos sujetos de los que no se poseía información sobre las HSB, ya que es una de las variables principales a analizar. Esto, resultó en una disminución en la cantidad de sujetos (cuadro 6.2).

Para analizar la distribución de las variables edad, años de educación e HSB, se realizó el test de Shapiro-Wilk, el cual, prueba la hipótesis nula de que los datos se extrajeron de una distribución normal. Al ejecutar este test, se obtienen el valor p que permite afirmar o negar si la distribución es normal. Valor p: el valor p mide la probabilidad de obtener los resultados observados asumiendo que la hipótesis nula es verdadera. Cuanto menor es el valor p, mayor es la evidencia de que se debería rechazar la hipótesis nula. Un valor menor a 0.05 es estadísticamente significativo.

Los resultados de dicho análisis se encuentran en el cuadro 6.1 en la sección de resultados.

Luego, se utilizó el test de OneWay-Anova para analizar la independencia de la variable edad para las tres clases. El objetivo fue lograr una base de datos en donde la edad sea independiente de la clase. Debido a esto, se acortó el rango de edad y se realizó el test de Anova recursivamente hasta lograr un *valor p* mayor a 0.05 que permitió aceptar la hipótesis nula, la cual, establece que no hay diferencia en la media de los grupos (*matchear* por edad). Dado que la clase con menor cantidad de sujetos corresponde a DFT, se priorizó conservar la mayor cantidad posible de dicha clase.

En la tabla 6.2 se encuentran las cantidades finales de sujetos que se utilizaron por clase.

Habiendo realizado el análisis de Anova, se procedió a analizar las métricas estadísticas sobre la edad, cuyos resultados se observan en el cuadro 6.3.

En cuanto al género, se analizó por medio del test de Chi-cuadrado. El mismo, sostiene una hipótesis nula de que no existe relación entre las clases y la variable categórica en cuestión. En la figura 6.3 se aprecia la distribución entre géneros para cada una de las clases.

Para analizar los años de educación, se utilizó el test de Kruskal-Wallis cuya hipótesis nula establece que los rangos medios de los grupos son iguales.

Al igual que para los años de educación, las HSB entre grupos se analizaron por medio de Kruskal-Wallis. Seguido de un análisis de Tukey-Kramer-HSD para comprender entre cuales dos grupos residen las diferencias.

A continuación, se procedió a analizar las correlaciones entre distintas variables y las HSB. Cabe aclarar que los valores de los exámenes cognitivos de la base de datos debieron ser completados con la media de la variable en aquellos casos que no se contaba esa información. Cuando se trabaja con datos faltantes hay dos posibilidades: eliminar los registros con datos faltantes o completarlos con suposiciones. Cuando el porcentaje de datos faltantes es bajo es conveniente completar los datos (una opción muy utilizada es con la media), en cambio si el porcentaje de datos faltantes es muy alto, los resultados no contarán con la variación natural propia de los

datos y resultará en un modelo erróneamente efectivo. Es por esto que, fue posible completar los datos faltantes por la media de cada variable ya que se contaba con un alto porcentaje de los datos completos [48].

En primer lugar, se estudió la correlación entre HSB y los tests cognitivos. En segundo lugar, la correlación entre HSB y la edad y, en tercer y último lugar, se analizó la correlación entre HSB y los años de educación. Para ello se contó con el test de Spearman-R, en donde se analiza la hipótesis nula de que el coeficiente de correlación de Spearman, ρ , es 0. Que dicho coeficiente sea 0 implica que los rangos de una variable no covarían con los rangos de la otra variable analizada.

Si el valor absoluto del coeficiente de correlación se encuentra entre 0.3 y 0.5, se puede afirmar que existe una correlación leve entre las variables estudiadas. Sin embargo, esta afirmación solo es válida si el valor p es estadísticamente significativo, es decir, menor a que 0.05. Por otro lado, si el valor absoluto es menor que 0.3, no se encuentra relación entre las variables estudiadas.

Por último, se graficaron las distribuciones de las HSB en cada una de las áreas mencionadas anteriormente. (Figuras 6.5, 6.6, 6.7)

Además, se realizó el análisis de Kruskal-Wallis para encontrar entre qué clases residen las diferencias significativas de las HSB para cada área. Dichos resultados se encuentran en el cuadro 6.6.

5.3. Machine Learning

5.3.1. Preprocesamiento

Los análisis y procedimientos explicados anteriormente fueron de utilidad para comprender la base de datos y definir el conjunto de sujetos que se utilizó para entrenar a los algoritmos.

Es por esto, que se utilizó la cantidad de sujetos especificados en el cuadro 6.2, los cuales fueron previamente filtrados para que la variable edad sea independiente de la clase a la que pertenezcan.

Además, para dichos sujetos, se completaron los valores faltantes con la media de la variable que se trate. Esto se pudo llevar a cabo ya que se contaba con un alto porcentaje de información en cada variable.

Finalmente se separaron los datos en grupos de a dos (EA-Control, EA-DFT y DFT-Control) para poder generar el conjunto de datos de *Train* y *Test* con una proporción de 80/20, dentro del cual *Train* se dividió en dos subgrupos: uno para entrenar los modelos y otro para validarlos.

5.3.2. Algoritmos

Para cada algoritmo mencionado en la sección *Algoritmos de Machine Learning* del *Marco Teórico*: SVM, RF, GB y XGB, se realizó el mismo procedimiento.

En primer lugar, se los entrenó con los parámetros por defecto de cada librería para obtener una primera aproximación de cada algoritmo. Para estabilizar los resultados obtenidos de cada algoritmo se comenzó a utilizar el *cross-validation* en la evaluación de cada uno.

A continuación, se realizó *grid-search* con la finalidad de encontrar la mejor combinación de parámetros para cada algoritmo. Este procedimiento se conoce como *hyperparameter tuning*.

Una vez obtenidos los mejores parámetros se avanzó en la búsqueda de la combinación óptima de *features* o variables de entrada para reducir el gasto computacional y evitar introducir variables que puedan interferir en la clasificación con impacto negativo.

Para la evaluación con los datos de test se iteró cinco mil veces cambiando el *random-state*, el cual corresponde a la semilla que se encarga de la separación de los datos, desde 42 e incrementándolo de a uno. Por último se procedió a promediar los cinco mil resultados para obtener métricas definitivas y estables.

El procedimiento detallado anteriormente se repitió para los siguientes tres casos con el fin de obtener una comparación y conclusión de los mismos:

1. Datos combinados: datos cognitivos + datos demográficos + datos neurorra-

diológicos

2. Datos neurorradiológicos
3. Datos cognitivos

Los datos cognitivos incluyen los datos provenientes de los tests psicológicos (MoCA, IFS y subtests del IFS). Los datos demográficos contienen la edad, género y años de educación. Finalmente los datos neurorradiológicos hacen referencia a los datos obtenidos a partir de imágenes de resonancia magnética (volúmenes de HSB normalizados por VIT).

Es decir, que en la sección de resultados se presentará el mejor algoritmo para cada uno de los 9 casos a analizar:

1. EA-Control: Datos combinados
2. EA-Control: Datos neurorradiológicos
3. EA-Control: Datos cognitivos
4. EA-DFT: Datos combinados
5. EA-DFT: Datos neurorradiológicos
6. EA-DFT: Datos cognitivos
7. DFT-Control: Datos combinados
8. DFT-Control: Datos neurorradiológicos
9. DFT-Control: Datos cognitivos

Sin embargo, al final del documento se encuentra el anexo con todos los resultados obtenidos, allí se puede apreciar el desempeño de cada algoritmo para las distintas situaciones planteadas.

Capítulo 6

Resultados

6.1. Estadística

Resultados del Test de Shapiro Wilk:

	Control	EA	DFT
	Valor-p	Valor-p	Valor-p
Edad	0.0909	0.0530	0.4807
Educación	0.0002	0.03367	0.0308
HSB	1.839e ⁻¹⁷	7.243e ⁻⁰⁷	1.905e ⁻⁰⁵

Cuadro 6.1: Valores del Test de Shapiro Wilk

De los resultados obtenidos en el cuadro 6.1, se observa que la edad presenta una distribución normal para las tres clases a analizar, mientras que la educación y las HSB normalizadas por VIT no presentan distribuciones normales para ninguna clase.

Por medio del test de Anova se obtuvo un *valor p* = 0.07, lo cual permitió aceptar la hipótesis nula y afirmar que las clases son independientes de la variable edad. Los rangos de edad seleccionados fueron los siguientes:

- Control: [62, 83]

- EA: [64, 81]
- DFT: [59, 83]

Con dichos rangos de edades, las cantidades finales de sujetos por clase que fueron utilizados son las siguientes:

	Cantidad original	Cantidad luego de filtrado sin valor	Cantidad luego de filtrado por edad
Control	105	96	69
EA	54	52	43
DFT	31	28	18

Cuadro 6.2: Cantidad de sujetos

En las figuras 6.1 y 6.2 se observa la distribución de las edades para cada clase antes y después de filtrar por edad.

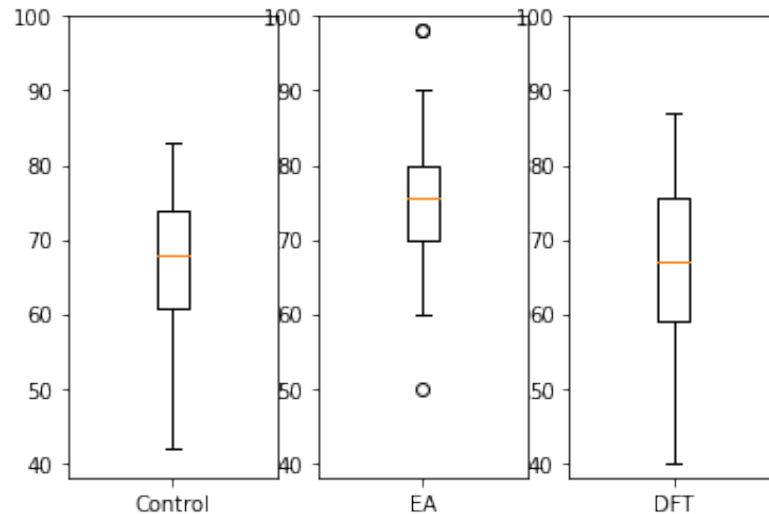


Figura 6.1: Boxplot de edades previo al filtrado

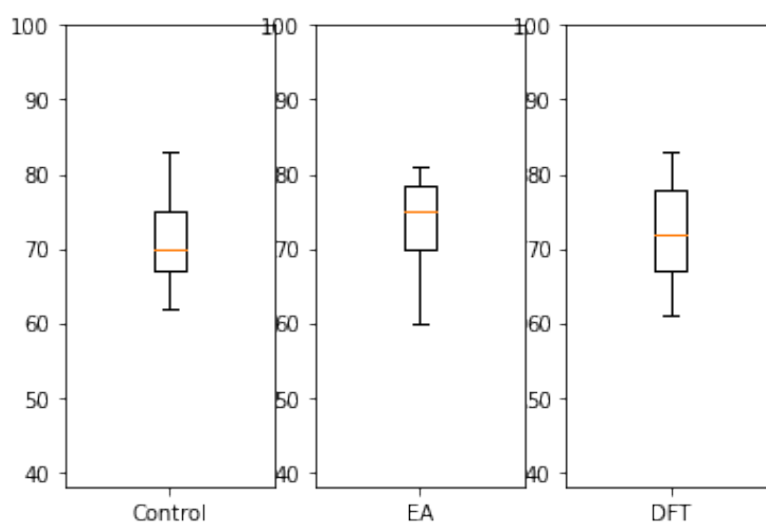


Figura 6.2: Boxplot de edades luego de filtrar

	Control	EA	DFT
Media	71.557	73.897	70.909
Mediana	72	75	70.5
Desvío estándar	5.394	4.442	8.201
Percentil 25	68	70	64.25
Percentil 75	76	77.5	77.75
Rango intercuartil	8	7.5	13.5

Cuadro 6.3: Estadísticas básicas de edad

En cuanto al género, se obtuvo un *valor-p* igual a 0.12 , por lo que se acepta la hipótesis nula del test de Chi-cuadrado que concluye que el género es independiente de la clase.

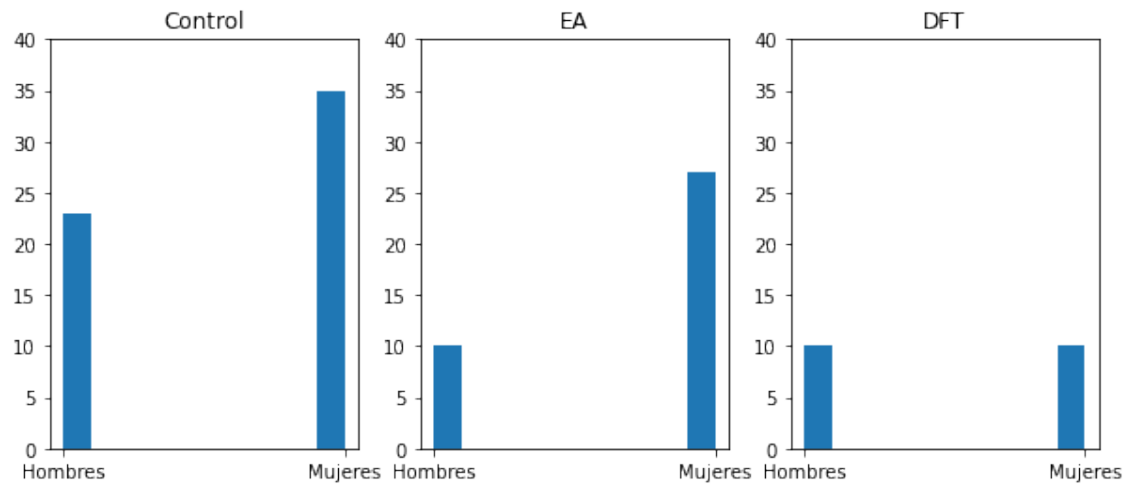


Figura 6.3: Histograma de la distribución por género para cada clase

A través del análisis de Kruskal-Wallis para estudiar los años de educación se obtiene un *valor-p* de $8.921e-05$. Es por esto que se rechaza la hipótesis nula y se concluye que los años de educación no son independientes de la clase.

El resultado obtenido del análisis de Kruskal-Wallis para las HSB fue de un *valor-p* igual a $3.673e-05$, por lo que también se rechaza la hipótesis nula. En la imagen 6.4 se puede observar la distribución de las HSB para cada clase.

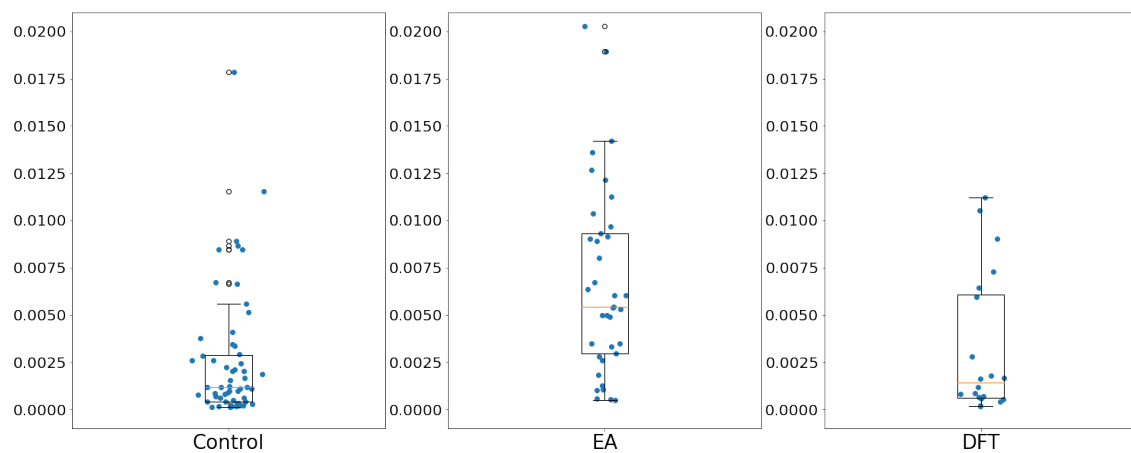


Figura 6.4: HSB según clase

Los resultados de la evaluación de Tukey-Kramer-HSD se encuentran en el cuadro 6.4. Por lo tanto, se concluye que existe una diferencia estadísticamente significativa entre las medias de los grupos de Control y EA y los grupos EA y DFT, pero no una

diferencia estadísticamente significativa entre las medias de los grupos de Control y DFT.

