



UNIVERSIDAD DE DEUSTO

ALGORITMOS PARA SELECCIÓN DE CARACTERÍSTICAS Y CLASIFICACIÓN DE SUJETOS PARA LAS PATOLOGÍAS DE DISLEXIA Y MIGRAÑA

Tesis doctoral presentada por Yolanda García Chimeno
dentro del Programa de Doctorado en Ingeniería para la Sociedad de la
Información y Desarrollo Sostenible

Dirigida por Dra. Begoña García-Zapirain Soto



UNIVERSIDAD DE DEUSTO

ALGORITMOS PARA SELECCIÓN DE CARACTERÍSTICAS Y CLASIFICACIÓN DE SUJETOS PARA LAS PATOLOGÍAS DE DISLEXIA Y MIGRAÑA

Tesis doctoral presentada por Yolanda García Chimeno
dentro del Programa de Doctorado en Ingeniería para la Sociedad de la
Información y Desarrollo Sostenible

Dirigida por Dra. Begoña García-Zapirain Soto

El doctorando

El director

Bilbao, Julio de 2020

Algoritmos para selección de características y clasificación de sujetos para las patologías de Dislexia y Migraña

Autor: Yolanda García Chimeno

Directora: Begoña García-Zapirain Soto

Texto impreso en in Bilbao

Primera edición, septiembre 2020

*A mi familia, y a la persona que ha recorrido junto a mi 12 años de
mi vida, y por muchos más...*

Resumen

Las técnicas de Machine Learning y Deep Learning se han extendido en el uso tanto para la predicción como la clasificación en múltiples campos tales como la automoción, energías renovables, finanzas, etc. Pero un campo que está continuamente en desarrollo e investigación es su aplicación en el campo de la medicina. Estas técnicas de aprendizaje automático permite un nuevo método de poder tratar a un paciente en función de unas características generales o particulares en función de la patología que se quiera investigar. Además, también dentro de estas técnicas de Machine y Deep Learning, es muy importante el campo de la selección de características, lo que permite tener presente la importancia y relevancia de cada una de ellas para poder realizar una correcta clasificación. En la realización de esta tesis se ha hecho uso de dos bases de datos, cada una perteneciente a un caso de estudio: una de ellas cuenta con 57 sujetos (dislexia), y la perteneciente al otro caso de estudio con 52 sujetos (migraña); donde en cada una de ellas se encuentran diferentes grupos de sujetos. El primer caso de estudio es el de Dislexia, en el que se encuentran un grupo de controles, un grupo de sujetos con Dislexia y otro grupo con visión monocular. En el segundo caso de estudio, el de Migraña, se cuenta con un grupo de controles, otro grupo de pacientes con migraña esporádica y el último grupo de pacientes con migraña crónica y abuso de medicamentos.

Para ello se ha diseñado una investigación centrada en cada uno de los dos casos de estudio, en el que se analizan los diversos algoritmos tanto de selección de características como de clasificación para poder clasificar a los sujetos en los diferentes grupos. En segundo lugar se han implementado estos algoritmos para poder aplicarlos a los conjuntos

de datos, provenientes de imágenes DTI (Diffusion Tensor Imaging) en los que se obtiene los tractos cerebrales de la materia blanca. Se han aplicado tanto técnicas de Machine Learning, como de Deep Learning. Finalmente, para la validación de esta tesis, se ha realizado un estudio en el que mediante medidas estadísticas, se mide el ratio de acierto de los clasificadores, y la importancia de seleccionar las características más relevantes para obtener un mejor ratio en la clasificación de cada uno de los sujetos. Se ha realizado una comparativa entre los resultados provenientes de los algoritmos de Machine y Deep Learning, obteniendo un resultado positivo para ambos, pero teniendo en cuenta que las técnicas de Deep Learning mejora el acierto en la clasificación. Estos resultados proporcionan un conjunto de algoritmos validados para la ayuda al diagnóstico por parte del profesional médico.

Agradecimientos

Toda meta que nos ponemos por delante no es algo individual, ya que sin la ayuda, apoyo y participación de numerosas personas, el camino sería mucho más arduo y complicado. Gracias a la dedicación desinteresada de todas estas personas, esta tesis doctoral consiguió darse forma y finalizarla dejando atrás momentos duros, pero también otros muy bonitos, gracias a los cuales he podido crecer como persona y como profesional.

En primer lugar, quiero agradecer a mi directora de tesis, la Dra. María Begoña García-Zapirain Soto, por permitirme realizar esta tesis doctoral bajo su dirección, y que gracias a sus infinitas ayudas y apoyo moral, no hubiera podido finalizar con éxito la presente tesis.

También quiero agradecer el tiempo empleado y esfuerzo dedicado a los especialistas del grupo de investigación de Dislexia y de Migraña, ya que ha sido una colaboración mutua para obtener resultados significativos tanto en la medicina, como en mi campo de Ingeniería y que han servido para poder finalizar con éxito esta tesis. Por estos motivos quiero agradecer a Ibone Saralegui Prieto, Begoña Fernández Ruanova, Marian Gómez Beldarrain, Alberto Cabrera y Juan Carlos García Moncó, todas las horas de trabajo que han permitido conseguir que sigamos apostando por el campo de la investigación, pese a los momentos en los que podamos flaquear.

Otras personas a las cuales quiero agradecer el tiempo y esfuerzo que me han brindado es a mis ex-compañeros de departamento, Fernando Jorge, Alain Sánchez, Zelai Sáenz de Urturi, Maite Frutos y Álvaro Muro. Han sido cómplices y amigos que han hecho que el camino sea

mucho más llevadero. Especial mención de este grupo tanto para Fernando como para Alain, muchas gracias, de verdad, ya que sin esas charlas, consejos, revisiones, todo tipo de ayuda incondicional y sobre todo muchas risas esto no hubiera sido posible.

También quiero agradecer todo el apoyo a mi familia, que no sólo en esta aventura, han sido un pilar fundamental, haciendo que mi mente se liberara en los momentos más críticos. Y gracias Abuela, que tu siempre dices que esto no lo entiendes, pero sabes el esfuerzo que conlleva.

Otros dos pilares fundamentales a los que no puedo dejar de agradecer durante toda mi vida son mis padres. Gracias Ama, siempre confías y siempre sabes que antes o después llegaré a la meta, sin esas charlas y ayudas a levantarme tras todas las caídas, han hecho posible que pueda conseguir el doctorado. Gracias Aita, tampoco tú has dejado de confiar en todo lo que me quería proponer y siempre entendiendo el por qué de mis decisiones. De verdad, muchas gracias a los dos, pese a mis malos momentos, siempre estáis ahí.

Por último, un “eskerrik asko“ muy especial a Iñigo Ortiz-Vigón, que no sólo en esta tesis, sino durante los 12 años de camino juntos ha hecho que siga creciendo día a día. Gracias por las revisiones, y sobre todo por tu sinceridad en la vida, y sacándome siempre una sonrisa en los peores momentos. Y sobre todo por hacer que conozca mundo y gastronomía, y que podamos seguir haciéndolo por muchos años más. “Eskerrik asko laztana“.

Eskerrik asko,

Yolanda García Chimeno

septiembre 2020

Índice general

Índice de figuras	xi
Índice de tablas	xiii
1 Introducción	1
1.1 Justificación de la investigación	3
1.2 Hipótesis	5
1.3 Objetivos	5
1.4 Metodología de la investigación	6
1.5 Estructura del documento	7
2 Estudio del estado del arte	11
2.1 Conceptos y definiciones	11
2.1.1 Resonancia magnética e imágenes DTI	12
2.1.2 Inteligencia Artificial	12
2.1.3 Términos estadísticos	13
2.1.3.1 Accuracy	13
2.1.3.2 Precision	14
2.1.3.3 Recall	14
2.1.3.4 F1 score	14
2.2 Antecedentes tecnológicos	15
2.2.1 Selección de características	15
2.2.1.1 Gradient Tree Boosting	15
2.2.1.2 Random Forest	16

ÍNDICE GENERAL

2.2.1.3	Univariate feature selection	18
2.2.1.4	Tree-Based	19
2.2.1.5	L1-based feature selection	20
2.2.2	Machine Learning	21
2.2.2.1	SVM	21
2.2.2.2	Adaboost	22
2.2.2.3	Naïve Bayes	23
2.2.3	Deep Learning	23
2.2.3.1	Tensorflow	24
2.2.3.2	Convolutional Neural Network	25
2.3	Antecedentes médicos	25
2.3.1	DTI	25
2.3.2	Dislexia	26
2.3.3	Migraña	27
2.4	Revisión de investigaciones	28
2.4.1	Machine Learning	29
2.4.2	Deep Learning	29
2.4.3	Conclusiones	29
2.5	Sumario	30
3	Base de datos	33
3.1	Población de referencia	34
3.2	Criterios de selección	35
3.2.1	Caso de estudio 1: Dislexia	35
3.2.2	Caso de estudio 2: Migraña	37
3.3	Tamaño muestral	38
3.3.1	Caso de estudio 1: Dislexia	38
3.3.2	Caso de estudio 2: Migraña	39
3.4	Método	39
3.4.1	Caso de estudio 1: Dislexia	40
3.4.1.1	Evaluación Neuropsicológica	40
3.4.1.2	Exploración DTI	41
3.4.2	Caso de estudio 2: Migraña	42

3.4.2.1	Evaluación Neuropsicológica	42
3.4.2.2	Exploración DTI	43
3.5	Sumario	44
4	Diseño	45
4.1	Diseño de alto nivel	46
4.1.1	Pre-procesado imágenes DTI (A)	46
4.1.2	Selección de características (B)	46
4.1.3	Algoritmos Machine Learning (C)	48
4.1.4	Comité (D)	48
4.1.5	Algoritmos Deep Learning (E)	49
4.2	Diseño de bajo nivel	49
4.2.1	Diseño de bajo nivel	49
4.2.1.1	Conversión (A1)	50
4.2.1.2	Selección de Regiones de Interés (A2)	51
4.2.1.3	Características (A3)	54
4.2.2	Selección de características (B)	55
4.2.2.1	Gradient Tree Boosting (B1)	56
4.2.2.2	Random Forest (B2)	57
4.2.2.3	Univariate feature selection (B3)	58
4.2.2.4	Tree-Based (B4)	58
4.2.2.5	L1-based feature selection (B5)	59
4.2.2.6	Método ensemble (B6)	59
4.2.3	Algoritmos Machine Learning (C)	59
4.2.3.1	SVM (C1)	61
4.2.3.2	Adaboost (C2)	62
4.2.3.3	Naïve Bayes (C3)	63
4.2.4	Comité (D)	64
4.2.4.1	Método de Ensemble algoritmos de Machine Learning(D1)	64
4.2.5	Algoritmos Deep Learning (E)	66
4.2.5.1	Convolutional Neural Network (E1)	66
4.3	Sumario	69

ÍNDICE GENERAL

5	Resultados	71
5.1	Definición del experimento	71
5.2	Evaluación de los resultados	73
5.2.1	Pre-procesado imágenes DTI (A)	73
5.2.2	Selección de características (B)	78
5.2.3	Algoritmos Machine Learning (C)	83
5.2.4	Comité (D)	85
5.2.5	Algoritmos Deep Learning (E)	91
5.3	Comparativa de Machine Learning vs. Deep Learning	93
5.3.1	Ratios de acierto	93
5.3.2	Recursos Hardware	94
5.3.3	Tiempos de procesamiento	95
5.3.4	Conclusión	96
5.4	Sumario	96
6	Conclusiones	99
6.1	Conclusiones técnicas	101
6.2	Conclusiones médicas	102
6.3	Impacto científico	102
6.3.1	Artículos en revistas científicas	103
6.3.2	Comunicación en congresos	103
6.3.3	Proyectos relacionados	103
6.4	Líneas futuras	104
	Bibliografía	107

Índice de figuras

2.1	Matriz de confusión.	14
2.2	Sobreajuste en un modelo.	17
3.1	Población universo a la muestra.	35
4.1	Diseño de alto nivel.	47
4.2	Tractos del cerebro.	50
4.3	Bajo nivel (A).	50
4.4	Región de interés en las 3 dimensiones.	52
4.5	Salida de la función Featquery.	54
4.6	Bajo nivel (B).	56
4.7	Método ensemble (B6).	60
4.8	Algoritmo SVM (C1).	61
4.9	Algoritmo Adaboost (C2).	63
4.10	Algoritmo Näive Bayes (C3).	63
4.11	Método ensemble de Machine Learning (D1).	65
4.12	Diseño de la red Convolutional Neural Network (E1).	68
5.1	Diseño de alto nivel.	72
5.2	Matriz con datos numéricos de las zonas cerebrales.	73
5.3	Arcuate Fasciculus.	74
5.4	Genu (Anterior Corpus Callosum).	74
5.5	Splenium (Posterior Corpus Callosum).	75
5.6	Fasciculus Occipitofrontalis Inferior.	75
5.7	Cápsula Interna.	76