Clase 1	Clase 2	Valor-p	Rechazo
EA	Control	0.001	Verdadero
EA	DFT	0.0059	Verdadero
Control	DFT	0.7794	Falso

Cuadro 6.4: Resultados del test de Tukey-Kramer-HSD

		Correlación (ρ)	Valor-p
Control	Moca	-0.1143	0.3929
	IFS	-0.0772	0.5646
	Edad	0.4493	0.0004
	Educación	-0.0518	0.6988
EA	Moca	-0.1143	0.3929
	IFS	-0.3439	0.0372
	Edad	0.1469	0.3854
	Educación	-0.3635	0.0269
DFT	Moca	-0.2060	0.3835
	IFS	-0.3262	0.1605
	Edad	0.3310	0.1539
	Educación	-0.3568	0.1224

Cuadro 6.5: Resultados del test Spearman-R

Utilizando el test de Spearman-R, cuyos resultados son mostrados en el cuadro 6.5, se observó que la variable HSB se encuentra correlacionada con la edad en el grupo de Control y, en el grupo de EA se correlaciona con las variables de educación y IFS.

Finalmente se analizaron las 20 áreas del cerebro y se graficaron las distribuciones de las HSB. Observando las figuras 6.5, 6.6 y 6.7 se distingue que la distribución de HSB a través de los tractos seleccionados corresponde a un patrón similar en las tres clases. Sin embargo, la clase que mayor suma acumulada de HSB posee es EA mientras que la que menor suma acumulada de HSB posee es la clase de DFT.

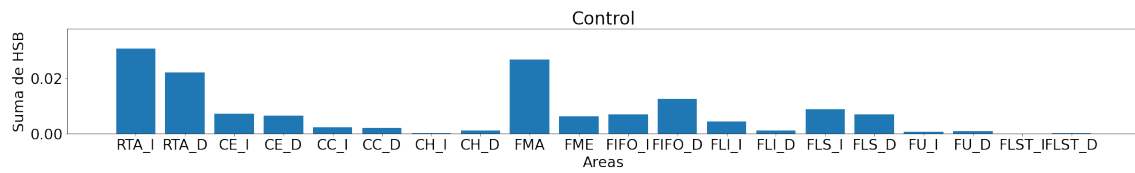


Figura 6.5: Distribución de las HSB distribuidas por área en la clase Control.

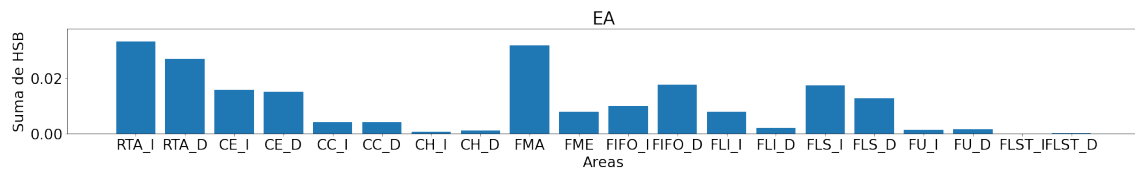


Figura 6.6: Distribución de las HSB distribuidas por área en la clase EA.

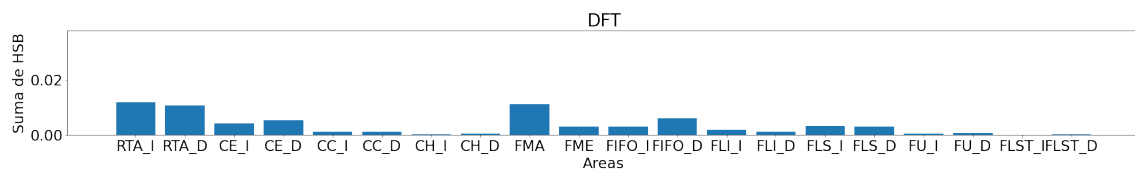


Figura 6.7: Distribución de las HSB distribuidas por área en la clase DFT.

A continuación se encuentran los resultados del test de Kruskal-Wallis analizando la dependencia de cada clase con las HSB de cada área. La V representa que hay diferencia significativa mientras que con una F se representa que no hay diferencia significativa.

	EA-Control	EA-DFT	DFT-Control
RTA_I	V	V	F
RTA_D	F	F	F
CE_I	V	V	F
CE_D	V	F	F
CC_I	V	V	F
CC_D	V	F	F
CH_I	V	V	F
CH_D	F	F	F
FMA	V	F	F
FME	V	F	F
FIFO_I	V	V	F
FIFO_D	V	F	F
FLI_I	V	V	F
FLI_D	F	F	F
FLS_I	V	V	F
FLS_D	V	V	F
FU_I	V	F	F
FU_D	V	F	F
FLST_I	F	F	F
FLST_D	F	F	F

Cuadro 6.6: Resultados del test Tukey-Kramer para cada área

6.2. EA-Control

6.2.1. Datos combinados: SVM

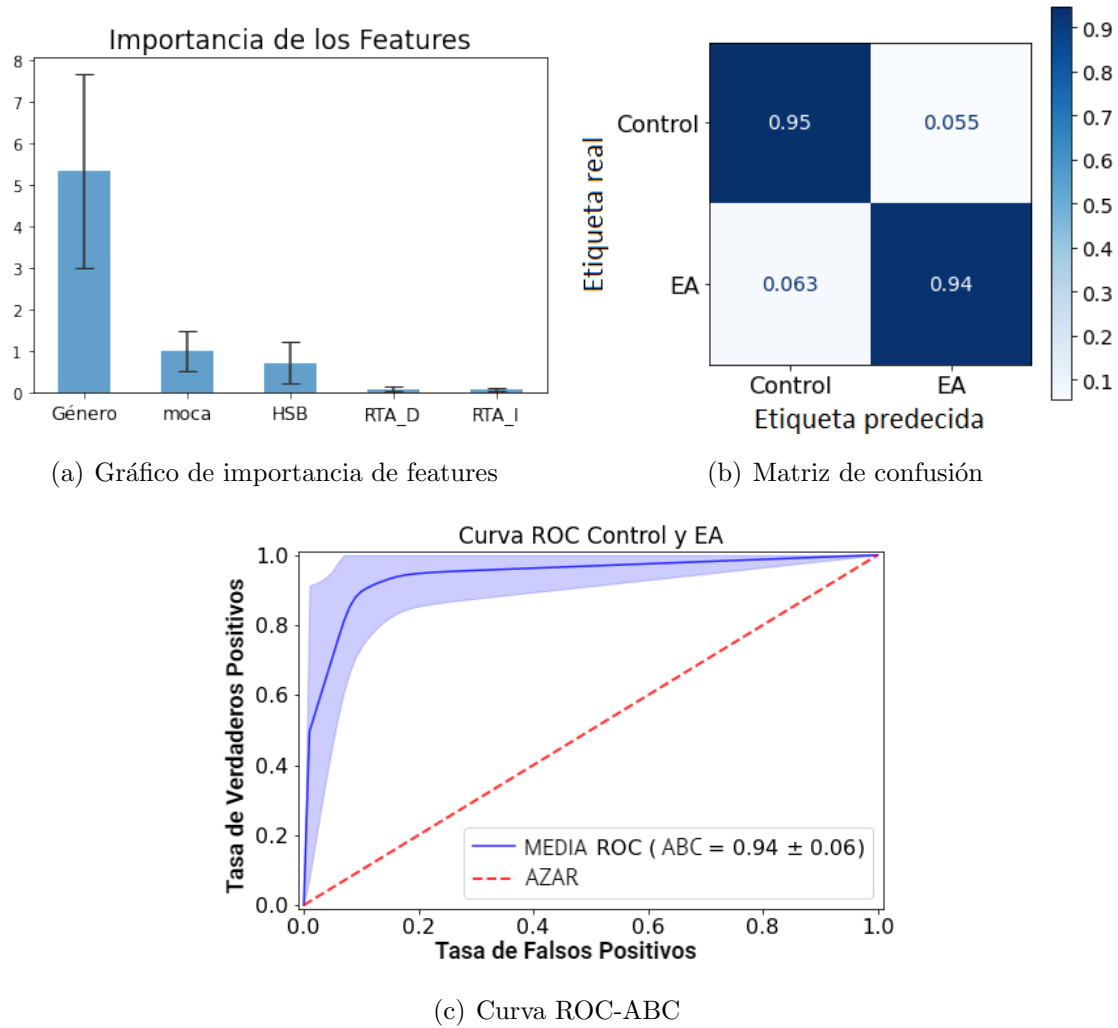


Figura 6.8: Resultados del algoritmo SVM para los grupos de EA y Control con datos combinados

6.2.2. Datos neurorradiológicos: RF

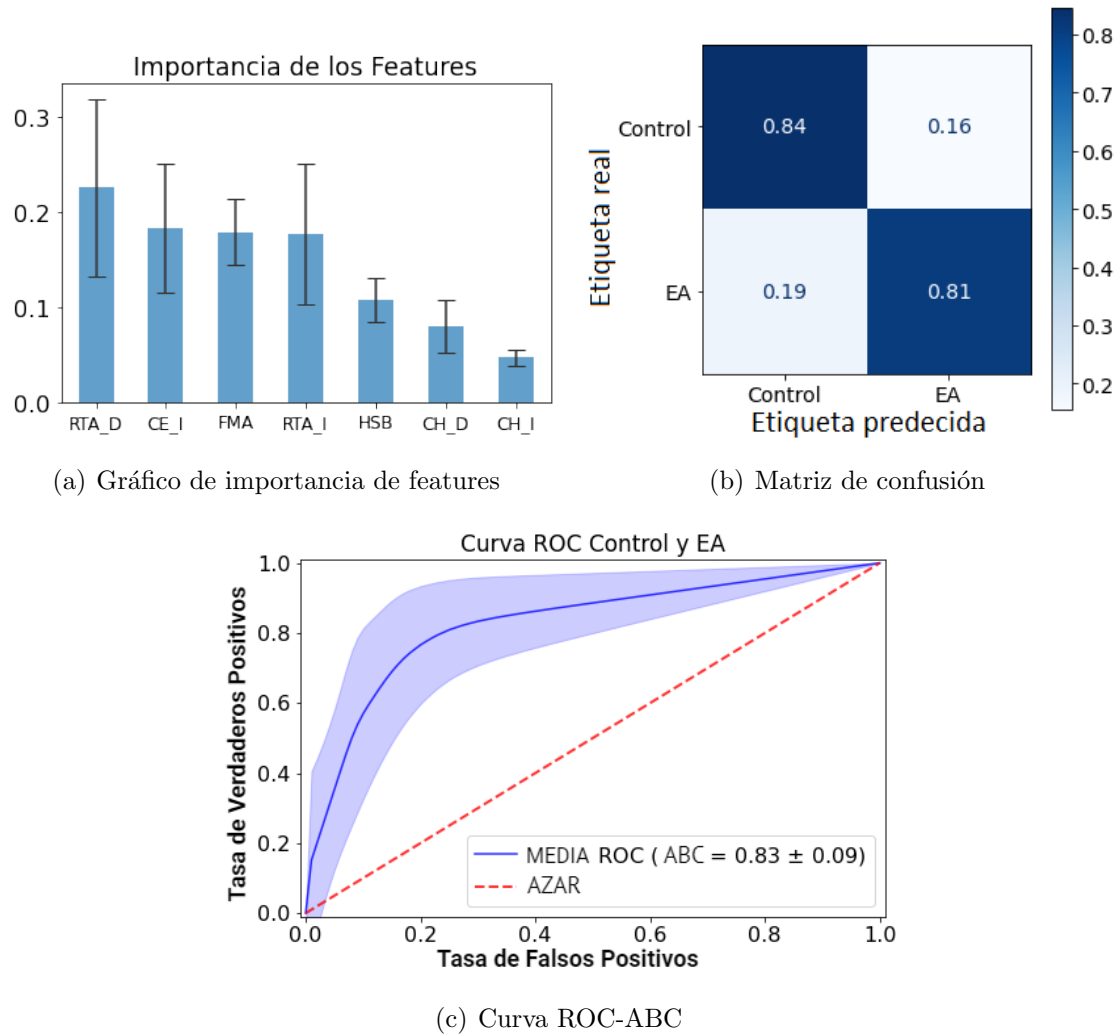


Figura 6.9: Resultados del algoritmo RF para los grupos de EA y Control con datos neurorradiológicos

6.2.3. Datos cognitivos: XGB

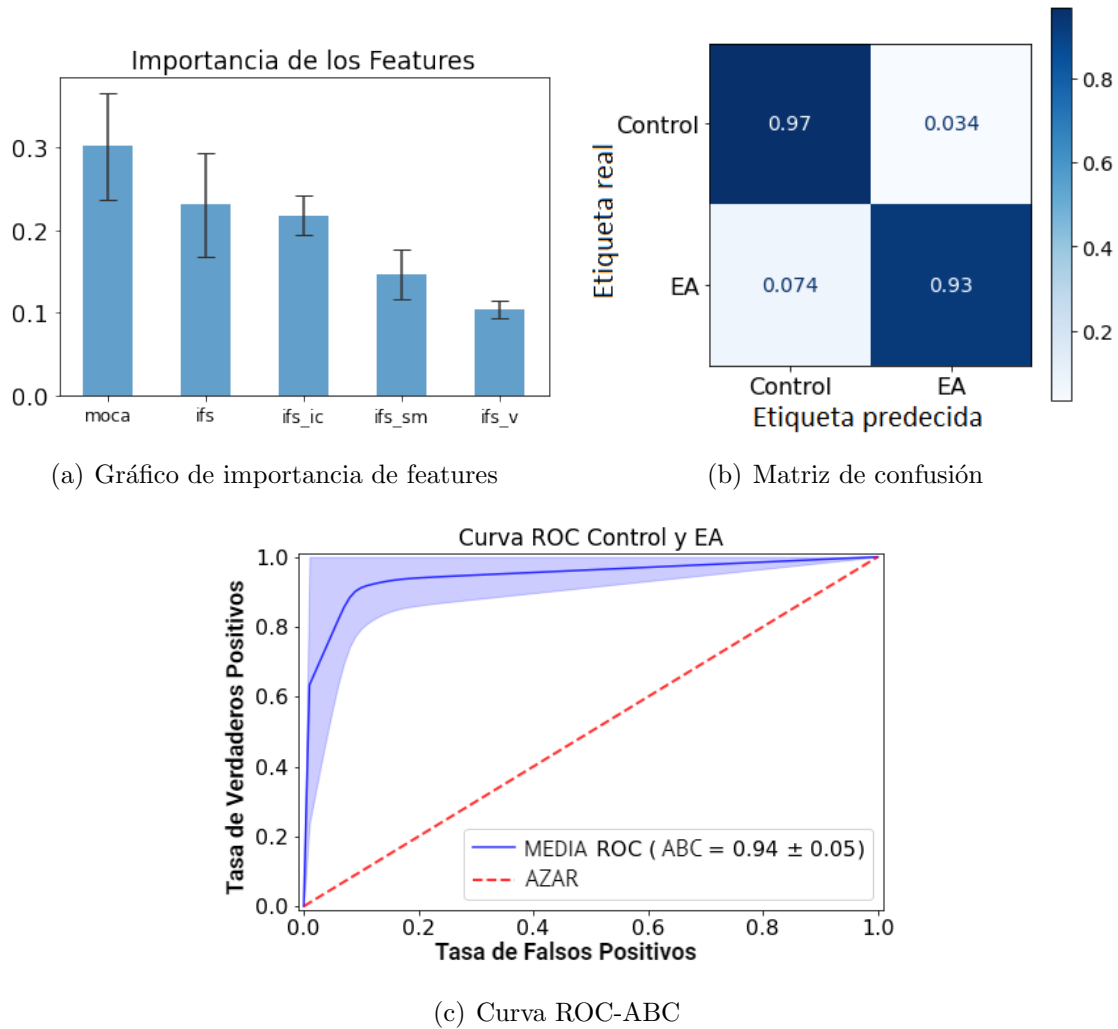


Figura 6.10: Resultados del algoritmo XGB para los grupos de EA y Control con datos cognitivos

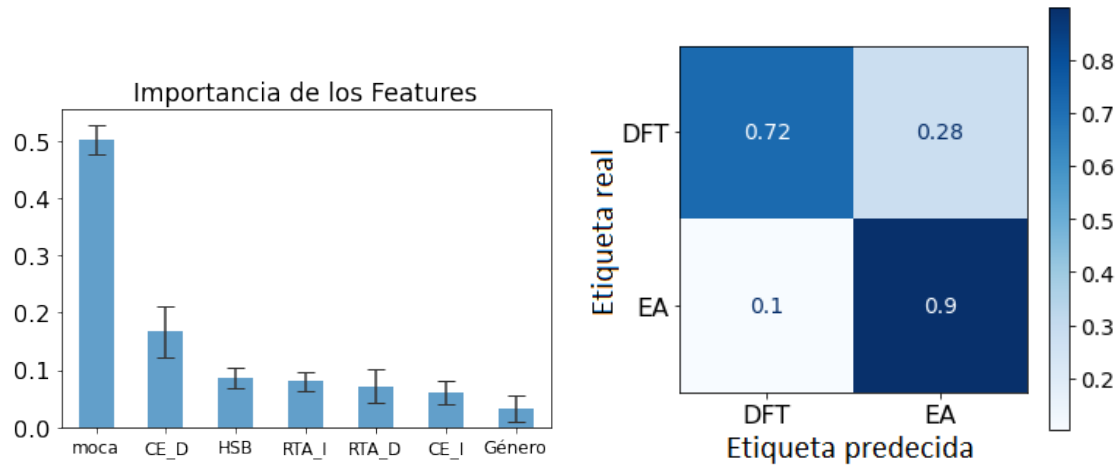
6.2.4. Métricas y comparación

EA-CONTROL				
	Datos combinados (Imágenes-test cognitivos-datos demográficos)			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.9221 (0.9200-0.9242)	0.9212 (0.9193-0.9231)	0.9239 (0.9220-0.9258)	0.9203 (0.9184-0.9223)
F1 Score Control (IC 95 %)	0.9504 (0.9491-0.9517)	0.9493 (0.9481-0.9505)	0.9512 (0.9500-0.9524)	0.9513 (0.9501-0.9525)
	Signos neurorradiológicos provenientes de imágenes de RM			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	-*	0.7796 (0.7764-0.7828)	0.6338 (0.6299-0.6378)	0.8829 (0.8803-0.8855)
F1 Score Control (IC 95 %)	-*	0.8547 (0.8526-0.8568)	0.7624 (0.7598-0.7650)	0.9344 (0.9331-0.9358)
	Test Cognitivos			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.9420 (0.9403-0.9437)	0.9859 (0.9851-0.9867)	0.9322 (0.9304-0.9341)	0.9322 (0.9304-0.9340)
F1 Score Control (IC 95 %)	0.8616 (0.9605-0.9626)	0.9910 (0.9905-0.9915)	0.9545 (0.9532-0.9557)	0.9576 (0.9865-0.9587)

*No se ha podido calcular el valor de la métrica F1 dado que el algoritmo predijo la clase mayoritaria en su totalidad.

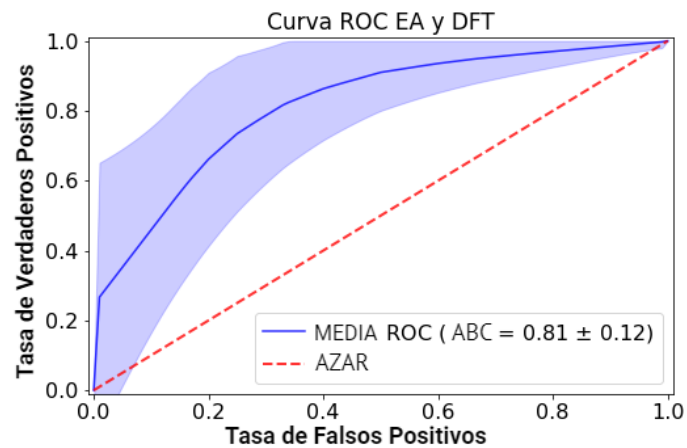
6.3. EA-DFT

6.3.1. Datos combinados: GB



(a) Gráfico de importancia de features

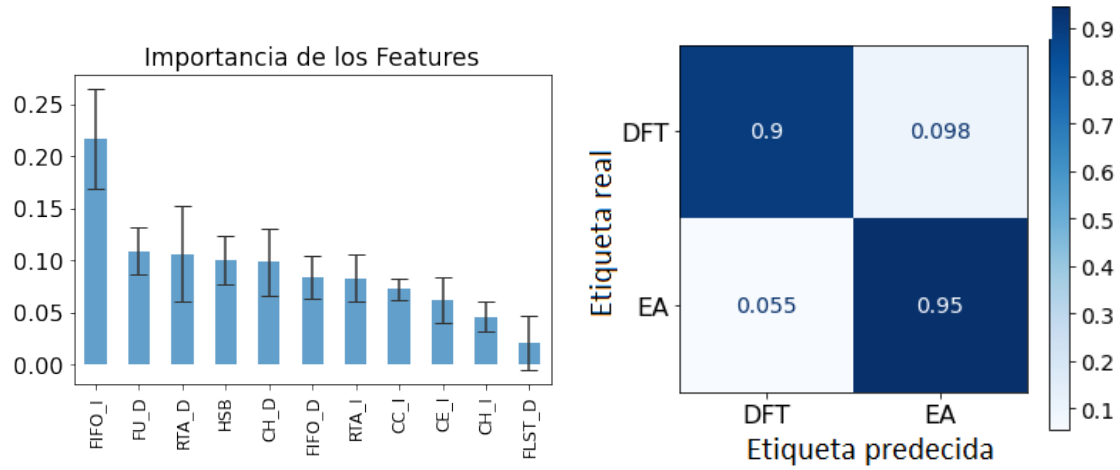
(b) Matriz de confusión



(c) Curva ROC-ABC

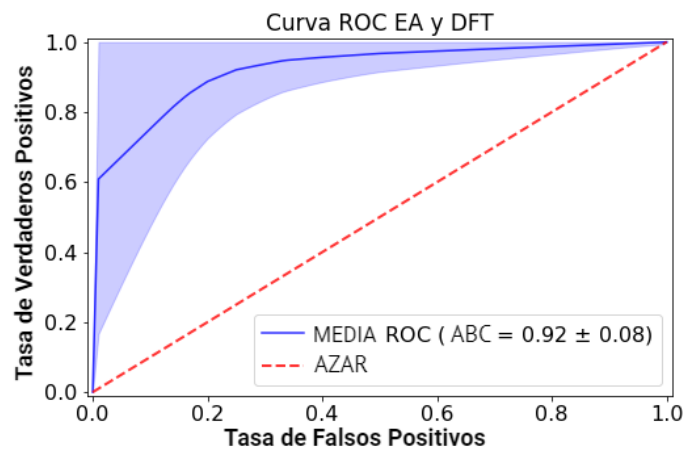
Figura 6.11: Resultados del algoritmo GB para los grupos de EA y DFT con datos combinados

6.3.2. Datos neurorradiológicos: GB



(a) Gráfico de importancia de features

(b) Matriz de confusión



(c) Curva ROC-ABC

Figura 6.12: Resultados del algoritmo GB para los grupos de EA y DFT con datos neurorradiológicos

6.3.3. Datos cognitivos: GB

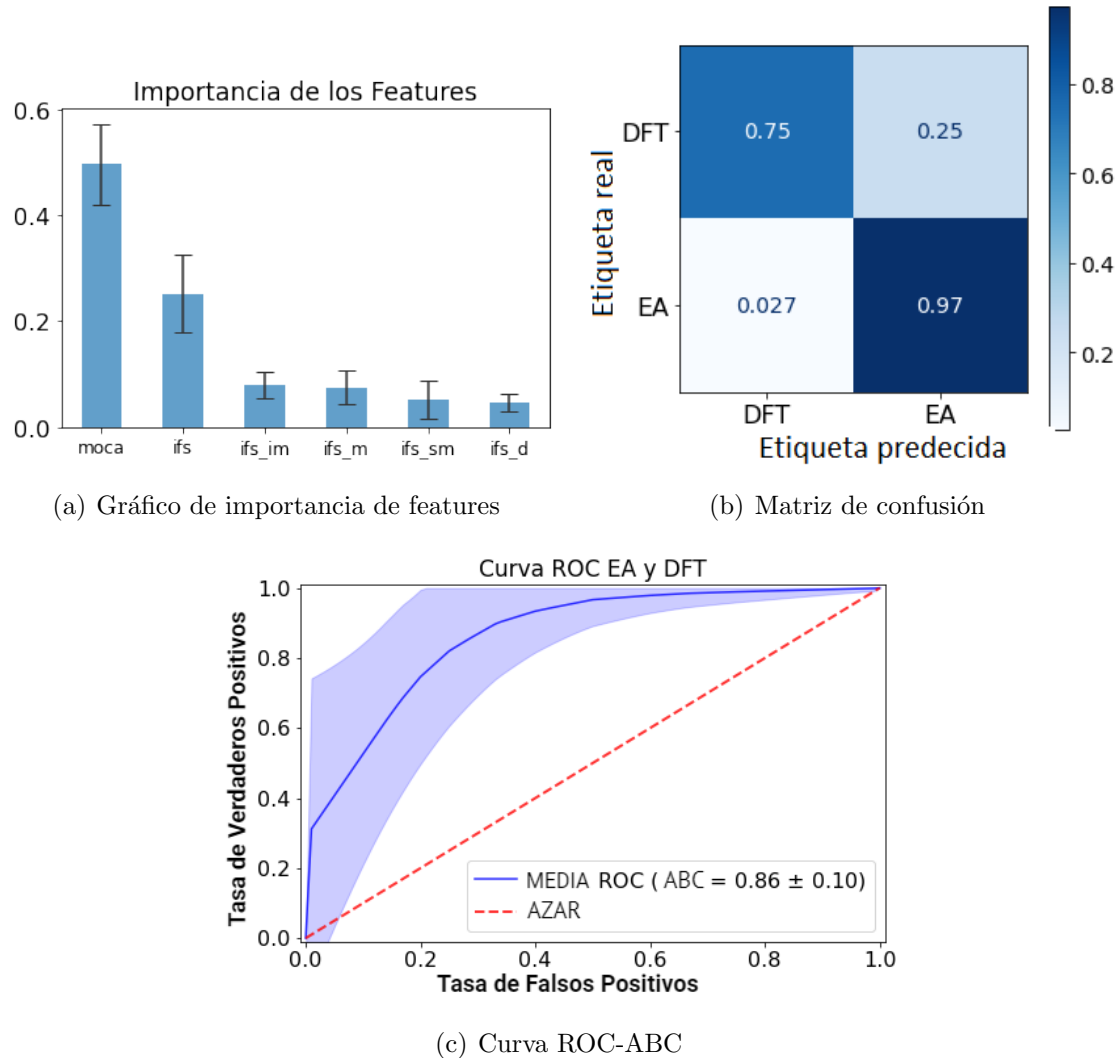


Figura 6.13: Resultados del algoritmo GB para los grupos de EA y DFT con datos cognitivos

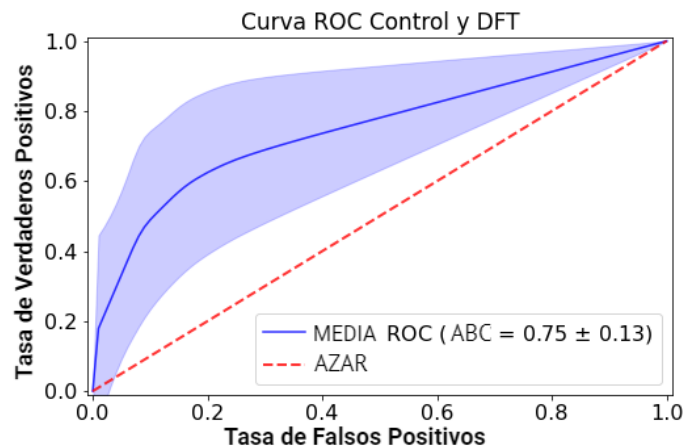
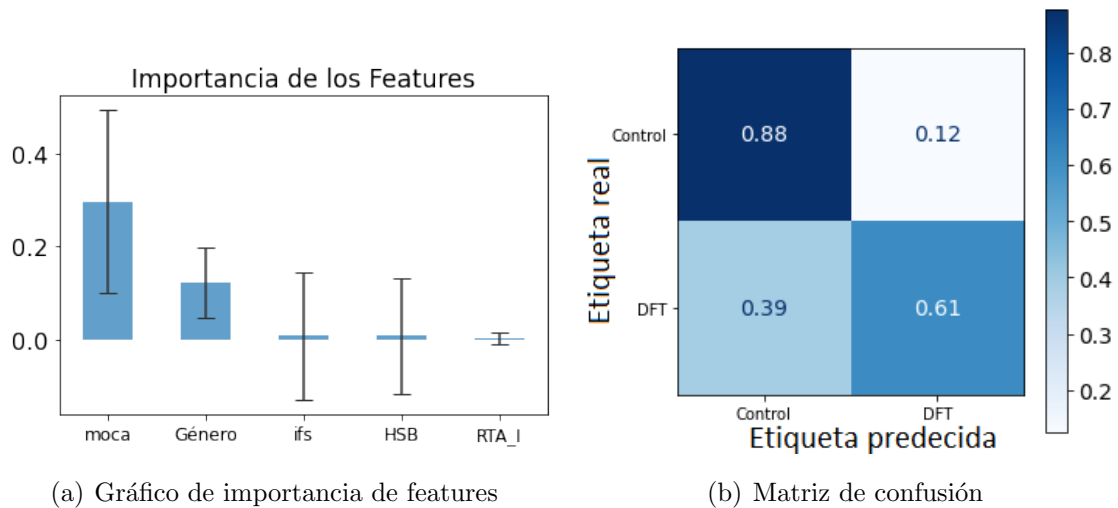
6.3.4. Métricas y comparación

EA-DFT				
	Datos combinados (Imágenes-test cognitivos-datos demográficos)			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.6718 (0.6668-0.6767)	0.7111 (0.7057-0.7165)	0.7361 (0.7313-0.7409)	0.7134 (0.7082-0.7186)
F1 Score DFT (IC 95 %)	0.8184 (0.8157-0.8211)	0.8733 (0.8708-0.8757)	0.8705 (0.8682-0.8728)	0.8810 (0.8788-0.8832)
	Signos neurorradiológicos provenientes de imágenes de RM			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	-*	0.6633 (0.6576-0.6689)	0.8910 (0.8877-0.8943)	0.5136 (0.5077-0.5195)
F1 Score DFT (IC 95 %)	-*	0.8533 (0.8509-0.8557)	0.9430 (0.9413-0.9446)	0.7927 (0.7900-0.7954)
	Test Cognitivos			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.7115 (0.7063-0.7167)	0.8153 (0.8112-0.8194)	0.8218 (0.8176-0.8261)	0.7322 (0.7271-0.7373)
F1 Score DFT (IC 95 %)	0.8583 (0.8558-0.8608)	0.9207 (0.9189-0.9226)	0.9191 (0.9172-0.9209)	0.8792 (0.8769-0.8815)

*No se ha podido calcular el valor de la métrica F1 dado que el algoritmo predijo la clase mayoritaria en su totalidad.