ÍNDICE DE FIGURAS

5.8	Corpus Callosum.	76
5.9	Body Corpus Callosum.	76
5.10	Fasciculus Occipitofrontalis Inferior.	77
5.11	Genu (Anterior Corpus Callosum).	77
5.12	Insula.	78
5.13	Lobus Frontalis.	78

Índice de tablas

2.1	Tabla publicaciones Machine Learning en DTI	29
2.2	Tabla publicaciones Deep Learning en DTI	29
3.1	Sujetos participantes en el estudio (Dislexia)	38
3.2	Sujetos participantes en el estudio (Migraña)	39
5.1	Número y porcentaje de reducción de características (Dislexia) . .	79
5.2	Características seleccionadas (Dislexia)	80
5.3	Número y porcentaje de reducción de características (Migraña) . .	82
5.4	Características seleccionadas (Migraña)	82
5.5	Porcentaje de acierto en algoritmos con todo el conjunto de datos (Dislexia)	83
5.6	Porcentaje de acierto en algoritmos con la selección de característi- cas (Dislexia)	84
5.7	Porcentaje de acierto en algoritmos con todo el conjunto de datos (Migraña)	84
5.8	Porcentaje de acierto en algoritmos con la selección de característi- cas (Migraña)	85
5.9	Porcentaje de acierto con el comité (Dislexia)	86
5.10	Porcentaje de acierto con el comité (Migraña)	87
5.11	Porcentaje de acierto algoritmo Deep Learning con todas las carac- terísticas (Dislexia)	91
5.12	Porcentaje de acierto algoritmo Deep Learning con selección de características (Dislexia)	91

ÍNDICE DE TABLAS

5.13	Porcentaje de acierto algoritmo Deep Learning con todas las características (Migraña)	92
5.14	Porcentaje de acierto algoritmo Deep Learning con selección de características (Migraña)	92
5.15	Comparativa de porcentajes de acierto Machine Learning vs. Deep Learning	93
5.16	Comparativa de tiempo de entrenamiento Machine Learning vs. Deep Learning	95
6.1	Revistas científicas I	103
6.2	Revistas científicas II	103
6.3	Revistas científicas III	103
6.4	Revistas científicas IV	104
6.5	Comunicación en congreso I	104
6.6	Proyecto MIGREIN	104

Introducción

La exploración de diversas patologías las cuales tienen que ver con el cerebro, ha desembocado en un gran número de diferentes investigaciones, con el fin de buscar nuevas causas o revisar las que ya han sido enunciadas. Entre la multitud de investigaciones encontramos modos de representación, estudio de activaciones, conectividad, cambios estructurales han permitido una mayor comprensión del funcionamiento del cerebro. También estos estudios exhaustivos han sido posibles gracias a la evolución de las máquinas de resonancia magnética. Esta tesis se centra principalmente en algoritmos de Inteligencia Artificial aplicados para obtener una clasificación de sujetos que tienen una patología en función de sus imágenes de resonancia magnética. Para ello, se basa en las características físicas extraídas a partir de las imágenes DTI de resonancia magnética. La aplicación de algoritmos de Inteligencia Artificial puede ser muy extensa, debido al número alto de diferentes algoritmos existentes en la materia, cada uno basado en un método estadístico por detrás. Este tipo de técnicas, son capaces de resolver un amplio abanico de problemas, bien sea en temas de ingeniería, medicina e incluso en la capacidad de enfrentarse a un jugador de ajedrez. Pero con el avance tecnológico de los últimos años, la inteligencia artificial se está enfocando cada vez más en la resolución de problemas simulando la inteligencia humana, incluyendo sus aciertos y errores. Aquí es donde se pueden encontrar técnicas de Machine Learning, Data Mining,

1. INTRODUCCIÓN

procesamiento del lenguaje natural, como herramientas para el aprendizaje. Es en este punto en el que la inteligencia artificial, es una gran ayuda para mejorar la salud pública.

Cada vez más, las enfermedades neurológicas son las principales en cuanto a la causa de mortalidad y discapacidad, además de provocar un gasto económico tanto para las familias como para los servicios de salud de los países [Bru16] .

La inteligencia artificial aplicada al campo de la neurología ha sido recibida con mucho éxito en diversos campos, como su aplicación en pronósticos posteriores al diagnóstico, para el análisis de datos tanto en radiología como en anatomía patológica, y como una asistencia durante todo el proceso de diagnóstico del paciente.

Actualmente ya existen [OLM⁺12] sistemas de interpretación de neuroimágenes, los cuales son de gran ayuda ya que al ser tan precisos pueden apoyar al diagnóstico médico humano, enriqueciendo los métodos de diagnóstico y dejando atrás la subjetividad a la hora de analizar una patología.

Por tanto, esta tesis se centra en un primer análisis y aplicación de estos tipos de algoritmos para la clasificación de sujetos con una patología. Además, se realiza un análisis de extracción de características, obteniendo las más relevantes y detectando las zonas cerebrales que pueden estar afectadas por cada una de las patologías. Finalmente se realizará una validación de los algoritmos a través de una metodología de “ensemble”, en la cual se estudiará el resultado de cada algoritmo, y haciendo una conjunción de estos resultados para lograr una mayor precisión en cuanto a la clasificación.

Con la finalidad de validar y estudiar de manera correcta el proceso de clasificación mediante algoritmos de Inteligencia Artificial, toda esta investigación se encuentra enmarcada dentro de dos tipos de patologías: la dislexia y la migraña. De esta forma se pueden comprobar y establecer resultados reales a partir de un conjunto de entrada ya verificados previamente por los especialistas.

Por un lado, el estudio original de la dislexia que va a ser utilizado se centra en sujetos de corta edad divididos en niños disléxicos y niños con visión monocular. Por otro lado, el estudio de la migraña que va a ser utilizado también en esta tesis, se centra en sujetos que sufren migraña esporádica y sujetos con migraña crónica que además abusan de medicamentos. El resto del capítulo se encuentra organizado

de la siguiente manera: La sección 1.1 describe la justificación de la investigación realizada. La sección 1.2 contiene la hipótesis formulada. La sección 1.3 nombra el objetivo principal y los objetivos secundarios identificados para la ejecución de esta tesis. La sección 1.4 describe la metodología seguida durante esta investigación. Y por último, la sección 1.5 detalla la estructura de este documento.

1.1 Justificación de la investigación

La investigación se centra en la clasificación de distintas patologías en el cerebro de pacientes a través de un análisis y procesado de imágenes de resonancia magnética mediante algoritmos de Inteligencia Artificial. Esta tesis se ha centrado en dos tipos de patologías: la dislexia y la migraña.

Consecuentemente, existen varias justificaciones principales para la realización de esta investigación:

1. La dislexia es un trastorno que representa una dificultad en el habla o la dicción, es decir, entorpece y complica el proceso de aprendizaje e influye en la capacidad de lectura, escritura, realizar cálculos y memoria a corto plazo. Pero en esta investigación, se centra en la dislexia infantil, focalizando en la deficiencia de la lectura, escritura y aprendizaje. La causa es una alteración de las zonas cerebrales que controlan el lenguaje. La dislexia infantil, datos del 2016, afecta a un 5 % de niños entre 7 y 9 años, sobre todo a los varones. En la actualidad el diagnóstico ha avanzado bastante, pudiendo diagnosticar a los niños a partir de los 7 años, aunque hay numerosos casos que empiezan a detectarse antes. Es muy importante realizar un diagnóstico temprano y específico para cada niño. La obtención de resultados significativos en las zonas del cerebro afectadas por esta patología, suponen un buen punto de partida para diversas investigaciones, ayudando desde el primer día del diagnóstico a atenuar los síntomas.
2. La migraña es una patología muy frecuente en la actualidad, en España según datos la Asociación Española de Migraña y Cefalea (AEMICE) publicado en

1. INTRODUCCIÓN

el Atlas de la Migraña en España 2018, afecta a un 12 % de la población siendo más de cuatro millones de españoles. De estos afectados, el 80 % son mujeres. Esta patología tiene un gran impacto a nivel laboral, familiar y social. A medida que las exploraciones neuroradiológicas se van perfeccionando, se está comprobando que existen fármacos específicos para esta patología que producen un daño en algunas regiones cerebrales implicadas en la percepción del dolor y de la emoción. La obtención de resultados significativos en lo que se refiere a las áreas del cerebro afectadas, bien sea por los medicamentos o zonas relevantes de la propia patología, suponen un punto de partida para diversas investigaciones tanto para neurólogos como para otros especialistas en cuanto a la realización de diagnósticos.

3. Los algoritmos de Inteligencia Artificial son un conjunto de instrucciones que a partir de unos datos de entrada, construye otros de salida. Específicamente, en esta investigación se usan algoritmos de aprendizaje automatizado, los cuales permiten a las máquinas aprender, es decir, generalizar comportamientos a partir de una información dada. Gracias a estos algoritmos, se pueden obtener patrones los cuales permitan emitir una clasificación a partir de datos nunca antes vistos por la máquina. Una parte que sigue siendo importante, pese a que parezca que la máquina es la que realiza todo el análisis, es el aspecto humano, ya que se ha de especificar al algoritmo tanto la forma de representación de datos, como las posibles variables que se pueden manipular para la caracterización de dichos algoritmos. A partir de la combinación de los datos obtenidos de las imágenes de resonancia magnética Diffusion Tensor Imaging (DTI) y test específicos de cada patología, se desea proveer de un ratio de clasificación alto y estable que permita un apoyo al diagnóstico de los especialistas, ya que estos algoritmos son capaces de aprender patrones que el ojo humano no es capaz.
4. Actualmente, la clasificación de los pacientes se realiza mediante diversos test psicológicos, visitas, fármacos ingeridos, etc. Mediante esta investigación se pretende que los resultados obtenidos permitan converger a un conjunto de características (obtenidas de las imágenes de resonancia magnética o a través de test) que permitan realizar la clasificación de los sujetos.

Por último, resaltar que el uso de los métodos utilizados para el pre-procesado de las imágenes de resonancia magnética y posterior desarrollo de métodos de Inteligencia Artificial, no se limitan únicamente al uso en el ámbito de la dislexia y migraña. Estos pueden aplicarse a cualquier otro estudio tanto de MRI (Imagen de Resonancia Magnética), como de EEG (Electroencephalography), etc. Pero además cualquier tipo de dato, proveniente del ámbito que sea, no sólo médico puede utilizarse para el análisis de métodos clasificatorios de Inteligencia Artificial.

Por tanto, la investigación realizada, como el prototipo desarrollado, algoritmos de extracción de características y de clasificación, pueden ser utilizados en otras investigaciones.

1.2 Hipótesis

La justificación de la investigación muestra las razones más relevantes para la realización de esta tesis. Esto permite formular la siguiente hipótesis:

Es posible desarrollar un método de clasificación de patologías cerebrales a través del análisis de imágenes de resonancia magnética basado en tractografía, mediante métodos de Machine Learning y Deep Learning.

1.3 Objetivos

Una vez formulada la hipótesis, se determina el objetivo principal y los objetivos secundarios. El objetivo principal de la investigación es el siguiente:

Clasificar patologías cerebrales mediante el análisis de imágenes de resonancia magnética a través de algoritmos de Machine Learning y Deep Learning.

A partir del objetivo principal, se extraen los siguientes objetivos secundarios:

- OS1: Extraer características de las imágenes de resonancia magnética y de test específicos para cada patología.
- OS2: Diseñar un método que extraiga las características más relevantes a partir de las iniciales obtenidas de las imágenes de resonancia magnética y de los test específicos.

1. INTRODUCCIÓN

- OS3: Diseñar un método que provea la clasificación de sujetos en función de su patología.
- OS4: Validar los algoritmos de Machine Learning y Deep Learning aplicados a las características obtenidas de las imágenes de resonancia magnética.
- OS5: Validar la confiabilidad de los algoritmos de Machine Learning y Deep Learning tras la selección de características más relevantes a partir de este tipo de algoritmo.
- OS6: Validar la confiabilidad de los algoritmos de Machine Learning y Deep Learning aplicando el método “ensemble” a partir de los resultados obtenidos por cada uno de los algoritmos.