6.4. DFT-Control

6.4.1. Datos combinados: SVM



(c) Curva ROC-ABC

Figura 6.14: Resultados del algoritmo SVM para los grupos de DFT y Control con datos combinados

6.4.2. Datos neurorradiológicos: RF

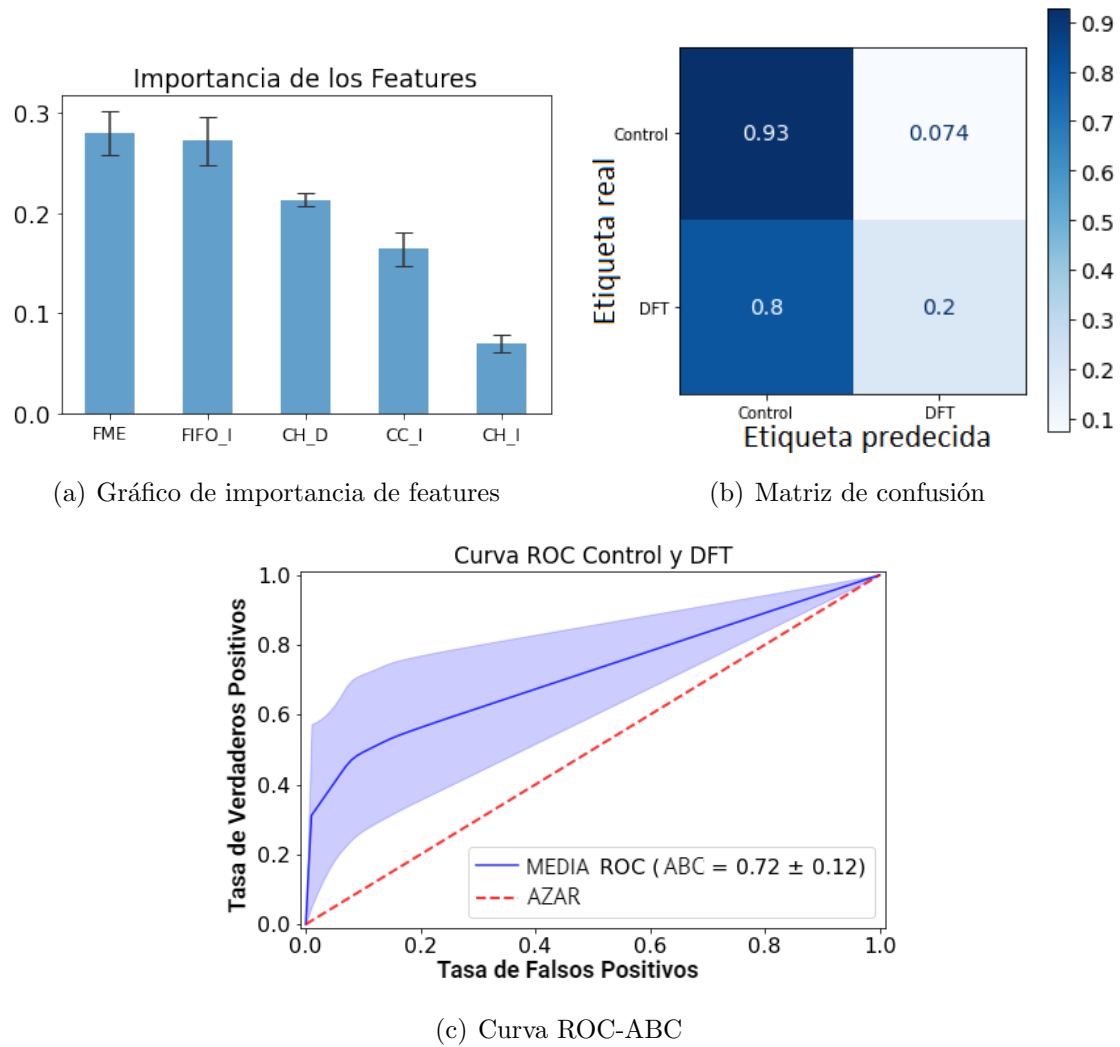


Figura 6.15: Resultados del algoritmo RF para los grupos de DFT y Control con datos neurorradiológicos

6.4.3. Datos cognitivos: RF

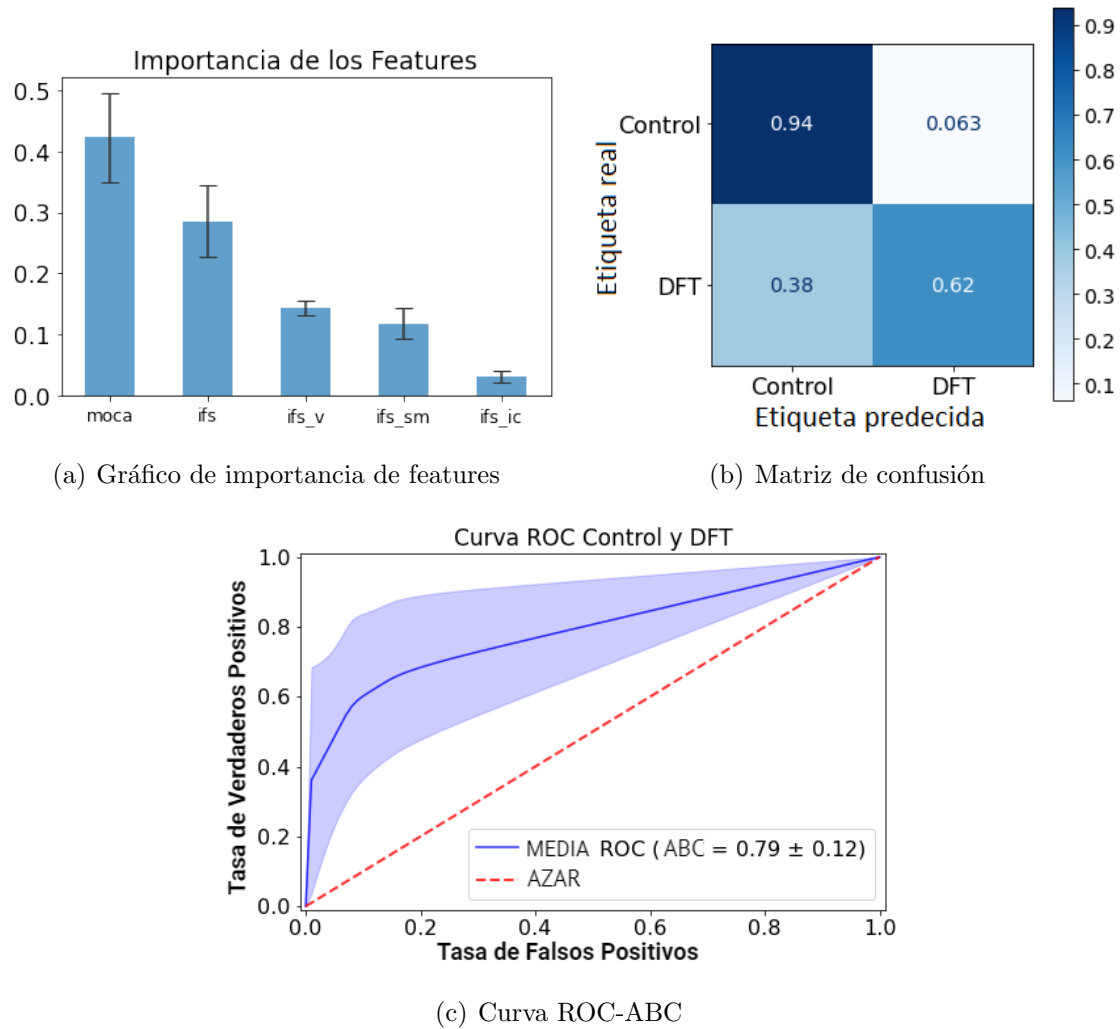


Figura 6.16: Resultados del algoritmo RF para los grupos de DFT y Control con datos cognitivos

6.4.4. Métricas y comparación

DFT-CONTROL				
	Datos combinados (Imágenes-test cognitivos-datos demográficos)			
	SVM	RF	GB	XGB
F1 Score DFT (IC 95 %)	0.5991 (0.5932-0.6050)	0.5381 (0.5318-0.5445)	0.5586 (0.5527-0.5645)	0.6011 (0.5952-0.6070)
F1 Score Control (IC 95 %)	0.8657 (0.8637-0.8678)	0.8834 (0.8815-0.8852)	0.8764 (0.8746-0.8782)	0.8918 (0.8901-0.8935)
	Signos neurorradiológicos provenientes de imágenes de RM			
	SVM	RF	GB	XGB
F1 Score DFT (IC 95 %)	-*	0.2735 (0.2665-0.2804)	0.0650 (0.0565-0.0736)	0.1617 (0.1554-0.1679)
F1 Score Control (IC 95 %)	-*	0.8372 (0.8350-0.8394)	0.8316 (0.8282-0.8350)	0.8086 (0.8063-0.8110)
	Test Cognitivos			
	SVM	RF	GB	XGB
F1 Score DFT (IC 95 %)	0.6476 (0.6422-0.6531)	0.6749 (0.6694-0.6804)	0.6060 (0.6005-0.6115)	0.6313 (0.6258-0.6367)
F1 Score Control (IC 95 %)	0.8960 (0.8943-0.8977)	0.9026 (0.9009-0.9043)	0.8774 (0.8756-0.8793)	0.8869 (0.8851-0.8886)

*No se ha podido calcular el valor de la métrica F1 dado que el algoritmo predijo la clase mayoritaria en su totalidad.

6.5. Métricas y comparación general

EA-CONTROL				
	Datos combinados (Imágenes-test cognitivos-datos demográficos)			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.9221 (0.9200-0.9242)	0.9212 (0.9193-0.9231)	0.9239 (0.9220-0.9258)	0.9203 (0.9184-0.9223)
F1 Score Control (IC 95 %)	0.9504 (0.9491-0.9517)	0.9493 (0.9481-0.9505)	0.9512 (0.9500-0.9524)	0.9513 (0.9501-0.9525)
	Signos neurorradiológicos provenientes de imágenes de RM			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	-*	0.7796 (0.7764-0.7828)	0.6338 (0.6299-0.6378)	0.8829 (0.8803-0.8855)
F1 Score Control (IC 95 %)	-*	0.8547 (0.8526-0.8568)	0.7624 (0.7598-0.7650)	0.9344 (0.9331-0.9358)
	Test Cognitivos			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.9420 (0.9403-0.9437)	0.9859 (0.9851-0.9867)	0.9322 (0.9304-0.9341)	0.9322 (0.9304-0.9340)
F1 Score Control (IC 95 %)	0.8616 (0.9605-0.9626)	0.9910 (0.9905-0.9915)	0.9545 (0.9532-0.9557)	0.9576 (0.9865-0.9587)
EA-DFT				
	Datos combinados (Imágenes-test cognitivos-datos demográficos)			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.6718 (0.6668-0.6767)	0.7111 (0.7057-0.7165)	0.7361 (0.7313-0.7409)	0.7134 (0.7082-0.7186)
F1 Score DFT (IC 95 %)	0.8184 (0.8157-0.8211)	0.8733 (0.8708-0.8757)	0.8705 (0.8682-0.8728)	0.8810 (0.8788-0.8832)
	Signos neurorradiológicos provenientes de imágenes de RM			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	-*	0.6633 (0.6576-0.6689)	0.8910 (0.8877-0.8943)	0.5136 (0.5077-0.5195)
F1 Score DFT (IC 95 %)	-*	0.8533 (0.8509-0.8557)	0.9430 (0.9413-0.9446)	0.7927 (0.7900-0.7954)
	Test Cognitivos			
	SVM	RF	GB	XGB
F1 Score EA (IC 95 %)	0.7115 (0.7063-0.7167)	0.8153 (0.8112-0.8194)	0.8218 (0.8176-0.8261)	0.7322 (0.7271-0.7373)
F1 Score DFT (IC 95 %)	0.8583 (0.8558-0.8608)	0.9207 (0.9189-0.9226)	0.9191 (0.9172-0.9209)	0.8792 (0.8769-0.8815)
DFT-CONTROL				
	Datos combinados (Imágenes-test cognitivos-datos demográficos)			
	SVM	RF	GB	XGB
F1 Score DFT (IC 95 %)	0.5991 (0.5932-0.6050)	0.5381 (0.5318-0.5445)	0.5586 (0.5527-0.5645)	0.6011 (0.5952-0.6070)
F1 Score Control (IC 95 %)	0.8657 (0.8637-0.8678)	0.8834 (0.8815-0.8852)	0.8764 (0.8746-0.8782)	0.8918 (0.8901-0.8935)
	Signos neurorradiológicos provenientes de imágenes de RM			
	SVM	RF	GB	XGB
F1 Score DFT (IC 95 %)	-*	0.2735 (0.2665-0.2804)	0.0650 (0.0565-0.0736)	0.1617 (0.1554-0.1679)
F1 Score Control (IC 95 %)	-*	0.8372 (0.8350-0.8394)	0.8316 (0.8282-0.8350)	0.8086 (0.8063-0.8110)
	Test Cognitivos			
	SVM	RF	GB	XGB
F1 Score DFT (IC 95 %)	0.6476 (0.6422-0.6531)	0.6749 (0.6694-0.6804)	0.6060 (0.6005-0.6115)	0.6313 (0.6258-0.6367)
F1 Score Control (IC 95 %)	0.8960 (0.8943-0.8977)	0.9026 (0.9009-0.9043)	0.8774 (0.8756-0.8793)	0.8869 (0.8851-0.8886)

*No se ha podido calcular el valor de la métrica F1 dado que el algoritmo predijo la clase mayoritaria en su totalidad.

Capítulo 7

Conclusión y discusión

El objetivo general de este trabajo fue la comparación entre cuatro algoritmos de ML en cuanto a su efectividad en la clasificación diferencial de las clases: EA, DFT y Control. Dichas clases fueron agrupadas de a pares y, a partir de la base de datos, se dividió la información en tres categorías: datos demográficos, datos neurorradiológicos y datos cognitivos. De esta manera, para cada clase se realizó la clasificación teniendo distinta información de entrada. Los tres escenarios planteados fueron: estudiar con datos combinados (consiste de datos demográficos, datos neurorradiológicos y datos cognitivos), estudiar con datos neurorradiológicos y en tercer lugar, estudiar con datos cognitivos. El fin de esta comparación fue seleccionar el mejor algoritmo según los valores de la métrica F1, la matriz de confusión y la curva ROC.

A partir de esta comparación, se han mostrado los mejores resultados que se han obtenido en los escenarios mencionados anteriormente. El resultado más relevante y novedoso obtenido fue el modelo GB utilizando datos provenientes únicamente de imágenes de RM (datos neurorradiológicos) en la clasificación de EA-DFT.

En conjunto con esta comparación, se realizó previamente un análisis estadístico tradicional para analizar las diferencias presentes en la base de datos. A partir de este análisis fue posible concluir los siguientes resultados.

Analizando las distribuciones de las HSB para cada clase, cabe notar que EA posee un incremento de la HSB en comparación con las dos clases restantes. Esto podría deberse en parte, a una mayor proporción de patología mixta en sujetos con

EA.

Ahora bien, teniendo en cuenta los resultados del test de Spearman-R, se encontró una correlación positiva de la HSB con la edad únicamente para la clase de Control. Dicho resultado coincide con la teoría y revisión bibliográfica ya que en pacientes de Control la única causa de las HSB es la carga vascular que aumenta conforme aumenta la edad. Por el contrario, la falta de correlación entre HSB y la edad para las clases patológicas puede haber sido enmascarada por la presencia prematura o previa de HSB a raíz de la degeneración patológica.

Es posible concluir que se ha cumplido el objetivo en su totalidad para las clases: EA-Control y EA-DFT. Sin embargo, no se han podido obtener resultados favorables en la clasificación DFT-Control.

Esto, probablemente, este relacionado a dos factores a tener en cuenta.

El primero de los mismos es el hecho de que las muestras para cada condición se encuentran desbalanceadas provocando que, los algoritmos tengan una tendencia a clasificar hacia la etiqueta con mayor cantidad de muestras. Si bien la diferencia entre cantidad de muestras estaba presente en todas las clases, la desproporción fue más significativa en la comparación DFT-Control ($n= 18$ vs 69) que en las comparaciones EA-DFT ($n= 43$ vs 18) y EA-Control ($n= 43$ vs 69).

El segundo factor a tener en cuenta en el caso de las clasificaciones donde se utilizan los signos neurorradiológicos es que, entre los grupos estudiados, no se han encontrado diferencias significativas en las mediciones de HSB tanto para el caso de volumen total como para las mediciones por tractos en el caso de la comparación DFT-Control. Esto, se puede apreciar en la tabla 6.6. Este cuadro muestra, además, que en el caso de la clasificación EA-Control la mayoría de los tractos poseen diferencias significativas, en tanto, en la clasificación EA-DFT se han encontrado diferencias significativas en 8 de los 20 tractos disponibles. Esto, probablemente este relacionado y explique porqué los mejores resultados se han obtenido en la clasificación EA-Control y EA-DFT.