1.4 Metodología de la investigación

La metodología utilizada en esta tesis se basa en el tipo de investigación-acción [L⁺92], específicamente en el modelo de Whitehead [Lat03]. Del término investigación-acción, se puede considerar que la palabra investigación se refiere a una mejora a la hora de entender la problemática que se quiere tratar; mientras que la palabra acción se refiere al cambio necesario que se debe poner en práctica para la obtención de unos resultados que aclaran la fase de la investigación.

Este método de investigación-acción es un ciclo tanto de investigación como de reflexión, siendo un proceso cíclico de exploración, actuación y valoración de los resultados. Durante este proceso cíclico se deben formular una serie de preguntas, las cuales se tienen que ir contestando a medida que se avanza en el ciclo.

En este modelo, la investigación se desglosa en los siguientes puntos:

- Identificar el problema. Este proceso empieza con la identificación de una pregunta o problema que se quiere abordar. Es necesario aprender a desarrollar las preguntas correctas que ayuden en el proceso de la investigación, permitiendo acotar el alcance, ya que en caso contrario la acción no será efectiva.

- Desarrollo de un plan de acción. Una vez que se ha identificado el problema, es importante tener claro el cómo se quiere resolver. Para ello se accederá a la literatura necesaria para obtener información tanto de técnicas y estrategias validadas previamente para poder abordar el problema. Por lo que este plan de acción, muestra el cómo se va a implementar la solución y cómo se abordan las variables definidas.
- Recopilación de datos. Antes de llevar a cabo el plan de acción, es necesario determinar los datos que ayuden a validar con éxito o no el propio plan de acción, además de decidir el cómo se van a recopilar.
- Analizar datos y extraer conclusiones. Una vez se han obtenidos los datos, estos mismos deberán ser analizados para poder extraer conclusiones. Estas conclusiones se confirmarán mediante el análisis de los propios datos.
- Modificar la práctica y repetir el ciclo. Tras la obtención de los resultados, se debe volver a empezar el ciclo, ajustando la teoría mentada en el inicio, precisándola y obteniendo nuevos datos que ayuden a validarla.
- Informar de los resultados. Este punto es el indispensable para cualquier investigación que se esté realizando. Permite una exposición a la comunidad científica para posibles futuros planteamientos del problema que se ha abordado, motivando a otros investigadores a ampliar dicha investigación.

1.5 Estructura del documento

Esta tesis doctoral se encuentra organizada en los capítulos enumerados a continuación:

- Capítulo 1. Introducción. En este primer capítulo se describe una primera aproximación de la investigación realizada en dicha tesis. Por lo que se describe primeramente la justificación de la hipótesis, la propia hipótesis, los objetivos necesarios para la ejecución de la tesis, y por último la metodología de la investigación.

1. INTRODUCCIÓN

- Capítulo 2. Estudio del estado del arte. En este capítulo se presenta un conjunto de conceptos y definiciones necesarios para enmarcar la temática de esta tesis doctoral. Posteriormente, se describen en detalle tanto los antecedentes médicos y técnicos evaluados previamente al desarrollo de esta investigación. Y por último se realiza una comparativa de los antecedentes más relevantes, enumerando en las conclusiones de este capítulo la necesidad existente en este campo de investigación.
- Capítulo 3. Base de datos. En este capítulo se expone el conjunto de imágenes de resonancia magnética (RM) necesarios para la realización de los dos estudios de esta tesis. Se detallan el número de sujetos, grupos, patologías, el tipo de imágenes y parámetros específicos.
- Capítulo 4. Diseño. Este capítulo especifica los métodos seguidos en esta tesis tanto en alto nivel como a bajo nivel. Para ello, en primer lugar se define el pre-procesado necesario para la extracción de características de las imágenes de resonancia magnética. Seguidamente se explican los métodos necesarios para una extracción de características, las más relevantes para la correcta diferenciación entre los grupos de cada patología. En tercer lugar se detallan los algoritmos de clasificación de Machine Learning aplicados para la clasificación de los sujetos. Posteriormente, al igual que en el bloque anterior, se explican los algoritmos en esta ocasión de Deep Learning. Y finalmente se detalla la realización de un método de “ensemble” o unión de los resultados de la clasificación tanto de los algoritmos de Machine Learning como de Deep Learning, para alcanzar un mayor ratio de acierto en la clasificación de los grupos de los dos estudios.
- Capítulo 5. Resultados. En este quinto capítulo se muestran los resultados obtenidos en cada uno de los métodos propuestos en el capítulo anterior. Además se detallan los resultados parciales generados y su consecuente evaluación.
- Capítulo 6. Conclusiones. Por último, en este capítulo se reúnen las conclusiones obtenidas durante el transcurso de esta tesis, junto a una discusión de

1.5 Estructura del documento

los propios resultados obtenidos, destacando los más relevantes. Finalmente se exponen las líneas futuras que deben llevarse a cabo para una mayor profundización en el análisis de los algoritmos de Machine y Deep Learning para una mejora en la riqueza de la clasificación de los grupos bajo estudio.

Estudio del estado del arte

Este capítulo contiene la definición de los conceptos y términos necesarios para la comprensión de la tesis, seguidamente del estudio de los antecedentes técnicos y médicos que se han dado durante estos últimos años.

El resto del capítulo se encuentra organizado de la siguiente manera. La sección 2.1 se enumeran los conceptos y definiciones básicas que se deben tener en cuenta en la presente tesis. La sección 2.2 muestra los antecedentes técnicos, especificando las técnicas de Inteligencia Artificial, entrando en detalle de algoritmos de Machine Learning y Deep Learning. La sección 2.3 se detallan los antecedentes médicos existentes relacionados con el procesamiento de imágenes de resonancia magnética, específicamente DTI, además de la resumir la patología de Dislexia y Migraña. La sección 2.4 se realiza una revisión de investigaciones realizadas previamente a esta tesis. Y para finalizar, la sección 2.5 se incluye un sumario del capítulo, detallando las conclusiones extraídas del estudio del arte analizado.

2.1 Conceptos y definiciones

En esta primera sección se indican algunos conceptos y definiciones acerca de las imágenes DTI a través de la resonancia magnética (MRI) y sobre las técnicas de aprendizaje automático. De esta manera se trata de establecer unos conocimientos

2. ESTUDIO DEL ESTADO DEL ARTE

previos básicos que doten de sentido al ámbito de la investigación llevada a cabo en esta tesis, previo a la lectura completa de este documento.

2.1.1 Resonancia magnética e imágenes DTI

Dentro de la investigación médica, la rama dedicada a la neurología que estudia el cerebro ha tenido grandes avances durante los últimos tiempos.

El órgano del cerebro es el más grande dentro del sistema nervioso central, y el que se encarga de controlar todo el cuerpo. Pero además tiene otras funciones muy importantes tales como las emociones, el lenguaje, la memoria, etc. Pese a que ha habido grandes avances, sigue siendo una de las ramas menos evolucionadas dentro de la investigación, desconociendo aspectos sobre el propio cerebro.

El cerebro está formado por diversas zonas, áreas o regiones, las cuales de manera independiente tienen funciones específicas, teniendo conexiones entre las propias zonas de interés y transportando la información necesaria, lo que conforma una red neuronal.

Dentro del estudio de la neurociencia, se han hecho avances respecto a las relaciones existentes entre las enfermedades y las lesiones cerebrales que pueda tener el paciente. Esta rama de investigación es posible gracias a las máquinas de resonancia magnética, las cuales pueden analizar las materias blancas y grises del cerebro, en las cuales se puede obtener información para detectar posibles enfermedades en el paciente [ONWK03] .

Específicamente las imágenes DTI (Diffusion Tensor Image), son las cuales recogen visualmente las fibras nerviosas en la materia blanca del cerebro, todo ello a través del análisis del movimiento de las moléculas de agua a través de estas fibras. Sobre todo se puede observar y evaluar el daño cerebral derivados de diferentes enfermedades o traumatismos [PWP⁺15, TMK⁺17].

2.1.2 Inteligencia Artificial

Dentro de las revoluciones tecnológicas habidas en los últimos años ha sido la provocada por la Inteligencia Artificial, siendo la encargada de que las máquinas u ordenadores piensen y razonen por su propia cuenta.

Aunque no existe una definición aceptada por todos los expertos, ya que es una nueva ciencia y continuamente sometida a cambios y experimentos, pero se puede determinar como el intento de imitar la inteligencia humana [Jor19] .

De esta nueva ciencia de la Inteligencia Artificial, deriva el aprendizaje automático, popularizado como Machine Learning y Deep Learning que ha aparecido en los últimos años. Los algoritmos dedicados al aprendizaje automático son los encargados de encontrar patrones y descubrir nuevos a través de los datos, gracias a tareas automáticas que pueden resolver problemas de la vida real en cuestión de segundos gracias a su alta capacidad de computación [Weh12] .

Una de las misiones de estos algoritmos de aprendizaje automático, es la clasificación de clases, siendo capaces de encontrar patrones en los datos que caracterizan a un grupo de individuos u objetos de manera sencilla y rápida [BRH04] , facilitando la labor del humano, incluso llegando a evitar los posibles errores humanos.

2.1.3 Términos estadísticos

La estadística es la ciencia que va a apoyar esta tesis, consiguiendo confirmar la hipótesis a partir de los datos obtenidos tras la clasificación automática de los sujetos.

Los términos estadísticos que se van a utilizar para poder medir de manera cuantitativa los resultados obtenidos en esta tesis, son los que se presentan en las siguientes secciones [GG05] .

Para ello, se tiene en cuenta la matriz de confusión, la cual es una herramienta que permite evaluar los algoritmos de aprendizaje supervisado [Arc] .

2.1.3.1 Accuracy

El término de Accuracy, o también llamada verdadero positivo, es el más simple de todos, ya que mide la tasa de acierto en la predicción. Se representa por la proporción entre los positivos reales predichos por el algoritmo y todos los casos positivos.

2. ESTUDIO DEL ESTADO DEL ARTE

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

Figura 2.1: Matriz de confusión.

Fuente Juan Ignacio Barrios Arce

2.1.3.2 Precision

La medida de Precision, es la que determina el porcentaje de casos positivos detectados, es decir, la proporción entre el número de predicciones correctas y el total de predicciones.

2.1.3.3 Recall

El término Recall, también llamado Sensitivity, se corresponde con la tasa de verdaderos positivos, es decir, la proporción de casos positivos son correctamente identificados por el algoritmo de aprendizaje automático.

2.1.3.4 F1 score

Este término, combina en una sola métrica los valores de Precision y Recall, gracias al cual se puede medir la eficacia del algoritmo en la detección de los casos positivos y negativos. Por ejemplo un valor alto de F1 score indica que se detectan correctamente las clases.

2.2 Antecedentes tecnológicos

En esta sección de la tesis, se presentan los antecedentes existentes desde el punto de vista técnico. Se detallan los conceptos de selección de características y algoritmos de Machine Learning y Deep Learning, de los cuales se compone esta tesis.

2.2.1 Selección de características

La selección de características es el método por el que se seleccionan las características más importantes y relevantes dentro de un conjunto de datos. El objetivo de dicho método es el de mejorar el rendimiento de los algoritmos de predicción, tanto a nivel de tasa de acierto como de rapidez.

En las siguientes secciones se indican los algoritmos de selección de características que se han utilizado en esta tesis.

2.2.1.1 Gradient Tree Boosting

El algoritmo Gradient Tree Boosting o también llamado Gradient Boosting o Potenciación del gradiente, es una técnica que se utiliza para cuando existen errores en la clasificación [RSR⁺19] .

La característica de este algoritmo es la construcción del modelo de manera escalonada, y que gracias a la generalización permite la optimización arbitraria en función de una pérdida diferenciable.

Gradient Tree Boosting funciona como clasificador [LYZ⁺17] , aunque contiene un atributo gracias al cual se pueden obtener las características más relevantes en base a la propia clasificación que ejecuta.

Para ello, se realiza la llamada al algoritmo y el atributo específico que contiene una función de decisión [PLD17] . Gracias a la cual, y en función de la clasificación que se haya especificado de los grupos del conjunto de training, devuelve un nuevo conjunto de training pero únicamente con las características que ha considerado significativas para que la clasificación posterior del conjunto de test sea la más óptima.