Además, se puede observar que en los casos donde la clasificación fue exitosa, se han utilizado algunos de los siguientes tractos mencionados anteriormente en reiteradas ocasiones utilizando distintos modelos: FU, RTA, CC, FIFO, FLS, etc. De manera notable, los tractos empleados en las clasificaciones exitosas, han mostrado estar

implicados con las patologías de interés según las siguientes referencias:

En la variante conductual de DFT, diferentes estudios han demostrado alteraciones en la mayoría de los tractos de materia blanca frontales utilizando RM con modalidad DTI (imágenes con tensor de difusión), incluyendo los tractos: CC, FLS, FU, FLI, FIFO. (Agosta et al., 2012b; Avants et al., 2010; Borroni et al., 2007; Matsuo et al., 2008; Whitwell et al., 2010a; Zhang et al., 2009, 2013). [51]

Los resultados del estudio realizado por Mario Torso y su equipo, a partir de imágenes FLAIR ponderadas en T2 demuestran que el daño provocado al tracto RTA es estratégico para las ocurrencias de apatía en pacientes con deterioro cognitivo leve que, como se ha mencionado anteriormente, es uno de los síntomas de la EA. [52]

Varios estudios han sido capaces de replicar la relación entre los cambios microestructurales en el tracto CC y la EA por medio de imágenes de RM con modalidad DTI. (Bozzali et al., 2012; Choo et al., 2010; Metzler-Baddeley et al., 2012a; Remy et al., 2015; Wang et al., 2017; Zhang et al., 2007; Zhuang et al., 2012; 2013, for review see Chua et al., 2008; Yu et al., 2017). [53]

Además, cabe notar que para los casos en donde solo se tuvieron en cuenta los datos neurorradiológicos el algoritmo de SVM no fue satisfactorio. Esto se debió a que en los tres escenarios planteados, el algoritmo predijo la clase mayoritaria en su totalidad (Control en dos casos y EA para el caso de EA-DFT). Es por esto que no se pudo extraer información pertinente ni calcular la métrica de F1. En las tablas de resultados se dejó el espacio en blanco.

Finalmente es posible concluir que, si bien los test cognitivos han sido capaces de realizar una buena clasificación en la clase de mayor interés, EA-DFT, el hallazgo más pertinente del estudio reside en que el mejor resultado se ha encontrado en el caso de utilizar el modelo GB con solamente signos neurorradiológicos provenientes de imágenes de RM como datos de entrada. Este es un hallazgo novedoso que podría resultar en un nuevo aporte/herramienta para el diagnóstico diferencial de la EA y la DFT. Además dicho resultado es una prueba adicional de que la medicina avanza hacia una digitalización masiva, mejorando la calidad y velocidad del cuidado de la salud.

A pesar de que las métricas evaluadas y presentadas fueron altas y es posible categorizar a los algoritmos como buenos clasificadores, éstas podrían ser mejoradas.

Una forma de lograr esto, sería utilizando una base de datos mucho mayor cuyas tres clases estén balanceadas para lograr un mejor entrenamiento y mejores resultados consecuentemente. Otra implementación a futuro, sería entrenar y testear al algoritmo con otras modalidades de RM, como DTI, para que otras características fueran tenidas en cuenta (por ejemplo, pérdida de sustancia blanca). Además, los análisis futuros se podrían complementar con datos no provenientes de imágenes, como ser biomarcadores fluidos de la sangre o del fluido cerebroespinal. Por último, el algoritmo propuesto podría ser de utilidad para la asistencia al diagnóstico de otras patologías, entre ellas la enfermedad de Parkinson.

Capítulo 8

Tutores

Los tutores escogidos para brindar asistencia y soporte fueron los siguientes:

Cecilia Gonzalez Campo es médica especialista en neuroimágenes e investigadora del Conicet sobre el tema “estudio de las interacciones conductuales, estructurales y funcionales en enfermedades neurodegenerativas: hacia una mejor precisión en el diagnóstico precoz” en el departamento de Matemáticas y Ciencias de la Universidad de San Andrés [49]. Contacto: cegcampo@gmail.com

Sebastián Moguilner, por su parte, es licenciado en Ciencias Físicas y Máster en Física Médica graduado del Instituto Balseiro. Recientemente ha finalizado su Doctorado en Biología, en el que desarrolló nuevas medidas para evaluar la conectividad funcional cerebral con el fin de utilizarlas como variables de entrada en modelos de machine-learning. [50]. Contacto sebagam@gmail.com

Bibliografía

- [1] M. FILIPPI, IMAGING DEMENTIA. [S.l.]: SPRINGER NATURE, 2022.
- [2] “ADI - Dementia statistics”, Alzint.org, 2022. [Online]. Disponible en: <https://www.alzint.org/about/dementia-facts-figures/dementia-statistics/>.
- [3] A. International, M. Guerchet, M. Prince and M. Prina, “Numbers of people with dementia worldwide: An update to the estimates in the World Alzheimer Report 2015”, Alzint.org, 2022. [Online]. Disponible en: <https://www.alzint.org/resource/numbers-of-people-with-dementia-worldwide/>.
- [4] “ADI - Alois Alzheimer”, Alzint.org, 2022. [Online]. Disponible en: <https://www.alzint.org/about/dementia-facts-figures/types-of-dementia/alzheimers-disease/alois-alzheimer/>.
- [5] C. Onyike and J. Diehl-Schmid, “The epidemiology of frontotemporal dementia”, International Review of Psychiatry, vol. 25, no. 2, pp. 130-137, 2013. Disponible en: 10.3109/09540261.2013.776523.
- [6] FRONTOTEMPORAL DEMENTIAS. [S.l.]: SPRINGER, 2022.
- [7] K. Rascovsky et al., “Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia”, Brain, vol. 134, no. 9, pp. 2456-2477, 2011. Disponible en: 10.1093/brain/awr179.
- [8] E. Roberson, Alzheimer’s disease and frontotemporal dementia. New York: Humana Press, 2011.

- [9] M. Brown and R. Semelka, MRI, 3rd ed. Hoboken, N.J.: Wiley-Blackwell/John Wiley Sons, 2010.
- [10] “MRI Basics”, Case.edu, 2022. [Online]. Disponible en: <https://case.edu/med/neurology/NR/MRI%20Basics.htm..>
- [11] M. Niknejad, “Fluid attenuated inversion recovery — Radiology Reference Article”, [Radiopaedia.org](https://radiopaedia.org), 2022. [Online]. Disponible en: <https://radiopaedia.org/articles/fluid-attenuated-inversion-recovery>.
- [12] N. Prins and P. Scheltens, “White matter hyperintensities, cognitive impairment and dementia: an update”, *Nature Reviews Neurology*, vol. 11, no. 3, pp. 157-165, 2015. Disponible en: 10.1038/nrneurol.2015.10.
- [13] J. Alber et al., “White matter hyperintensities in vascular contributions to cognitive impairment and dementia (VCID): Knowledge gaps and opportunities”, *Alzheimer’s amp; Dementia: Translational Research amp; Clinical Interventions*, vol. 5, no. 1, pp. 107-117, 2019. Disponible en: 10.1016/j.trci.2019.02.001.
- [14] A. Kapasi, C. DeCarli and J. Schneider, “Impact of multiple pathologies on the threshold for clinically overt dementia”, *Acta Neuropathologica*, vol. 134, no. 2, pp. 171-186, 2017. Disponible en: 10.1007/s00401-017-1717-7.
- [15] Z. Nasreddine et al., “The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment”, *Journal of the American Geriatrics Society*, vol. 53, no. 4, pp. 695-699, 2005. Disponible en: 10.1111/j.1532-5415.2005.53221.x.
- [16] T. TORRALVA, M. ROCA, E. GLEICHGERRCHT, P. LÓPEZ and F. MANES, “INECO Frontal Screening (IFS): A brief, sensitive, and specific tool to assess executive functions in dementia—CORRECTED VERSION”, *Journal of the International Neuropsychological Society*, vol. 15, no. 5, pp. 777-786, 2009. Disponible en: 10.1017/s1355617709990415.
- [17] Z. Nasreddine et al., “The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment”, *Journal of the American Ge-*

- riatrics Society, vol. 53, no. 4, pp. 695-699, 2005. Disponible en: 10.1111/j.1532-5415.2005.53221.x.
- [18] J. Herron, Machine Learning: The Ultimate Guide for Beginners to Programming and Deep Learning With Python.. 2022.
- [19] M. Coding, Machine Learning with Python: A Step by Step Guide for Absolute Beginners to Program Artificial Intelligence with Python. Neural Networks and Data Science from Pre-Processing to Deep Learning. 2019.
- [20] “The Essential Guide to Quality Training Data for Machine Learning”, [Cloudfactory.com](https://www.cloudfactory.com/training-data-guide), 2022. [Online]. Disponible en: <https://www.cloudfactory.com/training-data-guide>.
- [21] “What is data splitting and why is it important?”, SearchEnterpriseAI, 2022. [Online]. Disponible en: <https://www.techtarget.com/searchenterpriseai/definition/data-splitting#:~:text=In%20machine%20learning%2C%20data%20splitting,into%20three%20or%20four%20sets>.
- [22] “Train Test Validation Split: How To & Best Practices [2022]”, [V7labs.com](https://www.v7labs.com/blog/train-validation-test-set#h2), 2022. [Online]. Disponible en: <https://www.v7labs.com/blog/train-validation-test-set#h2>.
- [23] “Hyperparameter tuning for machine learning models.”, Jeremy Jordan, 2022. [Online]. Disponible en: <https://www.jeremyjordan.me/hyperparameter-tuning/>.
- [24] “Hyperparameter tuning - GeeksforGeeks”, GeeksforGeeks, 2022. [Online]. Disponible en: <https://www.geeksforgeeks.org/hyperparameter-tuning/>.
- [25] “Hyperparameter Tuning in Python: a Complete Guide”, neptune.ai, 2022. [Online]. Disponible en: <https://neptune.ai/blog/hyperparameter-tuning-in-python-complete-guide>.
- [26] “3.2. Tuning the hyper-parameters of an estimator”, scikit-learn, 2022. [Online]. Disponible en: https://scikit-learn.org/stable/modules/grid_search.html.

- [27] “Python: qué es y por qué deberías aprender a utilizarlo”, [Becas-santander.com](https://www.becas-santander.com), 2022. [Online]. Disponible en: <https://www.becas-santander.com/es/blog/python-que-es.html>.
- [28] “NumPy user guide — NumPy v1.23 Manual”, [Numpy.org](https://numpy.org), 2022. [Online]. Disponible en: <https://numpy.org/doc/stable/user/index.html#user>.
- [29] “User Guide — pandas 1.4.3 documentation”, [Pandas.pydata.org](https://pandas.pydata.org), 2022. [Online]. Disponible en: https://pandas.pydata.org/docs/user_guide/index.html#user-guide.
- [30] “Scikit-Lear User guide: contents”, [scikit-learn](https://scikit-learn.org), 2022. [Online]. Disponible en: https://scikit-learn.org/stable/user_guide.html.
- [31] “Scikit-Learn, Herramienta básica para el Data Science en Python”, 2022. [Online]. Disponible en: <https://www.master-data-scientist.com/scikit-learn-data-science/>.
- [32] Marvin. L, MACHINE LEARNING: Neural Networks, Decision Trees and Support Vector Machine with IBM SPSS Modeler. Scientific Books, 2022.
- [33] “The Complete Guide to Support Vector Machine (SVM)”, towardsdatascience.com, 2022. [Online]. Disponible en: <https://towardsdatascience.com/the-complete-guide-to-support-vector-machine-svm-f1a820d8af0b>.
- [34] “Random Forest: Bosque aleatorio. Definición y funcionamiento”, datascientest.com, 2022. [Online]. Disponible en: <https://datascientest.com/es/random-forest-bosque-aleatorio-definicion-y-funcionamiento>.
- [35] J. Brownlee, XGBoost With Python. 2018.
- [36] “Gradient Boosting Algorithm: A Complete Guide for Beginners”, Analytics Vidhya, 2022. [Online]. Disponible en: <https://www.analyticsvidhya.com/blog/2021/09/>

[gradient-boosting-algorithm-a-complete-guide-for-beginners/](#).

- [37] “XGBoost Documentation — xgboost 1.6.1 documentation”, [Xgboost.readthedocs.io](#), 2022. [Online]. Disponible en: <https://xgboost.readthedocs.io/en/stable/>.
- [38] W. Sullivan, Machine learning Beginners Guide Algorithms: Supervised Un-supervised learning, Decision Tree Random Forest Introduction. CreateSpace Independent Publishing Platform, 2017.
- [39] J. Rueda, “Aprendizaje supervisado y no supervisado”, [healthdataminer.com](#), 2022. [Online]. Disponible en: <https://healthdataminer.com/data-mining/aprendizaje-supervisado-y-no-supervisado/#:~:text=El%20aprendizaje%20supervisado%20supone%20que,de%20datos%20no%20etiquetados%20previamente>.
- [40] “Train/Test Split and Cross Validation in Python”, Medium, 2022. [Online]. Disponible en: <https://towardsdatascience.com/train-test-split-and-cross-validation-in-python-80b61beca4b6>.
- [41] “3.1. Cross-validation: evaluating estimator performance”, scikit-learn, 2022. [Online]. Disponible en: https://scikit-learn.org/stable/modules/cross_validation.html.
- [42] J. Brownlee, “A Gentle Introduction to k-fold Cross-Validation”, Machine Learning Mastery, 2022. [Online]. Disponible en: <https://machinelearningmastery.com/k-fold-cross-validation/>.
- [43] J. Brownlee, “Metrics To Evaluate Machine Learning Algorithms in Python”, Machine Learning Mastery, 2022. [Online]. Disponible en: <https://machinelearningmastery.com/metrics-evaluate-machine-learning-algorithms-python/>.
- [44] R. Díaz, “Métricas de Clasificación - Aprende a EVALUAR tu modelo”, The Machine Learners, 2022. [Online]. Disponible en: <https://www.themachinelearners.com/metricas-de-clasificacion/#Accuracy>.

- [45] “Evaluation Metrics for Classification Problems with Implementation in Python”, Medium, 2022. [Online]. Disponible en: <https://medium.com/analytics-vidhya/evaluation-metrics-for-classification-problems-with-implementation-in-python-a2>
- [46] Huynh, K., Piguet, O., Kwok, J., Dobson-Stone, C., Halliday, G., Hodges, J. and Landin-Romero, R., 2021. Clinical and Biological Correlates of White Matter Hyperintensities in Patients With Behavioral-Variant Frontotemporal Dementia and Alzheimer Disease. *Neurology*, 96(13), pp.e1743-e1754.
- [47] Zhao, H., Cheng, J., Liu, T., Jiang, J., Koch, F., Sachdev, P., Basser, P. and Wen, W., 2021. Orientational changes of white matter fibers in Alzheimer’s disease and amnesic mild cognitive impairment. *Human Brain Mapping*, 42(16), pp.5397-5408.
- [48] “How to Deal with Missing Data”, CORP-MIDS1 (MDS), 2022. [Online]. Disponible en: <https://www.mastersindatascience.org/learning/how-to-deal-with-missing-data/>.
- [49] “Cecilia Gonzalez Campo”, Conicet.gov.ar, 2022. [Online]. Disponible en: https://www.conicet.gov.ar/new_scp/detalle.php?id=53008&datos_academicos=yes
- [50] “Sebastián Moguilner”, Incyt.conicet.gov.ar, 2022. [Online]. Disponible en: <https://incyt.conicet.gov.ar/sebastian-moguilner/>
- [51] J. Rohrer and H. Rosen, “Neuroimaging in frontotemporal dementia”, *International Review of Psychiatry*, vol. 25, no. 2, pp. 221-229, 2013. Disponible en: 10.3109/09540261.2013.778822.
- [52] M. Torso et al., “Strategic Lesions in the Anterior Thalamic Radiation and Apathy in Early Alzheimer’s Disease”, *PLOS ONE*, vol. 10, no. 5, p. e0124998, 2015. Disponible en: 10.1371/journal.pone.0124998.
- [53] E. Bubb, C. Metzler-Baddeley and J. Aggleton, “The cingulum bundle: Anatomy, function, and dysfunction”, *Neuroscience & Biobehavioral Reviews*, vol. 92, pp. 104-127, 2018. Disponible en: 10.1016/j.neubiorev.2018.05.008.