Los parámetros son los siguientes:

2. ESTUDIO DEL ESTADO DEL ARTE

- Número de estimadores: es el número de iteraciones a realizar por el clasificador. Siempre un aumento de este valor genera un mejor rendimiento, ya que tiene más posibilidades de mejorar la clasificación.
- Ratio de aprendizaje: gracias a esta tasa de aprendizaje se reduce la contribución de cada árbol. Este ratio decrece en función del tiempo y permite que en cada iteración del algoritmo mejore el ratio de clasificación.
- Máxima profundidad: es la profundidad máxima de los estimadores de regresión. Es el valor que limita el número de nodos en el árbol.
- Estado: es la semilla utilizada por el propio algoritmo, provocando que siempre que se haga una llamada al algoritmo cree el mismo estado y se pueda reproducir la ejecución. En cambio, si no se indica, cada vez el algoritmo seleccionará un estado, pudiendo provocar diferentes resultados cada vez que se ejecuta.

2.2.1.2 Random Forest

Esta técnica también llamada Bosques Aleatorios, siendo una mejora sustancial a las técnicas de Árboles de Decisión [ZZH08] . El motivo de la mejora frente a las técnicas de árboles simples, es que se combina una gran cantidad de árboles de decisión independientes, probados cada uno de ellos y de manera independiente sobre conjuntos de datos aleatorios con la misma distribución [LWC⁺14] .

Este algoritmo hace uso de dichos árboles para mejorar la precisión en la predicción y controla el sobreajuste [HZNF07] , es decir, el efecto que puede tener sobreentrenar un algoritmo con ciertos datos para los que se conoce el resultado deseado. El propio algoritmo debe alcanzar un estado en el que es capaz de predecir el resultado de un conjunto nunca visto a partir de lo aprendido durante el entrenamiento. Pero cuando el modelo se sobreentrena, el algoritmo puede quedar ajustado a unas características muy específicas de los propios datos de entrenamiento, haciendo que empeore la predicción durante la fase de test.

En la Figura 2.2, se puede observar ver el sobreajuste en un modelo. La línea azul muestra el error de entrenamiento, mientras que la línea roja muestra el error

en la validación. Si el error de validación incrementa, mientras que el del entrenamiento sigue decreciendo, se está dando dicha situación. Para que no exista esta situación, siempre deben ir decreciendo paralelamente tanto el error en el entrenamiento como el error en validación.

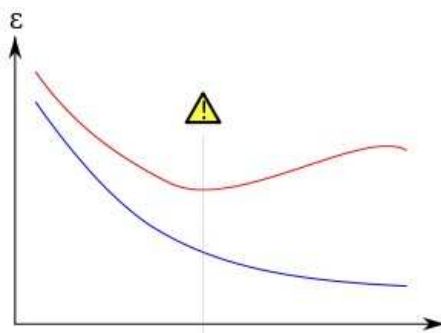


Figura 2.2: Sobreajuste en un modelo.

Fuente Wikipedia

Los parámetros que se han ajustado [PWB19] para conseguir las características más importantes son:

- Número de estimadores: es el número de iteraciones a realizar por el clasificador. Siempre un aumento de este valor genera un mejor rendimiento, ya que tiene más posibilidades de mejorar la clasificación.
- Máxima profundidad: es la profundidad máxima de los estimadores de regresión. Es el valor que limita el número de nodos en el árbol.
- Estado: es la semilla utilizada por el propio algoritmo, provocando que siempre que se haga una llamada al algoritmo cree el mismo estado y se pueda reproducir la ejecución. En cambio, si no se indica, cada vez el algoritmo seleccionará un estado, pudiendo provocar diferentes resultados cada vez que se ejecuta.
- Mínimo de muestras: es el mínimo número de muestras requeridas para poder dividir un nodo.
- Mínimo de muestras en la hoja: es el mínimo número de muestras requeridas que pueden estar en la hoja.

2. ESTUDIO DEL ESTADO DEL ARTE

2.2.1.3 Univariate feature selection

El algoritmo llamado Univariate feature selection, permite seleccionar las mejores características del conjunto de muestras basándose en el test estadístico univariado de datos [TMA14] .

Este test estadístico permite el análisis de cada una de las variables estudiados por separado, por lo que se encuentra basado en el análisis de una única variable o característica. Dentro de este tipo de test estadístico, hay diferentes técnicas las cuales se pueden aplicar para conseguir el hito de obtener las características más relevantes.

Las diferentes técnicas del análisis univariado de datos son:

- K: número de las características que se quieren seleccionar.
- Función de puntuación: esta función devuelve el listado de las puntuaciones de las características, seleccionando las que hayan tenido una mejor puntuación hasta el número K, el cual indica el máximo que se quieren seleccionar.

Esta función de puntuación tiene diferentes métodos para calcular dicha puntuación teniendo en cuenta la técnica de selección de las K mejores características:

- ANOVA: se tiene en cuenta la varianza de las características.
- Información mutua: se mide la dependencia de dos variables.
- Chi cuadrado: también llamada ji cuadrado, en la que se mide la distribución de probabilidad continua de las variables.
- Regresión: en este test se mide los errores y divergencias del propio algoritmo, descartando en las diferentes iteraciones las características que empeoran el rendimiento del mismo.

En cambio las posibilidades cuando se aplica la técnica de selección mediante percentiles son:

- Percentil: teniendo en cuenta que se ordenan las puntuaciones de las características de menor a mayor, es el valor de la variable por debajo del cual se encuentra el porcentaje de dichas puntuaciones.

- Ratio de falsos positivos: se tiene en cuenta los falsos positivos resultantes de la clasificación del algoritmo para la selección de las características significativas.
- Ratio de error: al igual que en el método anterior, en esta ocasión se tiene en cuenta el ratio de error que genera la clasificación con unas u otras características.

2.2.1.4 Tree-Based

Este algoritmo es otro de los que se encuentran basados en árboles de decisión, al igual que Random Forest. Es un estimador que ajusta a una serie de árboles de decisión aleatorios [MG09], también llamados extra-trees, seleccionando varias submuestras del conjunto de datos y haciendo uso del promedio para mejorar la precisión en la predicción y de nuevo el sobreajuste u over-fitting.

Se han incluido, junto al Random Forest, dos técnicas que hacen uso de los árboles de decisión debido a que cada uno se ha configurado de manera diferente, y son eficientes a la hora de seleccionar las características más importantes para conseguir el mejor ratio de clasificación posible [ZP15].

De nuevo, para que este algoritmo se convierta en un selector de características, se hace uso del atributo de función de decisión, para que en su resultado retorne el listado.

Los parámetros que se han ajustado para conseguir las características más importantes son:

- Número de estimadores: es el número de iteraciones a realizar por el clasificador. Siempre un aumento de este valor genera un mejor rendimiento, ya que tiene más posibilidades de mejorar la clasificación.
- Máxima profundidad: es la profundidad máxima de los estimadores de regresión. Es el valor que limita el número de nodos en el árbol.
- Estado: es la semilla utilizada por el propio algoritmo, provocando que siempre que se haga una llamada al algoritmo cree el mismo estado y se pueda

2. ESTUDIO DEL ESTADO DEL ARTE

reproducir la ejecución. En cambio, si no se indica, cada vez el algoritmo seleccionará un estado, pudiendo provocar diferentes resultados cada vez que se ejecuta.

2.2.1.5 L1-based feature selection

Este algoritmo pertenece a la clase de los modelos lineales, tales como SVC (Support Vector Classification), modelos de regresión lineal, Lasso, etc.

Este algoritmo se caracteriza por hacer uso de la geometría del taxista (también llamada distancia L1), reemplazando la métrica euclidiana por una nueva en la que la distancia entre dos puntos es la suma de las diferencias absolutas de sus coordenadas. Por tanto, las características se representan en un espacio para poder hacer el cálculo de dicha geometría del taxista.

Aplicando esta distancia L1 [WZZ⁺16], específicamente se ha seleccionado entre otros, el algoritmo SVC lineal. Se caracteriza por tener más flexibilidad en la elección de las penalizaciones y funciones de pérdida [HLM⁺19], escalando mejor a mayor número de muestras. Además también es compatible con datos densos como dispersos, además de soportar la clasificación multiclase (en un esquema de uno contra el resto).

De nuevo, se hace uso de este algoritmo de clasificación, para poder obtener las características más relevantes.

La parametrización que se ha llevado a cabo para este caso es la siguiente:

- C: este parámetro mide el cuánto quiere evitar errores en la clasificación del conjunto de entrenamiento. A mayor valor de C, el espacio en el que se encuentren las características tendrá menor margen, haciendo que todos las muestras del entrenamiento se clasifiquen correctamente. En cambio, si se selecciona un valor pequeño, el espacio será mayor provocando en ocasiones que existan mayores errores de las muestras de entrenamiento, provocando tanto falsos positivos como falsos negativos.
- Penalización: En este parámetro se selecciona el tipo de penalización L1, o L2 (también llamada distancia euclidiana, no utilizada en esta tesis).

- Dual: en este parámetro se selecciona el algoritmo para resolver posibles problemas de optimización. Los valores posibles son Verdadero o Falso, indicando Falso si hay mayor número de muestras que características.

2.2.2 Machine Learning

Como ya se ha comentado previamente, en la sección de conceptos sobre la Inteligencia Artificial, la disciplina de Machine Learning es una rama dentro de este universo, que está dedicada a la identificación de patrones complejos en millones de datos, bien sea para la clasificación o predicción a futuro de dichos patrones [Anz12].

Dentro de este grupo de algoritmos se engloban tres grupos:

- Aprendizaje supervisado: se utiliza en los casos en los que ya se tienen previamente etiquetadas las muestras, generando un subconjunto de entrenamiento. Son capaces de obtener patrones y propiedades de las características que definen las muestras, que no son fáciles de extraer sin este tipo de técnicas [CNM06].
- Aprendizaje no supervisado: este otro grupo de algoritmos son adecuados para descubrir relaciones implícitas dentro del conjunto de observaciones, pero que a priori son desconocidas, no estando etiquetadas previamente [Zhu05].
- Aprendizaje por refuerzo: este último grupo de algoritmos son una mezcla entre el aprendizaje supervisado y el no supervisado. En esta ocasión también se indican las etiquetas de clasificación ya conocidas [SHS⁺17]. Durante las iteraciones en las que los algoritmos buscan obtener los patrones, si se genera una clasificación correcta, aprende automáticamente a repetir esa decisión en el futuro. Mientras que si la clasificación es incorrecta, evita tomar dicha decisión.

2.2.2.1 SVM

Support Vector Machine es un algoritmo supervisado para la clasificación, regresión y detección de outliers. SVM permite clasificación multiclase [SFJI20]. Específicamente se ha hecho uso de la variante del algoritmo C-SVM.

2. ESTUDIO DEL ESTADO DEL ARTE

La parametrización de la validación cruzada es la siguiente:

- Número de división: es el número que indica la cantidad de combinaciones que tiene que generar tanto del conjunto de entrenamiento como el de test.
- Tamaño de test: es la proporción del conjunto total de muestras en la que se tiene que dividir para poder generar el conjunto de entrenamiento y test.
- Estado: es la semilla utilizada por el propio algoritmo, provocando que siempre que se haga una llamada al algoritmo cree el mismo estado y se pueda reproducir la ejecución. En cambio, si no se indica, cada vez el algoritmo seleccionará un estado, pudiendo provocar diferentes resultados cada vez que se ejecuta.

La parametrización para el algoritmo de aprendizaje supervisado SVM [HHW⁺20] es la siguiente:

- Kernel: Se especifica el tipo de kernel que utiliza el algoritmo. Ya que los datos a estudiar no presentan casos idílicos, y los algoritmos se tienen que enfrentar a datos con más de dos variables predictoras, o datos no lineales, se hace uso de esta función kernel. Permite proyectar los datos en un espacio de características de mayor dimensión, en la que aumenta la capacidad computacional de los algoritmos lineales.
- C: este parámetro mide el cuánto quiere evitar errores en la clasificación del conjunto de entrenamiento.

2.2.2.2 Adaboost

Este segundo algoritmo es del grupo de aprendizaje por refuerzo, por lo que es una mezcla de supervisado y no supervisado. También se le incluye las etiquetas de las observaciones del conjunto de entrenamiento [WJWS19] .

Este algoritmo, en su primera fase ajusta el clasificador con el conjunto de datos de entrenamiento, y posteriormente ajusta copias adicionales del clasificador sobre el mismo conjunto de datos. Pero en estas copias adicionales, los pesos de las muestras mal clasificadas, se ajustan de tal manera que los clasificadores posteriores se centren en esos casos difíciles de clasificar.