Apéndice A

Anexo

A.1. EA-Control

A.1.1. Datos combinados

SVM

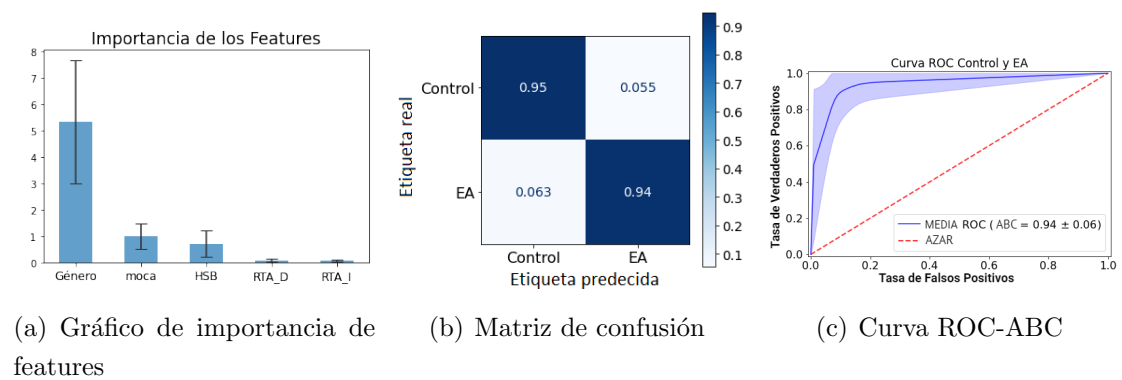


Figura A.1: Resultados del algoritmo SVM para los grupos de EA y Control

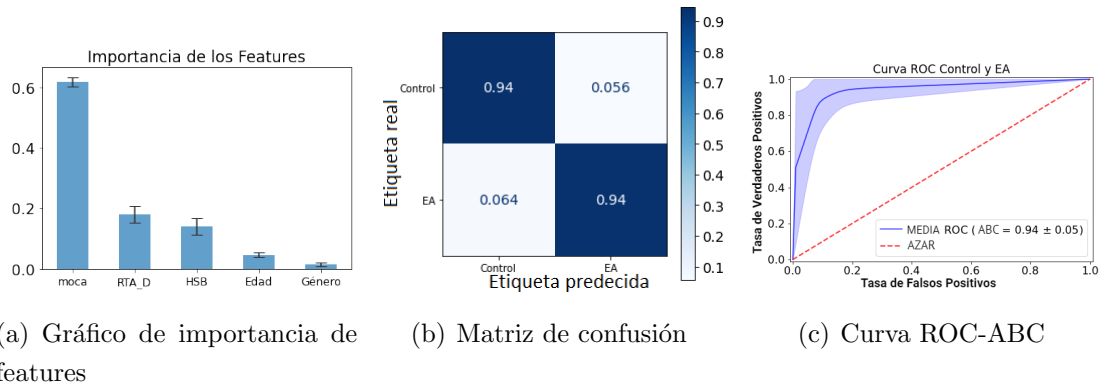
RF

Figura A.2: Resultados del algoritmo RF para los grupos de EA y Control

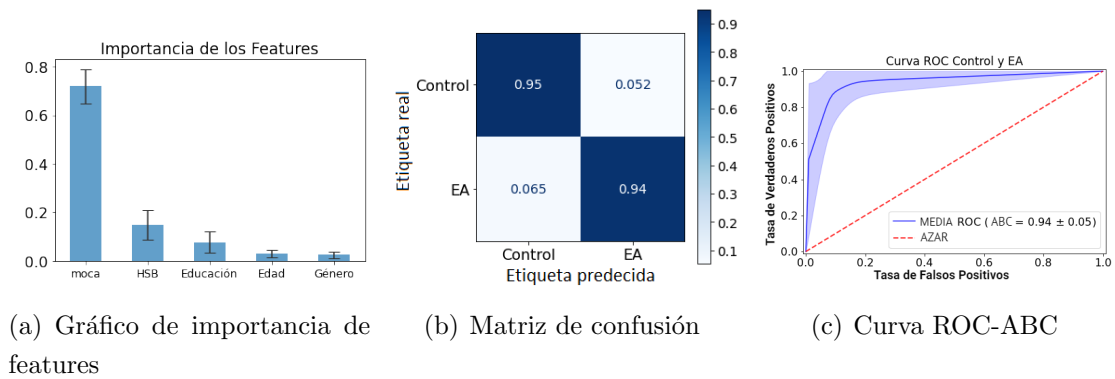
GB

Figura A.3: Resultados del algoritmo GB para los grupos de EA y Control

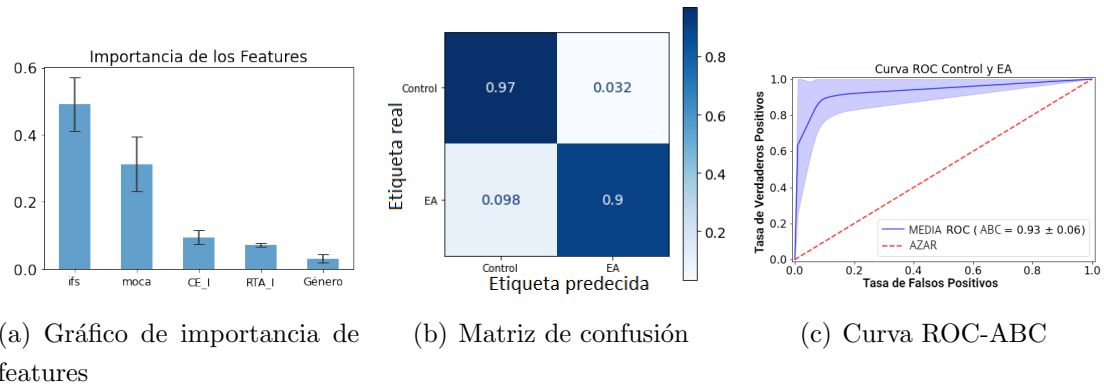
XGB

Figura A.4: Resultados del algoritmo XGB para los grupos de EA y Control

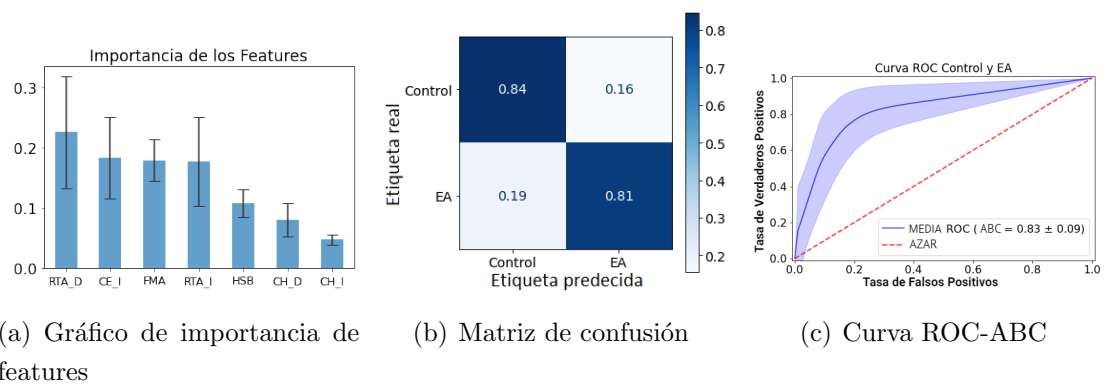
A.1.2. Datos neurorradiológicos**RF**

Figura A.5: Resultados del algoritmo RF para los grupos de EA y Control

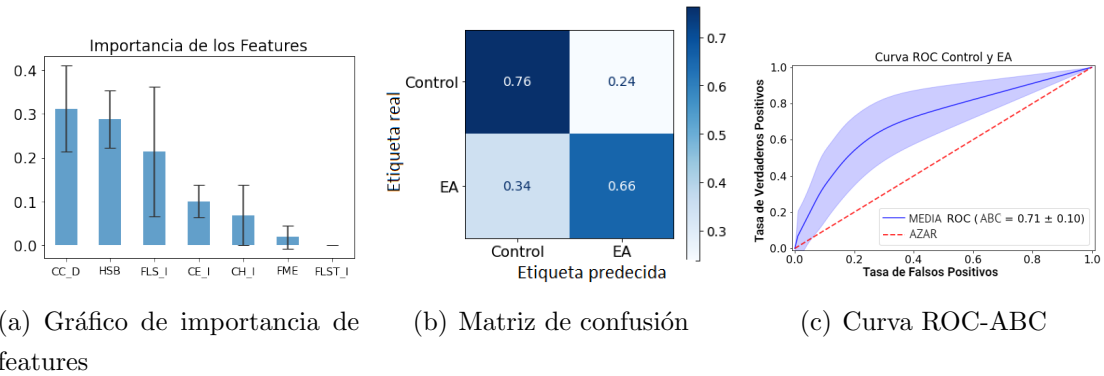
GB

Figura A.6: Resultados del algoritmo GB para los grupos de EA y Control

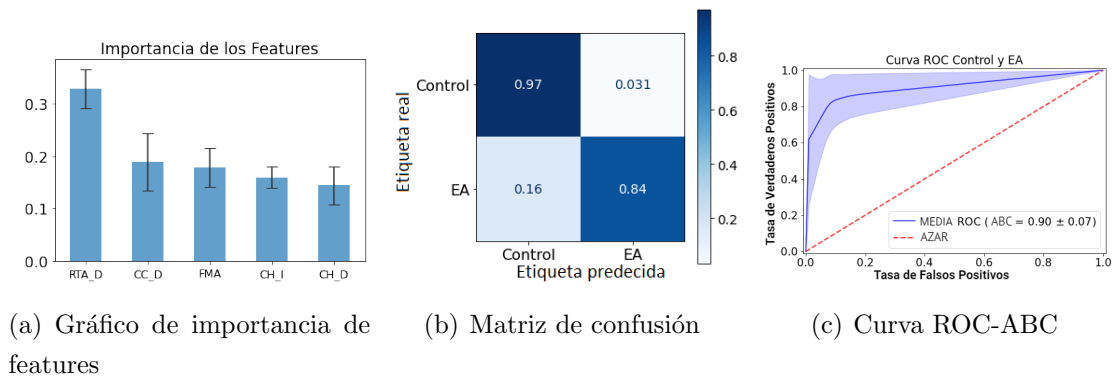
XGB

Figura A.7: Resultados del algoritmo XGB para los grupos de EA y Control

A.1.3. Datos cognitivos

SVM

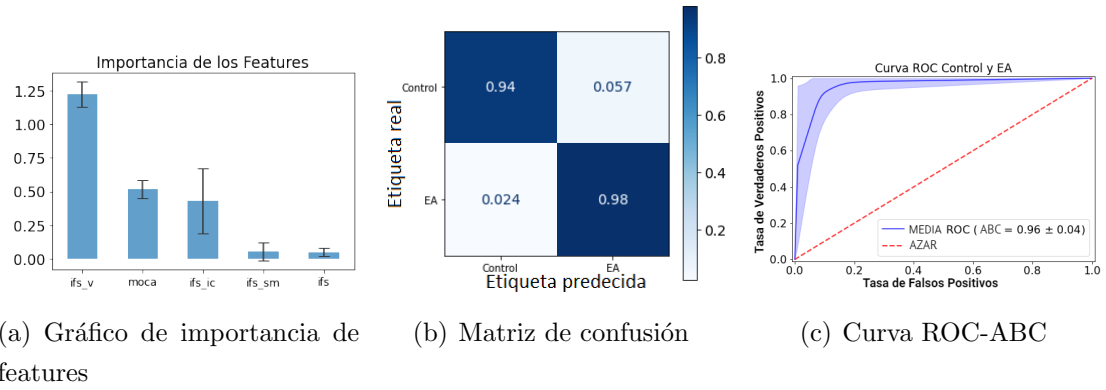


Figura A.8: Resultados del algoritmo SVM para los grupos de EA y Control

RF

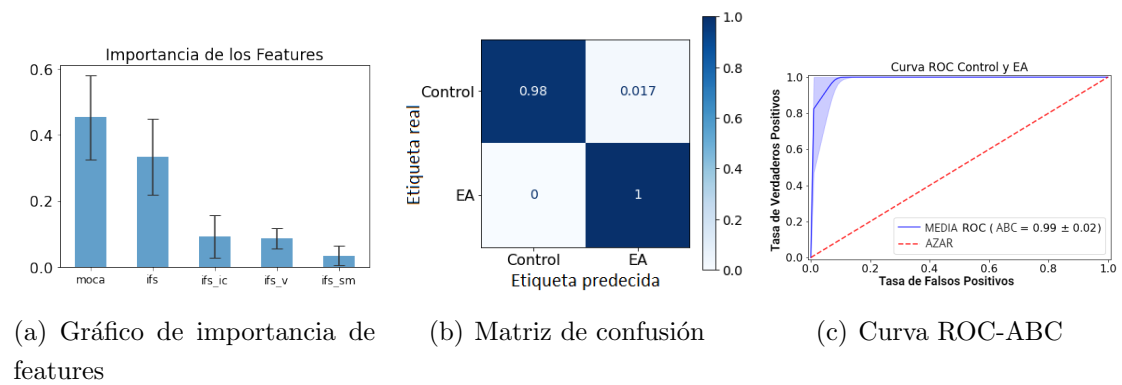


Figura A.9: Resultados del algoritmo RF para los grupos de EA y Control

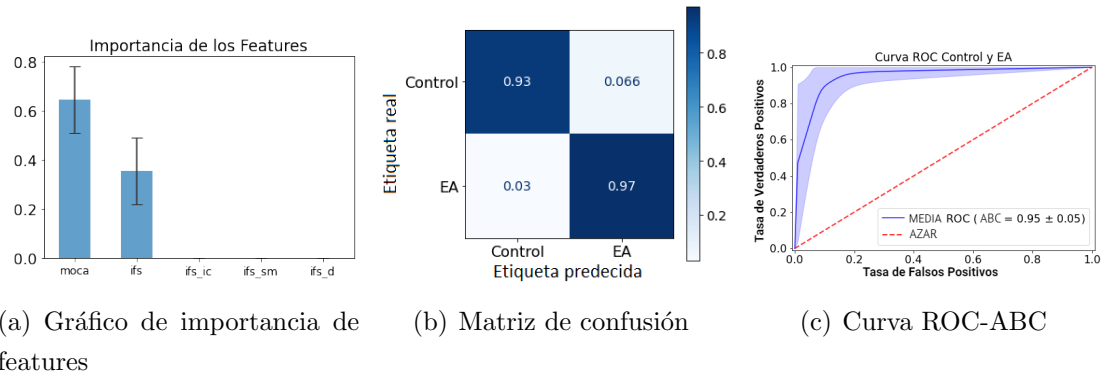
GB

Figura A.10: Resultados del algoritmo GB para los grupos de EA y Control

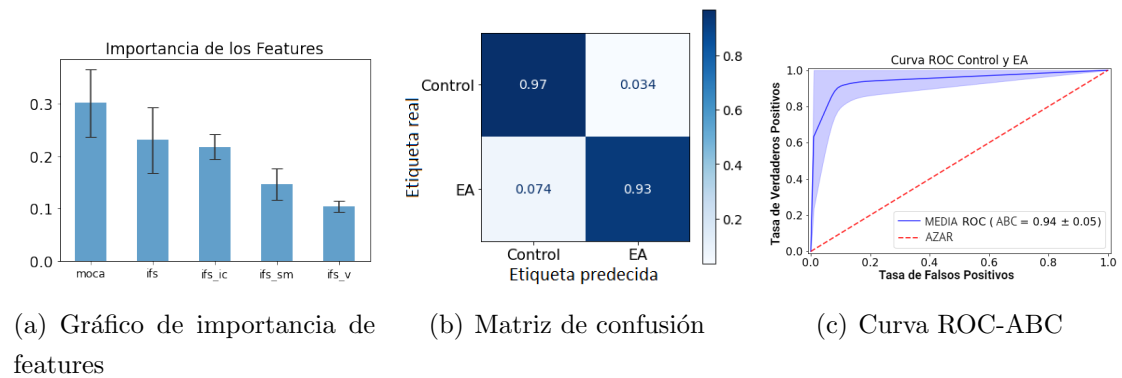
XGB

Figura A.11: Resultados del algoritmo XGB para los grupos de EA y Control

A.2. EA-DFT

A.2.1. Datos combinados

SVM

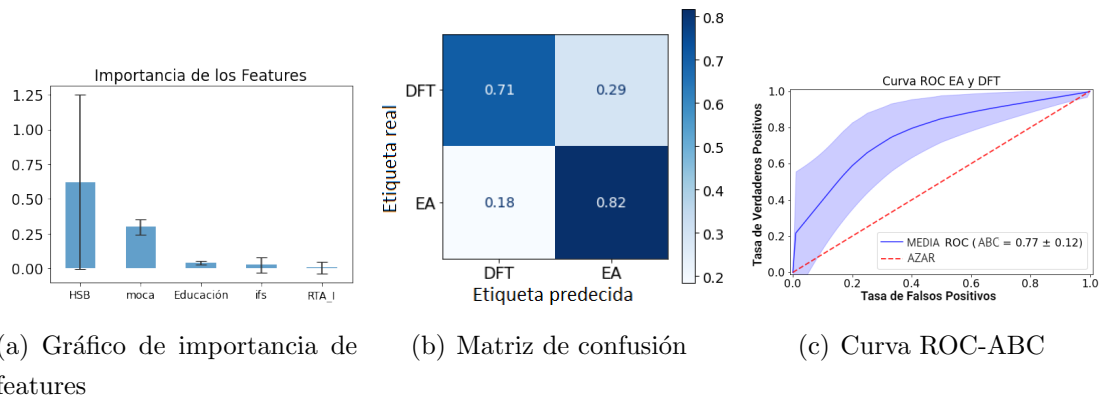


Figura A.12: Resultados del algoritmo SVM para los grupos de EA y DFT

RF

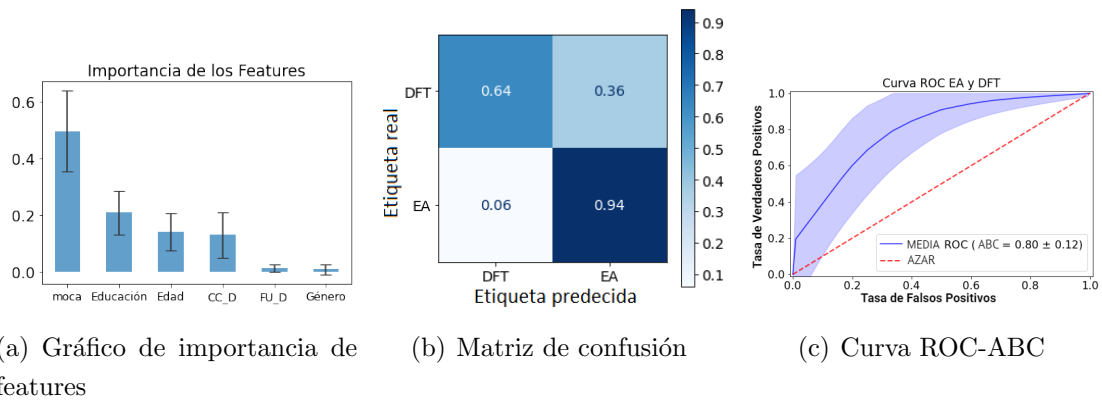