En este aspecto, se observa perfectamente el aprendizaje por refuerzo, ya que los casos correctamente clasificados, los toma en cuenta para el futuro, pero en los casos incorrectamente clasificados los pasa a la siguiente capa en la que hay una copia del mismo clasificador, indicándole los casos que ha clasificado de manera incorrecta.

La parametrización [LGF20] para el algoritmo Adaboost es la siguiente:

- Número de estimadores: es el número de iteraciones a realizar por el clasificador. Siempre un aumento de este valor genera un mejor rendimiento, ya que tiene más posibilidades de mejorar la clasificación.
- Estimador base: es el estimador a partir del cual se construye el conjunto de reforzamiento.
- Algoritmo: se indica el algoritmo que va a utilizar Adaboost para realizar la clasificación.

2.2.2.3 Naïve Bayes

Este último algoritmo del bloque de Machine Learning, es el llamado Naïve Bayes, el cual pertenece al grupo de aprendizaje supervisado.

Está basado en el teorema de Bayes con el supuesto “naïve” o “ingenuo” de independencia condicional entre cada par de características dada la etiqueta de clasificación. El teorema de Bayes expresa la probabilidad condicional de un evento aleatorio “X” dado “Y” con la probabilidad de “Y” dado “X” [NJF19].

Por tanto, en este algoritmo, aplicándose dicho teorema, se estudian las diferentes combinaciones de pares de características que determinan el conjunto de datos, y determina la probabilidad de que puedan pertenecer a la misma clase o no.

2.2.3 Deep Learning

Deep Learning, también llamado aprendizaje profundo, es una variante del aprendizaje automático que está basado en arquitecturas de redes neuronales. El funcionamiento es muy parecido al que pueda tener el cerebro humano, ya que se compone de redes neuronales artificiales [ZYST19].

2. ESTUDIO DEL ESTADO DEL ARTE

Los algoritmos vistos hasta ahora hacen uso de análisis con estructuras lineales, pero los algoritmos pertenecientes a Deep Learning, tienen estructuras jerárquicas que permiten procesar datos con un enfoque no lineal.

Otro de los puntos que caracteriza al Deep Learning es la capacidad de aprendizaje sin la intervención humana, puesto que no es necesaria una selección de características previas. El propio algoritmo, gracias al aprendizaje profundo y su enfoque no lineal, es capaz de sacar los patrones necesarios [YZW19] para caracterizar y etiquetar a una muestra.

Este tipo de algoritmos alcanzan una mayor eficiencia cuanto mayor es el tamaño de muestras que se le enseña. Adquiere el conocimiento del conjunto de datos por sí solo, siendo capaz de procesar mediante la red neuronal datos en bruto de cualquier tipo, desde datos numéricos hasta imágenes.

El porqué del término “deep” se debe al número de capas ocultas en la red neuronal, en la que puede albergar 150 capas, mientras que las redes neuronales tradicionales suelen contener 2 o 3 capas ocultas.

Dentro del Deep Learning, existen diferentes arquitecturas siempre formadas por modelos basados en redes neuronales.

- Deep Neural Network (DNN) o Redes Neuronales Profundas [BJSJ19] .
- Convolutional Neural Network (CNN) o Redes Neuronales Profundas Convolucionales [Zho20] .
- Deep Belief Network (DBN) o Redes de Creencia Profunda [HAU⁺19] .

2.2.3.1 Tensorflow

Tensorflow [Ten] es una biblioteca de código abierto desarrollado por Google para poder entrenar algoritmos de Redes Neuronales dentro de la metodología de Deep Learning [Gér19] . Su nombre deriva de las operaciones que las redes neuronales realizan sobre arrays multidimensionales de datos, que son referidos como “tensores”.

Esta librería es la utilizada para poder desarrollar el algoritmo de Deep Learning de esta tesis.

2.2.3.2 Convolutional Neural Network

Las Redes Neuronales Convolucionales es una arquitectura [TYO⁺17] en la que se aprenden características de orden superior haciendo uso de la operación de convolución. Este tipo de operación consiste en filtrar los datos usando una máscara. Aplicando diferentes máscaras a los resultados, se consiguen resultados distintos, permitiendo al algoritmo abstraerse de las características que se le muestran en el conjunto de entrenamiento [AOH⁺17] , y determinando unas nuevas reglas que permiten etiquetar a las muestras.

Como ya se ha comentado, las redes neuronales están compuestas de capas, las cuales aprenden progresivamente las características de orden superior sin procesar. Este es un proceso que permite aprender características de manera automática.

Este tipo de redes neuronales convolucionales se forman utilizando dos tipos de capas [WPS⁺18] :

- Convolutiva: esta capa transforma los datos de entrada aplicando la operación de convolución, fusionando conjuntos de información diferentes.
- Pooling: en esta otra capa se resume las respuestas de las salidas cercanas. Esta acción se realiza mediante la reducción progresiva del tamaño de los datos, y agrupando de tal manera que se consiga una representación de los datos invariable.

Como se ha comentado, los métodos de Deep Learning son capaces de realizar un análisis no lineal, mientras que la forma de los conjuntos de datos son lineales. Es necesario adecuar y dar forma al conjunto de datos para poder introducirse al algoritmo.

2.3 Antecedentes médicos

2.3.1 DTI

Actualmente la Diffusion Tensor Imaging (DTI) permite examinar la integridad de las vías de la materia blanca de forma no invasiva a través de cuatro medidas: Anisotropía fraccional (FA), como índice de integridad de la materia blanca;

2. ESTUDIO DEL ESTADO DEL ARTE

Difusividad media (MD) como medida del movimiento molecular promedio y/o el desplazamiento medio cuadrado general de las moléculas independientes de la direccionalidad de los tejidos; Difusividad radial (DR) que corresponde a la difusividad en direcciones perpendiculares al eje principal de difusión y se ha asociado con el grado de mielinización; y Difusividad axial (DA) que corresponde a la difusividad a lo largo del eje principal y se ha asociado con el diámetro del eje [LBMP⁺⁰¹, SYL⁺⁰⁵].

En la presente tesis se ha aplicado la técnica DTI, considerando una de las cuatro medidas mencionadas anteriormente (FA) en las principales zonas de interés determinadas para cada uno de los dos casos de estudio.

2.3.2 Dislexia

Desde una perspectiva evolutiva, la lectura y la escritura son inventos muy recientes, en contraste con el lenguaje. La lectura es un proceso complejo que implica varios procesos fonológicos, ortográficos, sintácticos y semánticos que permiten interpretar la temática de la información escrita.

La dislexia evolutiva es un trastorno que suele definirse como una dificultad inesperada de la lectura que afecta a niños y adultos sin ninguna otra deficiencia intelectual o falta de motivación y que tienen suficiente acceso a la educación. Según el estudio, se ha registrado una prevalencia excesiva que varía entre el 5 % y el 10 % de la población [SSP⁺⁹⁸]. Así pues, la dislexia es una discapacidad específica del aprendizaje que se ha establecido como un trastorno neurobiológico de origen genético [Hab00].

La dislexia es un trastorno primario de la lectura de alta prevalencia que varía en función de la transparencia ortográfica. Hoy en día, el aprendizaje y la educación se basan en la lectura, por lo que cualquier alteración de la lectura impide el correcto desarrollo educativo y el conocimiento. Esto significa que un gran número de niños, padres, maestros y médicos tienen que tratar este importante y complejo problema, por lo que una mejor comprensión de las causas subyacentes de esta perturbación facilitaría una pronta reparación y una mejora en el desarrollo del lenguaje y el aprendizaje de estos niños.

Esta patología es un trastorno neurológico de origen genético que se está investigando actualmente. A pesar de las intensas investigaciones, las causas biológicas y cognitivas subyacentes de la dificultad para leer siguen siendo objeto de debate. Aunque hoy en día existe un gran consenso en que la base de la dislexia es un déficit de procesamiento fonológico, se han postulado algunas otras teorías importantes para explicar las causas de esta deficiencia.

Existen diversas teorías tales como la Teoría Fonológica [PDF⁺01, SH11] , que sostiene que un déficit en el sistema del lenguaje es la causa de estas dificultades. Otra la llamada Teoría Visual [LRDG91] señala que el procesamiento visual en los disléxicos puede fallar debido a una interrupción en la vía magnocelular, aunque no se rechazan otras deficiencias fonológicas. Las últimas teorías relacionadas con esta hipótesis hacen hincapié en un déficit que afectaría a los mecanismos de atención controlados por el flujo visual dorsal [VP10, ZBF⁺12] .

Hasta la fecha se han realizado muchos estudios con el objetivo de identificar el sistema neurológico que permite la lectura de palabras [FP98] , y se ha demostrado la viabilidad de detectar las áreas responsables de esta complicada tarea con la neuroimagen. El dominio de las habilidades de lectura precisa y fluida requiere la activación de una red neuronal compleja que involucra la colaboración de regiones corticales distantes además de una comunicación adecuada entre las áreas. Estudios previos han mostrado especial interés en la conectividad dentro de la materia blanca que permite que haya conexiones entre las áreas corticales elocuentes de la red de lectura [LBCR⁺13] .

2.3.3 Migraña

La migraña es una condición incapacitante que afecta al 12 % de los población mundial, con un 1,4-2,2 % de la población que sufren de su forma crónica (es decir, dolor de cabeza durante 15 o más días al mes durante al menos 3 meses) [MG13] .

De acuerdo con el Estudio de la Carga de la Enfermedad 2013, la migraña fue la sexta causa de discapacidad en todo el mundo; cuando el dolor de cabeza por exceso de medicación está incluido en la encuesta, los trastornos del dolor de cabeza ocupan el tercer lugar entre los causas de la discapacidad [SBJ⁺15] .

2. ESTUDIO DEL ESTADO DEL ARTE

La migraña episódica supera a la migraña crónica en un 2,5 % por año, siendo los factores de riesgo tales como una alta tasa de ataques, medicación excesiva, obesidad, depresión, así como el estrés de la infancia [TBF⁺15] .

Aunque no hay pruebas bien establecidas de que la migraña o el dolor de cabeza por el uso excesivo de medicamentos representan un riesgo en los factores de deterioro cognitivo a lo largo del tiempo [TTT⁺13] , poco ha se ha investigado sobre el papel de la función cognitiva en el dolor de cabeza. Sin embargo, es ampliamente aceptado que el dolor es modulado por la cognición.

La reserva cognitiva (RC) es un concepto recientemente introducido, que se refiere a las diferencias de rendimiento en el procesamiento cognitivo que hacen que la persona sea más propensa a mantener funcionamiento cognitivo a pesar de la enfermedad o el daño [SBH⁺14] . Niveles de educación más altos, complejidad ocupacional, y/o la inteligencia pre-mórbida están asociadas con una menor niveles de deterioro cognitivo de lo que se esperaría de una determinada patología cerebral [Rob14] .

El concepto de RC se había aplicado principalmente a la demencia pero hay diferentes estudios sobre el efecto protector de CR sobre el curso y el pronóstico de otras enfermedades nerviosas como esclerosis múltiple [MDSCM⁺15] , o abuso de drogas [PPRMdL⁺14] .

Por otro lado, también se sabe que el dolor crónico puede inducir cambios anatómicos en las regiones del cerebro relacionadas con la percepción y el control del dolor [BČL13] . Su disfunción en pacientes con dolor crónico, incluyendo a los migrañosos, puede contribuir a disminuir el dolor modulación y al comportamiento de la adicción a los analgésicos [RCDS⁺13] .

Estas regiones de interés (ROI) también fueron elegidas porque son áreas en las que el dolor crónico y la cognición reserva puede transmitir y porque otros estudios han demostrado su deterioro en la migraña.

2.4 Revisión de investigaciones

A partir de la literatura recogida, a continuación se expone la revisión de los artículos consultados. Esta revisión se centra en las técnicas de Machine Learning y Deep Learning.

Tabla 2.1: Tabla publicaciones Machine Learning en DTI

Año	Patología	Algoritmo	Ratio de Acierto	N sujetos	Referencia
2019	Niños con desorden sensorio	Naive Bayes	71 %	41	[PPO ⁺ 19a]
2018	Alzheimer	SVM	92 %	213	[ZL18]
2018	Mielopatía	SVM	95 %	70	[WPS ⁺ 18]
2019	Glaucoma	Método ensemble	77 %	44	[QWC ⁺ 19]
2019	Daño por radioterapia	SVM	84 %	144	[LFL ⁺ 19]
2020	Alzheimer	SVM	77 %	77	[CCR ⁺ 20]
2019	Exceso de respuesta auditiva	SVM	73 %	39	[PPO ⁺ 19b]
2018	Esquizofrenia	SVM	73 %	126	[MHS ⁺ 18]
2018	Parkinson	Método ensemble	96 %	90	[RMM ⁺ 18]

Tabla 2.2: Tabla publicaciones Deep Learning en DTI

Año	Patología	Algoritmo	Ratio de Acierto	N sujetos	Referencia
2020	Esclerosis múltiple	Deep Neural Network	93 %	38	[YGW ⁺ 20]
2020	Problemas cognitivos vasculares	CNN	98 %	242	[CWQ ⁺ 20]
2018	Lesión traumática cerebral	CNN	84 %	58	[CWZ ⁺ 18]
2018	Alzheimer	Directed Acyclic Graph	95 %	68	[QLC ⁺ 18]
2019	Interacción con la droga	CNN	80 %	No disponible	[LKN19]

2.4.1 Machine Learning

Dentro de las publicaciones revisadas dentro de la metodología de Machine Learning para clasificación de sujetos con imágenes DTI, en la Tabla 2.1 se exponen los diversos métodos utilizados más relevantes en los dos últimos años.

2.4.2 Deep Learning

Dentro de las publicaciones revisadas de Deep Learning para clasificación de sujetos con imágenes DTI, en la Tabla 2.2 se exponen los diversos métodos utilizados más relevantes en los dos últimos años.

2.4.3 Conclusiones

Tras el análisis de la literatura, el uso de técnicas de clasificación está siendo comúnmente usado en el campo de la medicina, como apoyo a los profesionales médicos para bien sea como una ayuda a su diagnóstico, o para poder descubrir nuevos patrones en las características que permitan obtener una clasificación más

2. ESTUDIO DEL ESTADO DEL ARTE

exacta. Focalizando en el estudio de esta tesis, existen numerosas publicaciones dedicadas a la clasificación en el campo de la neurología.

La clasificación se puede realizar tanto de las imágenes de un escáner de resonancia magnética, como añadiendo características propias de la patología que se está tratando.

Los métodos más utilizados en la parte de Machine Learning es el algoritmo SVM, mientras que en la parte de Deep Learning se utiliza mayormente el algoritmo de Convolutional Neural Network. Se remarca que en la mayoría de las publicaciones hacen uso de algoritmos supervisados, ya que la revisión de la clasificación por parte de un profesional médico es importante a la hora de conseguir un ratio importante de acierto.

El implementar algoritmos de aprendizaje automático en vez de los métodos convencionales estadísticos, se debe a que se consigue una mayor precisión en el resultado final, como una mayor rapidez a la hora de dictaminar un diagnóstico al paciente. Además gracias a estos algoritmos, se pueden encontrar nuevos patrones que hasta ahora no se habían descubierto con los métodos tradicionales.

También se ha podido observar, el exponencial uso de los algoritmos de Deep Learning, ya que consiguen obtener mejores resultados que con la técnica de Machine Learning, debido a la mejora interna de la composición de la estructura del algoritmo, centrada en unas redes neuronales más profundas, que consiguen llevar el autoaprendizaje a un nivel más exhaustivo. Estas redes son mucho más complejas, por lo que ciertamente se consiguen mejores resultados y la obtención de nuevos patrones, pero también exige un mayor tiempo computacional, además de tener a disposición un hardware potente, ya que el ejecutar estos algoritmos, requiere de una mayor exigencia computacional.

2.5 Sumario

En este capítulo se ha revisado tanto los conceptos relevantes tecnológicos como médicos. Además de entrar de un estudio de la literatura existente sobre las técnicas de aprendizaje automático.

Por tanto, tras el estudio realizado del estado del arte, puede afirmarse que:

El estudio y clasificación de sujetos a partir de la materia blanca del cerebro es posible mediante técnicas de Inteligencia Artificial, dando la posibilidad a la obtención de un método con una mayor confiabilidad, llegando a un ratio elevado de acierto en cuanto a la extracción de características importantes para cada caso de estudio, y a la clasificación de los sujetos de cada una de las patologías estudiadas.

Base de datos

Para la realización de esta tesis, ha sido necesaria la utilización de 2 bases de datos diferentes de imágenes de resonancia magnética, específicamente una variación en la toma de este tipo de imágenes las cuales capturan el recorrido de las fibras nerviosas en la materia blanca del cerebro.

Las 2 bases de datos utilizadas se consideran por igual las principales para esta investigación. Se encuentran enmarcadas dentro de un estudio sobre la dislexia en niños, y en personas con migraña.

Caso de estudio 1: Dislexia

Respecto a la base de datos de niños con dislexia, los sujetos se separan en 3 grupos:

- Sujetos control, los cuales no exhiben ningún signo de que su alfabetización se viera afectada por patologías oculares.
- Sujetos disléxicos, los cuales no habían recibido tratamiento o apoyo psicopedagógico para la alfabetización y cuando su puntuación en la escala Wechsles Intelligence Scale – Four edition (WISC-IV) mostró que su coeficiente intelectual (Intelligence Quotient - IQ) se encontraba dentro del rango normal, (> 75) [Wec03].

3. BASE DE DATOS

- Sujetos con visión monocular, los cuales tienen una capacidad de lectura normal.

Caso de estudio 2: Migraña

Por otro lado, la base de datos correspondiente al estudio de migraña, los grupos se separan de nuevo en 3:

- Sujetos control que no padecen de migraña.
- Sujetos con migraña esporádica.
- Sujetos con migraña crónica y abuso de fármacos.

Destacar que para los 2 estudios, la elaboración de una resonancia magnética es costosa. Por un lado en el estudio de dislexia que se introducen en la máquina niños de 9 a 12 años, lo cual es necesario en primera instancia un permiso por parte de los padres o tutores legales. Además de que los niños se pueden sentir incómodos dentro de la máquina y puede llevar más tiempo la realización de la resonancia.

Y por parte del estudio de migraña, es inviable la elaboración de una resonancia magnética en un paciente con migrañas durante el padecimiento del dolor del sujeto. El instrumental necesario emite un alto nivel de ruido, lo cual es imposible de soportar por parte del paciente.

El resto de este capítulo se encuentra organizado como sigue: La sección 3.1 describe la población de referencia existente en los 2 estudios. La sección 3.2 indica los criterios de selección seguidos para la inclusión y exclusión de sujetos en el estudio. La sección 3.3 expone el tamaño muestral finalmente recogido para el estudio. La sección 3.4 muestra el método utilizado. Finalmente la sección 3.5 resume el capítulo con los sujetos que componen las 2 bases de datos.

3.1 Población de referencia

La población o universo representa a un grupo de individuos los cuales comparten características similares, también llamados parámetros. Derivado de este grupo, se encuentra la población de referencia es definida como la población total que tiene posibilidad de pertenecer a este estudio. De esta población de referencia, se pasa a

3.2 Criterios de selección

la población blanco, la cual es la población a la cual se van a generalizar los resultados. Pasando al siguiente escalón, está la población elegible, la cual es la población concreta a la cual va dirigido el estudio según ciertos parámetros seleccionados. Y por último está la población estudio, siendo una parte de la población elegible que está disponible para la realización del estudio.

Esta población de estudio, también es llamada muestra, la cual permite inferir o extrapolar los resultados de las observaciones de la población elegible según los criterios de inclusión y exclusión seleccionados.

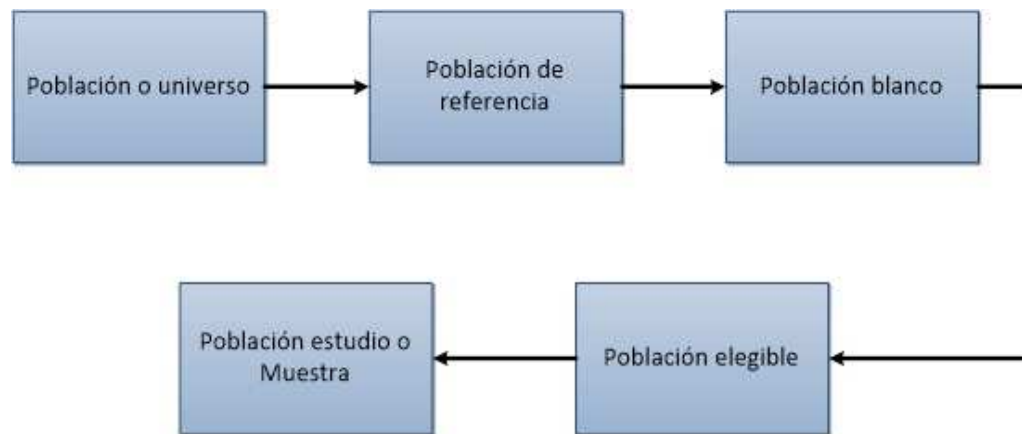


Figura 3.1: Población universo a la muestra.

Para el estudio de dislexia, la población de referencia se encuentra formada por estudiantes en la región de Bizkaia, además de niños y niñas de los departamentos de neurología y oftalmología pediátrica provenientes del hospital de Cruces, el cual cubre a una población de 385.000 habitantes.

Mientras que para el estudio de migraña, la población de referencia se encuentra formada por pacientes de la localidad de Bizkaia, atendidos específicamente en el hospital de Galdakao, el cual cubre a una población de unos 309.000 habitantes.

3.2 Criterios de selección

3.2.1 Caso de estudio 1: Dislexia

Para todos los grupos, se han seleccionado sujetos de las siguientes procedencias:

3. BASE DE DATOS

- Niños y niñas pacientes de los departamentos de neurología y oftalmología pediátrica del Hospital de Cruces.
- Para el grupo de control también se reclutaron estudiantes locales.

Los **criterios de inclusión** para todos los grupos y específicos de cada grupo son:

- Edades entre 9 y 12 años.
- Niños y niñas.
- Firma del Consentimiento informado, supervisado en primera instancia por el Comité de Ética del Hospital de Galdakao, y de acuerdo con el Código de Ética de la Asociación Médica Mundial (Declaración de Helsinki).
- Dislexia: no habían recibido tratamiento o apoyo psicopedagógico para la alfabetización y cuando mostraron una puntuación en Wechsler Intelligence Scale – Fourth Edition dentro del rango normal, > 75 .
- Visión monocular: tenían una capacidad de lectura normal.
- Control: no mostraron signos de que su alfabetización se viera afectada por patologías oculares.

Los **criterios de exclusión** para todos los grupos:

- Cualquier circunstancia contraria a los criterios de inclusión.
- Intolerancia a la exploración por resonancia magnética de cooperación, etc.
- Enfermedad previa neurológica o psiquiátrica.
- Traumatismo craneal grave.
- Alteración de la coordinación sensorio-motora.
- Tratamiento crónico con medicamentos.
- Privación social, escolarización inadecuada.

3.2.2 Caso de estudio 2: Migraña

Para los criterios de selección para el estudio de migraña que se han seguido, se ha tomado como referencia la clasificación de migrañas de la International Headache Society (IHS) 2013 [OBK⁺13] :

- Pacientes con migraña crónica y abuso de medicación: cefalea durante más de 15 días al mes con más de 3 meses de evolución, con uso de ergóticos, triptantes, opioides, y combinación de analgésicos durante más de 10 días al mes o uso de analgésicos simples durante más de 15 días al mes.
- Pacientes con migraña esporádica: 2 a 8 episodios de migraña al mes durante más de 3 meses y sin abuso de medicación.

Para el grupo control, se han seleccionado diversos voluntarios sanos sin migrañas de algún tipo y que aceptasen participar en el estudio. Su procedencia es de:

- Personal trabajador del hospital de Galdakao.
- Pacientes de las Consultas Externas de Neurología en los que no se observase lesión cerebral por Resonancia Magnética, ni ninguna patología del sistema nervioso central.

Los **criterios de inclusión** para todos los grupos son:

- Edades entre 18 y 50 años.
- Hombres y mujeres.
- Nivel de escolarización mínima de 10 años.
- Firma del Consentimiento informado, supervisado en primera instancia por el Comité de Ética del Hospital de Galdakao.
- No presentar deterioro intelectual, enfermedad psiquiátrica previa, ni seguimiento de tratamientos farmacológicos de tipo psiquiátrico.