Figura A.13: Resultados del algoritmo RF para los grupos de EA y DFT

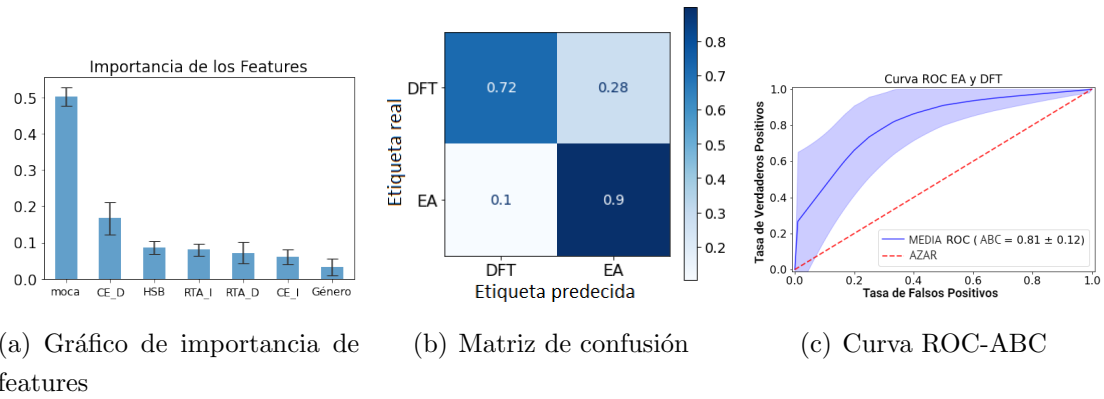
GB

Figura A.14: Resultados del algoritmo GB para los grupos de EA y DFT

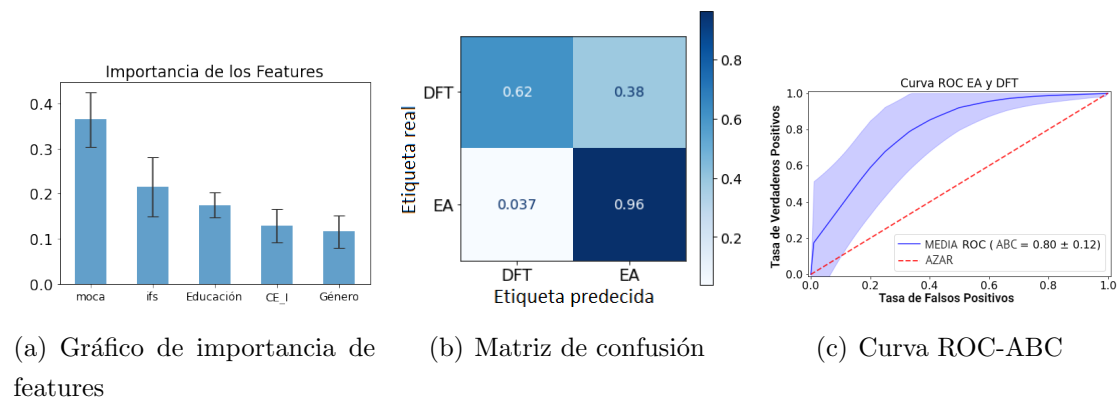
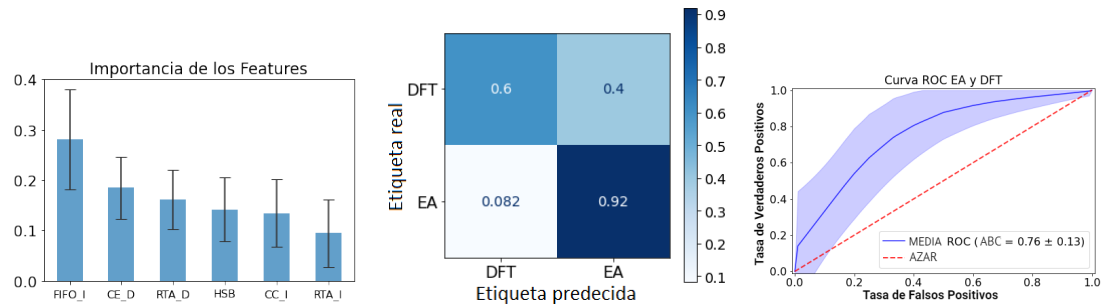
XGB

Figura A.15: Resultados del algoritmo XGB para los grupos de EA y DFT

A.2.2. Datos neurorradiológicos

RF



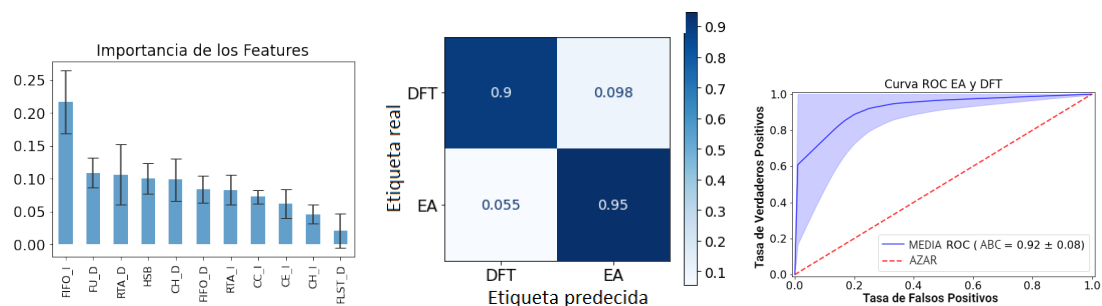
(a) Gráfico de importancia de features

(b) Matriz de confusión

(c) Curva ROC-ABC

Figura A.16: Resultados del algoritmo RF para los grupos de EA y DFT

GB



(a) Gráfico de importancia de features

(b) Matriz de confusión

(c) Curva ROC-ABC

Figura A.17: Resultados del algoritmo GB para los grupos de EA y DFT

XGB

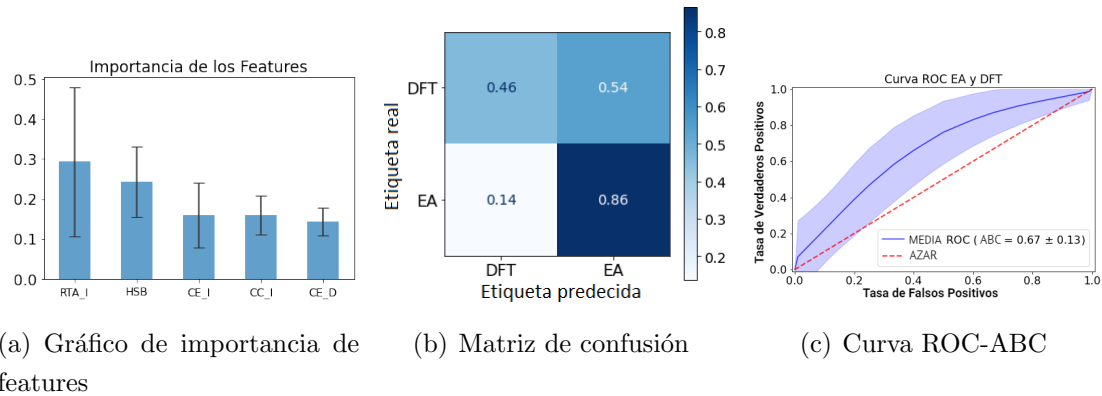


Figura A.18: Resultados del algoritmo XGB para los grupos de EA y DFT

A.2.3. Datos cognitivos

SVM

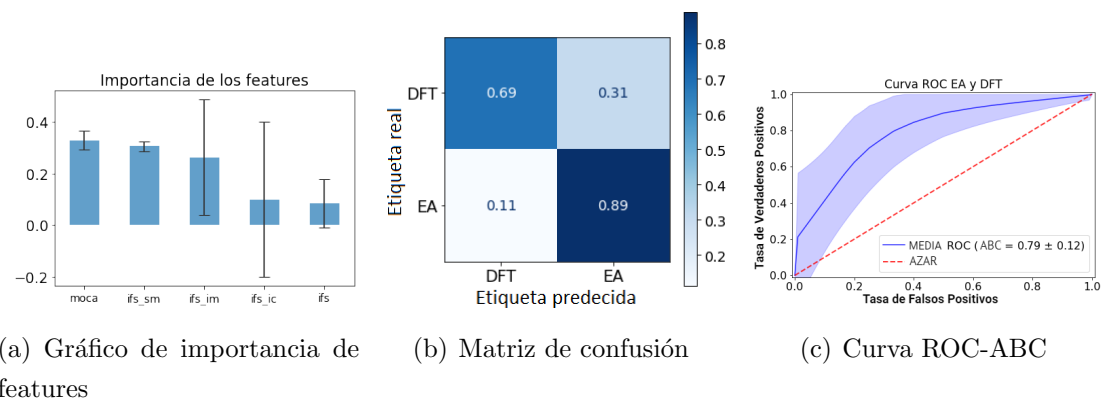
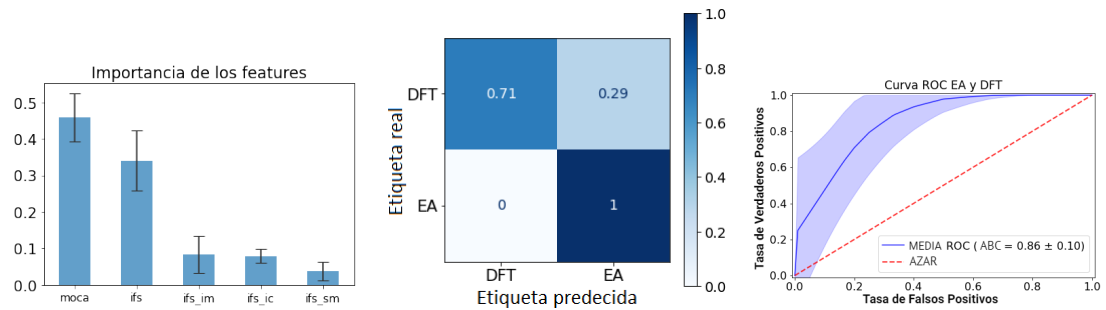


Figura A.19: Resultados del algoritmo SVM para los grupos de EA y DFT

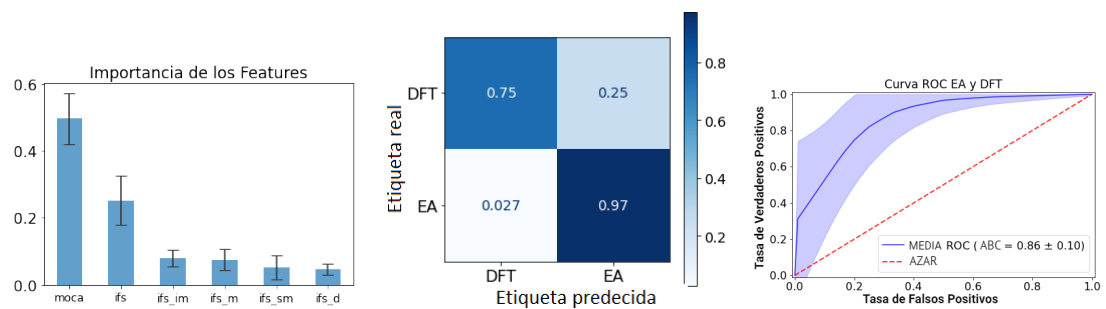
RF

(a) Gráfico de importancia de features

(b) Matriz de confusión

(c) Curva ROC-ABC

Figura A.20: Resultados del algoritmo RF para los grupos de EA y DFT

GB

(a) Gráfico de importancia de features

(b) Matriz de confusión

(c) Curva ROC-ABC

Figura A.21: Resultados del algoritmo GB para los grupos de EA y DFT

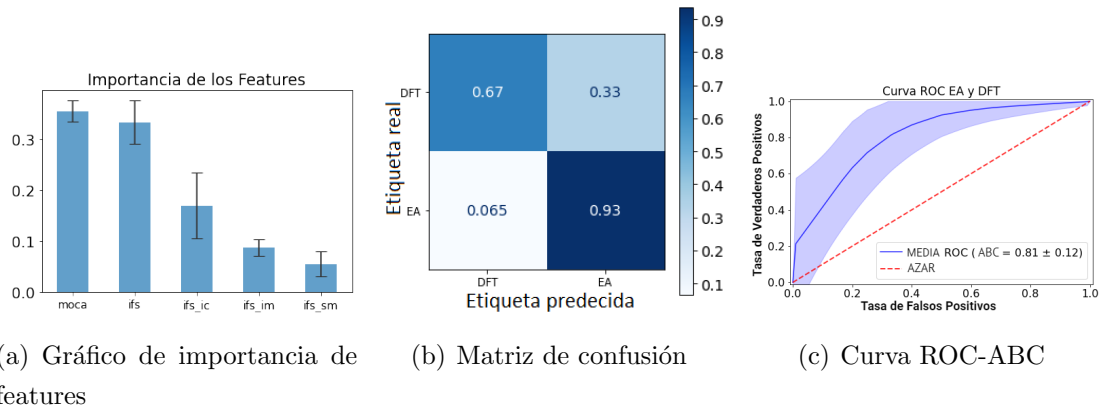
XGB

Figura A.22: Resultados del algoritmo XGB para los grupos de EA y DFT

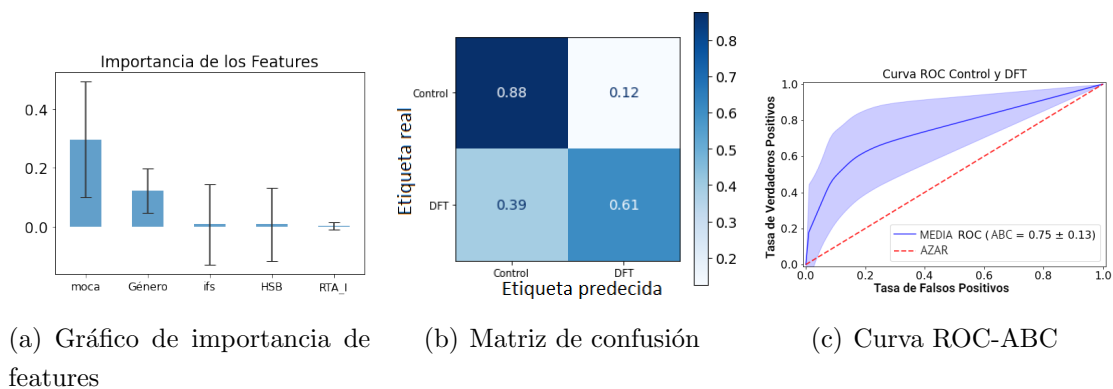
A.3. DFT-Control**A.3.1. Datos combinados****SVM**