3. BASE DE DATOS

Tabla 3.1: Sujetos participantes en el estudio (Dislexia)

	Control	Dislexia	Visión Monocular
Total	19 (33.3)	20 (35.1)	18 (31.6)
Datos socio-demográficos			
Edad	10.0 (0.9)	10.5 (1.1)	10.4 (0.9)
Sexo (Mujer/Hombre)	8/11	8/12	8/10
Datos clínicos			
Desorden de hiperactividad	2	2	1
Correcta agudeza visual	1.00	1.00	1.00
IQ Test			
Escala	107.9 (12.2)	95.9 (16.6)	103.5 (7.7)
Índice comprensión verbal	109.3 (14.1)	94.1 (18.0)	101.3 (10.4)
Índice de razonamiento perceptivo	106.4 (9.7)	103.4 (21.5)	101.3 (20.8)
Índice de velocidad de procesamiento	97.1 (9.9)	96.8 (8.2)	102.1 (10.6)
Índice de memoria de trabajo	100.9 (19.7)	97.8 (16.6)	107.8 (9.9)
BASC (rango)	40-60	40-60	40-60
Puntuación lectura (PROLEC-R)			
Precisión lectura palabras (n/40)	39.6 (0.7)	36.2 (3.4)	39.8 (0.8)
Precisión lectura pseudopalabras (n/40)	37.5 (1.5)	30.6 (5.4)	36.4 (1.8)
Velocidad en lectura de palabras (s)	41.3 (7.8)	85.8 (49.9)	32.3 (12.0)
Velocidad en lectura de pseudopalabras (s)	68.5 (15.1)	104.3 (54.8)	59.0 (18.6)
Habilidad lectura palabras (precisión/velocidad) 100	99.8 (20.5)	58.9 (30.9)	142.8 (44.9)
Habilidad lectura pseudopalabras (precisión/velocidad) 100	57.6 (12.2)	37.3 (17.7)	69.9 (24.5)

Los **criterios de exclusión** para todos los grupos son:

- Cualquier circunstancia contraria a los criterios de inclusión.
- Dominancia zurda.
- Abuso a otro tipo de sustancias.
- Presentar una lesión focal al realizar la prueba de resonancia magnética, o cualquier circunstancia que impida la realización de la misma, como pudiera ser marcapasos, claustrofobia, etc.

3.3 Tamaño muestral

3.3.1 Caso de estudio 1: Dislexia

La muestra utilizada para el estudio de dislexia es de 57 sujetos (19 controles, 20 sujetos con dislexia y 18 con visión monocular).

La tabla 3.1 muestra los sujetos que han participado en esta parte de la investigación. Se muestran los datos socio-demográficos y clínicos con su media y

Tabla 3.2: Sujetos participantes en el estudio (Migraña)

	Migraña Crónica	Migraña Esporádica	Control
Total	18 (34.6)	19 (36.5)	15 (28.9)
Datos socio-demográficos			
Edad	43.8(7.1)	41 (7.9)	45.7 (6.8)
Sexo (Mujer)	14 (77.8)	19 (100)	14 (93.3)
Datos clínicos			
Reserva cognitiva			
≤ 11 (Baja)	9 (50)	3 (15.8)	1 (6.7)
12-17 (Baja-Media)	4 (22.2)	5 (26.3)	4 (26.7)
16-18 (Media-Alta)	2 (11.1)	6 (31.6)	7 (46.7)
> 18 (Alta)	3 (16.7)	5 (26.3)	3 (20)

desviación típica. Excepto en el dato de edad que se muestra el número neto y el desorden de hiperactividad, correcta agudeza visual y BASC, que se muestran los valores específicos.

3.3.2 Caso de estudio 2: Migraña

La muestra actual utilizada en el estudio de migraña es de 52 sujetos (15 controles, 19 pacientes con migraña esporádica y 18 con migraña crónica y abuso de medicamentos).

La tabla 3.2 muestra los sujetos que han participado en esta parte de la investigación. Se muestran datos socio-demográficos y clínicos con su media y desviación típica.

3.4 Método

El método seguido para la inclusión de los sujetos, diagnóstico y toma de imágenes de resonancia magnética se describe en los siguientes puntos.

3. BASE DE DATOS

3.4.1 Caso de estudio 1: Dislexia

3.4.1.1 Evaluación Neuropsicológica

Los niños y niñas con visión monocular secundaria a los de trastornos de motilidad ocular, fueron identificados durante un exámen realizado por un oftalmólogo pediátrico. Durante este mismo proceso de examinación, también los pacientes con dislexia fueron revisados por posibles problemas oftalmológicos.

Además las pruebas pudieron comprobar la agudeza visual tanto lejana como cercana (espontánea y corregida), y la visión binocular lejana y cercana (haciendo uso de las pruebas Worth –permite estudiar el estado de la visión binocular–, pruebas de vectografía –permitiendo conocer si el paciente con los 2 ojos abiertos anula alguno de ellos– y la prueba TNO –detección precoz de la ambliopía–).

Por otro lado, también se realizaron pruebas de cobertura y de no cobertura, permitiendo medir la visión monocular secundaria y trastornos de la motilidad ocular. Otras evaluaciones fueron llevadas a cabo con una lámpara de hendidura, examen del fondo de pupila dilatada, y graduación con ciclopejía.

También los sujetos fueron evaluados por un neuropsicólogo pediátrico, realizando tres tipos de pruebas:

1. Prueba WISC-IV [Wec05] , la cual permite medir la inteligencia. El umbral necesario (criterio de inclusión) en la puntuación del CI tenía que ser por encima de 75.
2. Pruebas de lectura, incluyendo test de lectura con calificación de baterías PROLEC-R (de 8 a 11 años) o PROLEC-SE (de 12 a 16 años). Estas pruebas, en función de la puntuación obtenida, permiten clasificar a los sujetos como lectores disléxicos.
3. Detección de posibles trastornos psicológicos, sensoriales o neurológicos profundos. También posibles problemas familiares o educativos. Para la evaluación del comportamiento se hizo uso de la prueba BASC (Behaviour Assessment System for Children, 2004).

También se realizaron pruebas adicionales, cuando la dominancia lateral no estaba clara. Por ello sometieron a los sujetos al test de Harris (2001), para determinar la dominancia lateral.

Toda esta información recogida, servirá para correlacionarla con las imágenes de MRI.

3.4.1.2 Exploración DTI

La resonancia magnética se adquiere en un sistema de 3 Teslas (Philips Achieva TX 3.0T), en los cuales los sujetos se introducen en un decúbito supino. El sistema tiene una bobina de cabeza de 32 canales (Philips Medical Systems, Best, the Netherlands).

Se adquirieron imágenes anatómicas en 3D estructurales T1 de alta resolución (SPGR-3D, TR/TE 7.4/3.4 ms, ángulo de giro 8° , tamaño de matriz 228 x 227, campo de visión 250 x 250, número de cortes 304, en resolución plana 1 x 1 x mm, NSA 1, Tiempo total de adquisición 4'58"). Estas exploraciones anatómicas se utilizaron para confirmar la ubicación de la medición DTI con respecto a las marcas cerebrales convencionales.

Después de la exploración anatómica, los sujetos se sometieron a una secuencia DTI, la cual consistía en la toma única de una imagen SENSE-EPI (Echo Planar Imaging) con factor de reducción 2, la cual tuvo una duración de 4'17"; 60 cortes, 2 mm sin espacios. Estas imágenes fueron recogidas para cubrir la totalidad del cerebro y el tronco cerebral de cada sujeto (TR / TE 6819/81 ms).

La matriz de adquisición fue de 112 x 112, reconstruida a 128 x 128 con un cambio de visión de 224 x 224 mm para una resolución en el plano de 1.75 x 1.75 x 2 mm .

La ponderación de difusión se aplicó a lo largo de los 15 ejes distribuidos en forma de esfera con un valor de $b = 800 \text{ s/mm}^2$, además de una imagen menos ponderada.

3. BASE DE DATOS

3.4.2 Caso de estudio 2: Migraña

3.4.2.1 Evaluación Neuropsicológica

Todo el proceso de evaluación fue realizada por 2 neurólogos para detectar, seleccionar y ver criterios de inclusión o exclusión de los participantes en los diferentes grupos.

Para la evaluación se realizaron diferentes tareas al grupo seleccionado de migraña: una prueba de depresión y ansiedad; calendario de migrañas para que anotaran los ataques de migraña que tienen día a día, la medida tomada y el número de días laborales perdidos debido a la migraña.

Además el equipo de neurología evaluó cada 3 meses todas estas pruebas, excepto la prueba de depresión y ansiedad. De esta manera pudiendo determinar si la migraña crónica y el abuso de medicamentos disminuían con un tratamiento adecuado y/o si los pacientes con migraña esporádica terminaban recurriendo a la ingesta excesiva de medicamentos.

Fueron monitorizados durante un año, después del cual se realizaría una última evaluación, tomando de nuevo datos de las pruebas de depresión y ansiedad.

Gracias esta información recopilada, permitió la clasificación de los pacientes con migraña según tuvieran un buen pronóstico (esporádicos que siguen con ese tipo de migraña, y abusadores que han pasado a no abusar de los medicamentos). O por el contrario tener un mal pronóstico (abusadores que persisten en la ingesta excesiva de los medicamentos, o los esporádicos que se han empezado a abusar de los medicamentos). Tanto en los casos de buen o mal pronóstico, se podría comprobar a la finalización del estudio, si existía alguna diferencia en las imágenes de MRI correlacionándolo con el resultado clínico.

Respecto a los test que se incluyeron en el estudio tanto para clasificar como observar la evolución trimestral de los pacientes fueron los siguientes:

1. Prueba de Acentuación de Palabras, exigiendo un umbral de puntuación de 30 palabras en castellano de baja frecuencia a las cuales se les había eliminado los símbolos de acentuación. Gracias a esta prueba, medían el coeficiente de inteligencia.

2. SF-36, el cual medía la calidad de vida y el estado de salud a través de diversos dominios.
3. MSQoL, el test específico para la medición de la calidad de vida de sujetos con migraña.
4. Test MIDAS, referente a la gravedad del dolor, los días de dolor y la medición del sentimiento emocional.
5. Test BDI, el cual medía la dependencia; y el test BAI que medía la ansiedad.
6. Test MDQ-H, el cual medía la dependencia en la medicación durante el dolor de cabeza de los sujetos.

3.4.2.2 Exploración DTI

La resonancia magnética se adquiere en un sistema de 3 Teslas (Philips Achieva TX 3.0T), en los cuales los sujetos se introducen en un decúbito supino. El sistema tiene una bobina de cabeza de 32 canales (Philips Medical Systems, Best, the Netherlands).

Los sujetos se sometieron a una adquisición anatómica, siendo imágenes T1 3D de alta resolución, utilizando un gradiente de secuencia recuperada (SPGR-3D, TR / TE 7.4 / 3.4 ms; ángulo de giro 8 ° ; tamaño de matriz 228 x 227; campo de visión 250 x 250; número de cortes 301; en el plano de resolución 1 x 1 x 1 mm; NSA 1; tiempo total de adquisición 4'58").

Estas exploraciones anatómicas se utilizaron para la confirmación de la ubicación de la medición DTI con respecto a los puntos de referencia del cerebro convencionales.

Después de la exploración anatómica, los sujetos se sometieron a una secuencia DTI, la cual consistía en la toma única de una imagen SENSE-EPI (Echo Planar Imaging) con factor de reducción 2, la cual tuvo una duración de 4'17"; 60 cortes, 2 mm sin espacios. Estas imágenes fueron recogidas para cubrir la totalidad del cerebro y el tronco cerebral de cada sujeto (TR / TE 6819/81 ms).

La matriz de adquisición fue de 112 x 112, reconstruida a 128 x 128 con un cambio de visión de 224 x 224 mm para una resolución en el plano de 1.75 x 1.75 x 2 mm .

3. BASE DE DATOS

La ponderación de difusión se aplicó a lo largo de los 15 ejes distribuidos en forma de esfera con un valor de $b = 800 \text{ s/mm}^2$, además de una imagen menos ponderada.

3.5 Sumario

En este capítulo se ha descrito y definido las 2 poblaciones de referencia que se van a utilizar en la metodología de esta tesis. Se han enumerado tanto los criterios de selección, tamaño muestral y el método seguido para la obtención de las imágenes de resonancia magnética.

Capítulo

4

Diseño

Este capítulo contiene la descripción detallada de los métodos utilizados en esta tesis, los cuales han sido desarrollados y aplicados con la finalidad de validar la hipótesis planteada.