Figura A.23: Resultados del algoritmo SVM para los grupos de DFT y Control

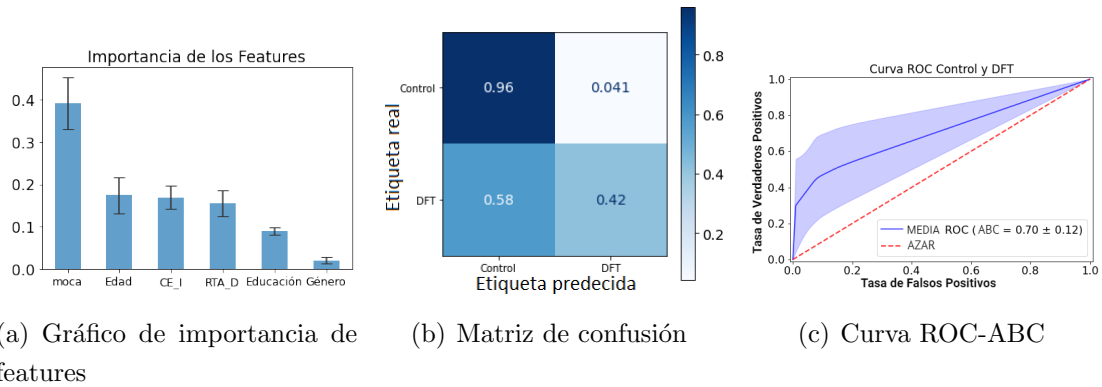
RF

Figura A.24: Resultados del algoritmo RF para los grupos de DFT y Control

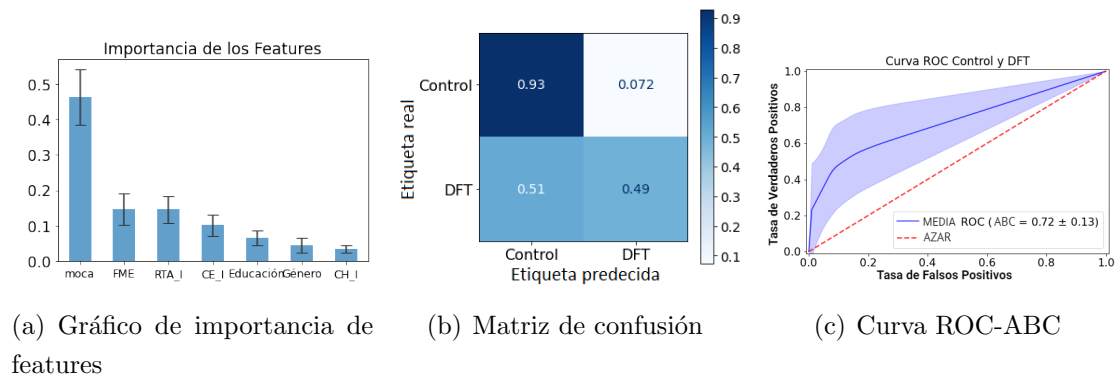
GB

Figura A.25: Resultados del algoritmo GB para los grupos de DFT y Control

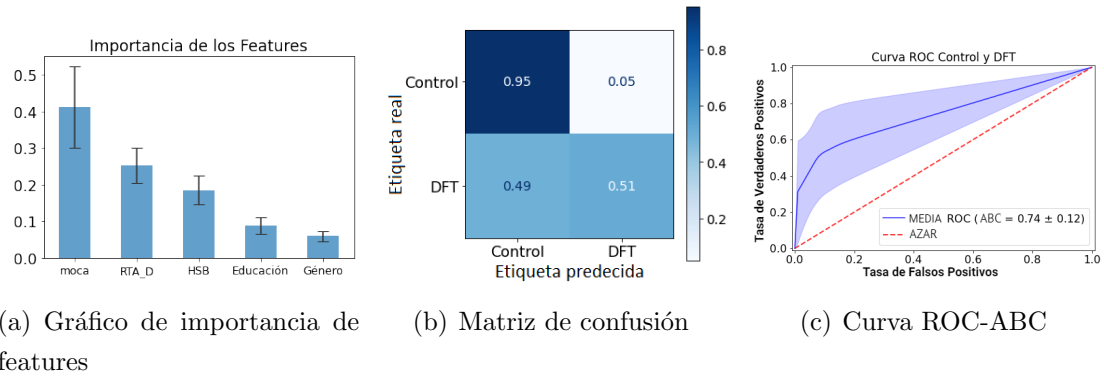
XGB

Figura A.26: Resultados del algoritmo XGB para los grupos de DFT y Control

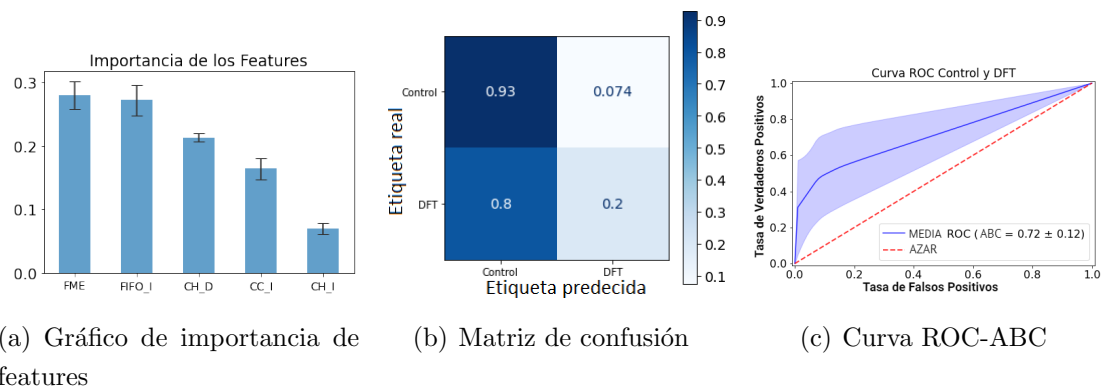
A.3.2. Datos neurorradiológicos**RF**

Figura A.27: Resultados del algoritmo RF para los grupos de DFT y Control

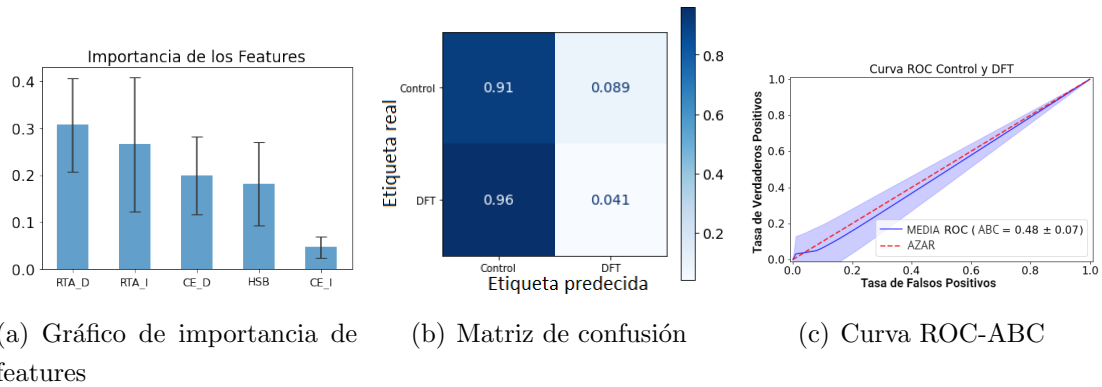
GB

Figura A.28: Resultados del algoritmo GB para los grupos de DFT y Control

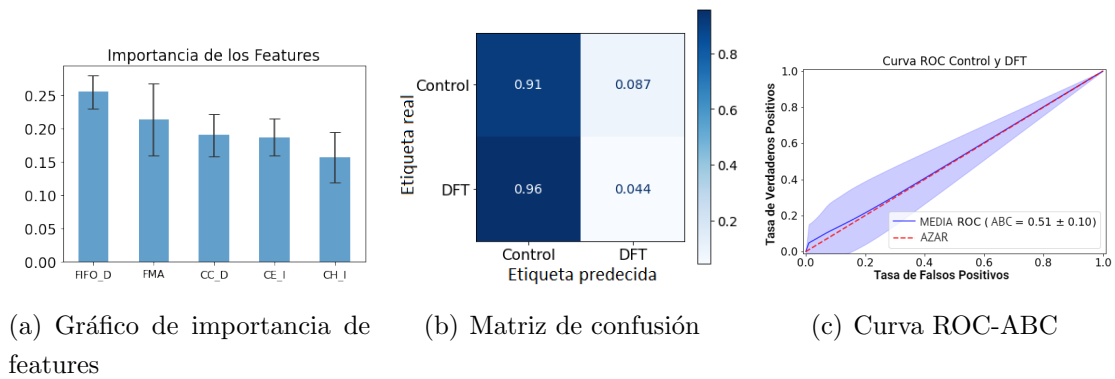
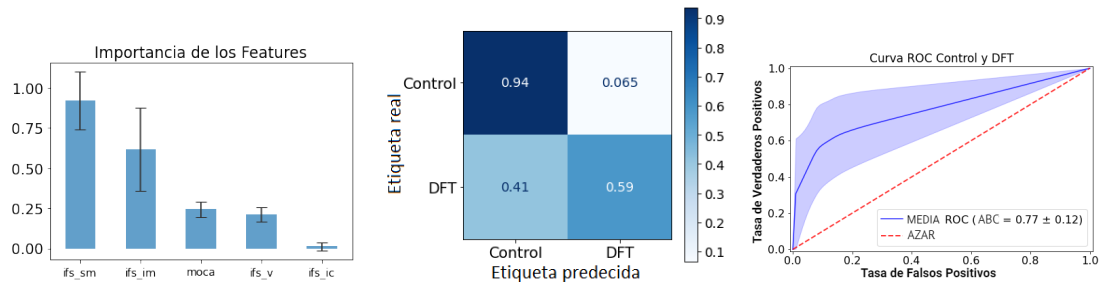
XGB

Figura A.29: Resultados del algoritmo XGB para los grupos de DFT y Control

A.3.3. Datos cognitivos

SVM



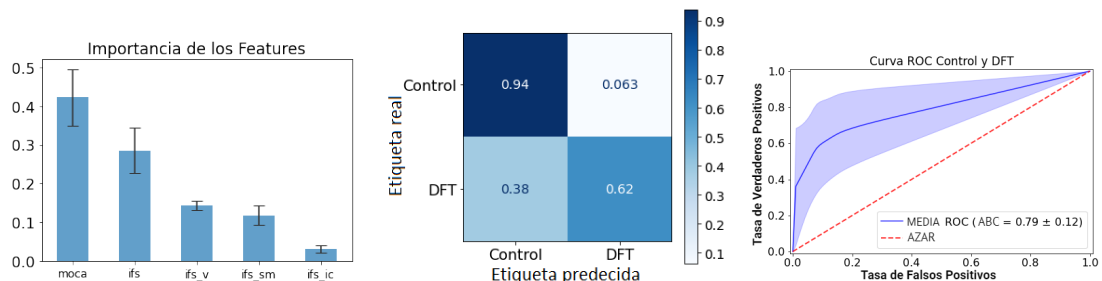
(a) Gráfico de importancia de features

(b) Matriz de confusión

(c) Curva ROC-ABC

Figura A.30: Resultados del algoritmo SVM para los grupos de DFT y Control

RF



(a) Gráfico de importancia de features

(b) Matriz de confusión

(c) Curva ROC-ABC

Figura A.31: Resultados del algoritmo RF para los grupos de DFT y Control

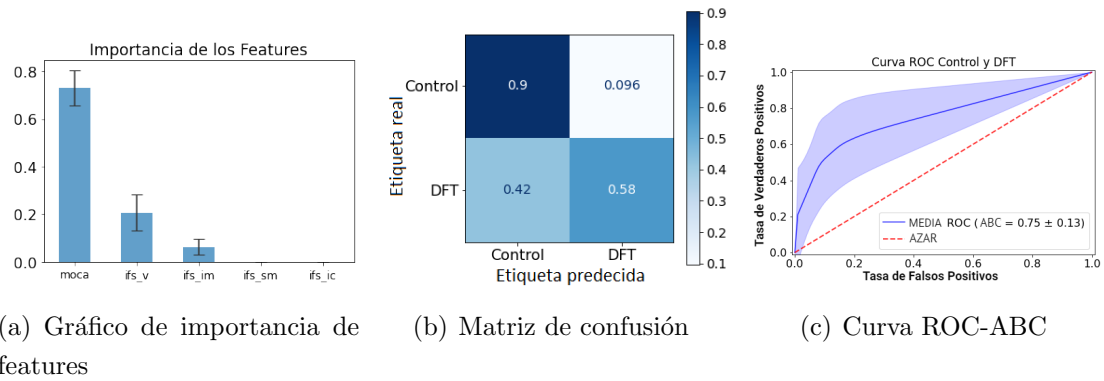
GB

Figura A.32: Resultados del algoritmo GB para los grupos de DFT y Control

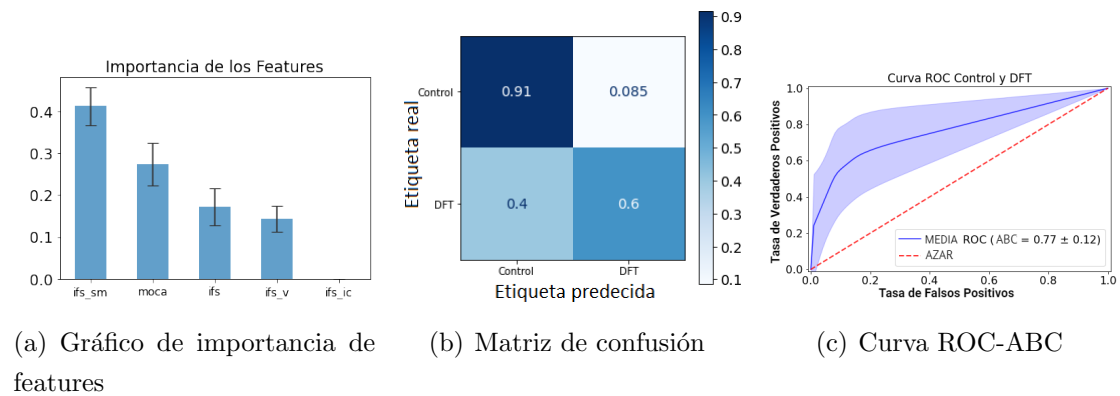
XGB

Figura A.33: Resultados del algoritmo XGB para los grupos de DFT y Control