A través del uso de dicha metodología se desea aplicar algoritmos de Machine Learning en imágenes de resonancia magnética, validando la confianza de su aplicación mediante diversos parámetros de los propios algoritmos, permitiendo una clasificación lo más precisa posible en las dos bases de datos utilizadas en esta tesis: estudio de dislexia y estudio de migraña. Se extraerán las características propias de las imágenes DTI y test específicos de cada estudio para la clasificación y distinción de cada grupo existente en cada estudio.

Añadir, que la metodología de estos bloques se aplicará a los dos estudios de esta tesis de la misma manera; esto es, al estudio de dislexia y al estudio de migraña.

El resto del capítulo se encuentra organizado de la siguiente manera. La sección 4.1 describe el diseño de alto nivel de los bloques que conforman esta tesis. La sección 4.2 afronta el diseño de bajo nivel para cada uno de los bloqueos, descomponiéndolos a su vez en diagramas más completos para su mejor comprensión. Y finalmente, la sección 4.3 describe el resumen de este capítulo.

4. DISEÑO

4.1 Diseño de alto nivel

El diseño de alto nivel de esta tesis se divide en 5 grandes bloques (Figura 4.1) que contienen la totalidad de la investigación. Estos bloques son: Pre-procesado de las imágenes DTI de resonancia magnética, Selección de la matriz de características, Interfaz de usuario y Algoritmos de Machine Learning.

A continuación, se describen cada uno de estos bloques que conforman el diagrama de alto nivel, indicando su funcionamiento, y la descripción detallada de las entradas y salidas de cada uno de estos bloques.

4.1.1 Pre-procesado imágenes DTI (A)

En este primer bloque se muestra el diseño del pre-procesado en imágenes DTI. Estas imágenes han sido adquiridas mediante una única toma y las cuales muestran los tractos del cerebro.

Para este procesado es necesaria una imagen de referencia del cerebro anatómica de alta resolución (imágenes estructurales T1). Gracias a estas exploraciones anatómicas se puede confirmar la ubicación de la medición DTI con respecto a las marcas cerebrales convencionales.

El objetivo en este bloque es obtener datos numéricos de cada área del cerebro relevante, obteniendo por cada sujeto del estudio un valor por cada zona cerebral. De esta manera se obtiene una matriz en la que se tienen todos los valores de cada área cerebral estudiada de cada uno de los participantes.

4.1.2 Selección de características (B)

Una vez obtenidos los datos de todos los sujetos y de cada una de las áreas cerebrales que intervienen en cada una de las dos patologías de esta tesis, se realiza un método para la selección o extracción de características relevantes.

En esta fase se tienen en cuenta tanto los datos obtenidos de las imágenes DTI, como de los test específicos de cada una de las dos patologías.

El fin de este bloque es obtener un subconjunto de características tanto del propio cerebro como de los test, las cuales sean relevantes y aporten una mayor in-

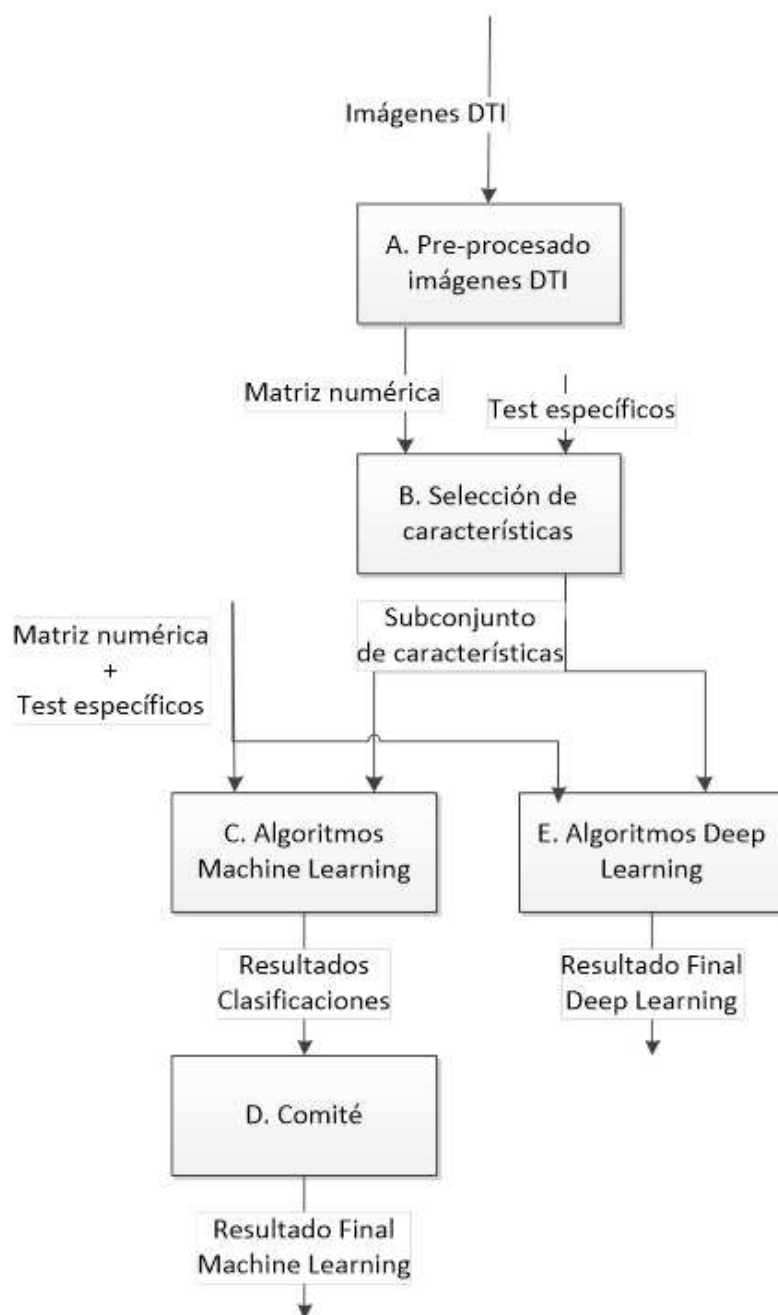


Figura 4.1: Diseño de alto nivel.

4. DISEÑO

formación. Es decir, características que ayudan a la diferenciación entre los grupos dentro de una misma patología.

Aunque se determine una selección de características, a la hora de aplicar los métodos de Machine y Deep Learning, se hará todo este proceso con los dos conjuntos: el que contiene toda la información (todas las zonas cerebrales y todos los test), y el obtenido en este bloque gracias a la selección de características.

4.1.3 Algoritmos Machine Learning (C)

Hasta ahora, concluidos los bloques anteriores, es necesario ordenar los resultados que se han obtenido en la salida de los bloques previos. Por un lado, se dispone del conjunto de características, es decir, la unión de los datos de las imágenes DTI y los resultados obtenidos de los diferentes bloques específicos.

Por otro lado, gracias al bloque previo, se obtiene una selección de características, las más relevantes y que aportan información necesaria para la correcta clasificación de los sujetos en diferentes grupos dentro de cada uno de los dos estudios.

Por tanto, con estos dos conjuntos de datos, el completo y la selección de los mismos, se necesita realizar un análisis el cual permita distinguir entre los 3 grupos catalogados en cada uno de los dos estudios: dislexia y migraña. Para ello, se hace uso de clasificadores automáticos dentro de la categoría de Machine Learning, permitiendo la diferenciación correcta de los sujetos y clasificarlos dentro de uno de los 3 grupos existentes en cada uno de los estudios de esta tesis.

4.1.4 Comité (D)

En este último bloque del diseño de alto nivel, se tiene en cuenta los resultados obtenidos del bloque de algoritmos de Machine Learning.

Este nuevo bloque consiste en la recogida de los resultados de clasificación para cada una de las dos patologías estudiadas en esta tesis, y aplicar a dichos resultados un método de comité o ensemble; teniendo en cuenta para el resultado final de clasificación el de cada uno de los algoritmos aplicados de Machine Learning. Esta metodología permite afianzar los resultados de los algoritmos, obteniendo unos resultados mejores y más seguros, ya que se tiene en cuenta la clasificación de cada

uno de ellos, y emitir el resultado final en función del resultado mayoritario de los diferentes algoritmos.

Finalmente, se obtiene un resultado final para cada uno de los sujetos que han participado en el estudio dentro de los 3 grupos diferenciados para cada una de las patologías de esta tesis.

4.1.5 Algoritmos Deep Learning (E)

De la misma manera que en el bloque de Algoritmos de Machine Learning, en este bloque se hace uso de algoritmos de clasificación automática, pero que se encuentran catalogados como Deep Learning.

Por ello, también se hace uso de los dos conjuntos de datos, el cual contiene la completitud de los datos de las imágenes DTI y los test específicos de cada una de las patologías estudiadas en esta tesis; y el conjunto de datos los cuales han sido obtenidos mediante una selección de los mismos, los cuales son los más relevantes dentro del conjunto completo de los datos. Al igual que en el bloque de Machine Learning, también se aplica a los dos estudios de esta tesis, dislexia y migraña.

La ventaja a la hora del uso de este tipo de algoritmos dentro de Deep Learning es la independencia a la hora de emitir una clasificación para cada sujeto, ya que este tipo de algoritmos son capaces de aprender de sus errores y modificar en cada iteración para optimizar al máximo la propia clasificación.

4.2 Diseño de bajo nivel

4.2.1 Diseño de bajo nivel

Las imágenes DTI son obtenidas a partir de las moléculas de agua que transcurren por los tractos del cerebro. En la figura 4.2 se puede observar un ejemplo de los tractos obtenidos gracias a una tractografía.

Gracias a un estudio de DTI probabilístico, se pueden obtener los tractos existentes entre dos puntos del cerebro. Este análisis tiene un coste computacional excesivamente alto, ya que por ejemplo, para un solo vóxel (unidad cúbica que compone un objeto tridimensional, siendo la unidad mínima procesable y equivalente a un

4. DISEÑO

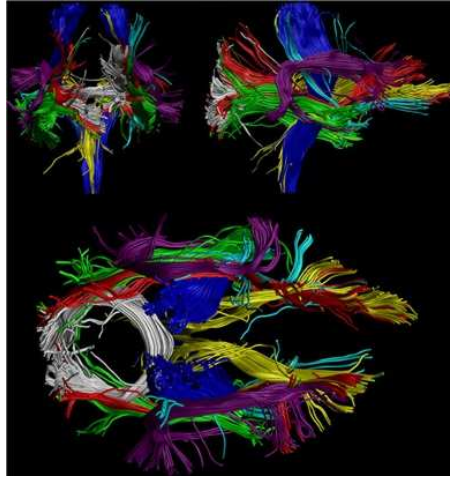


Figura 4.2: Tractos del cerebro.

Fuente ResearchGate por Timo Roine

píxel de un objeto en 2 dimensiones), el número de fibras que transcurren puede alcanzar el número de 10.000.

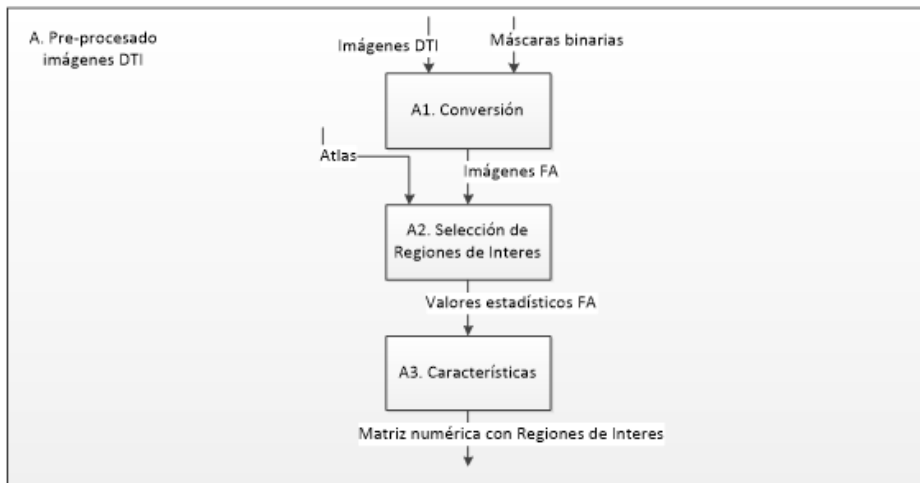


Figura 4.3: Bajo nivel (A).

4.2.1.1 Conversión (A1)

En este bloque en el cual se realiza la conversión de las imágenes originales obtenidas del escáner de resonancia magnética. Para esta conversión se realizó gracias a

la herramienta Oxford Centre for Functional MRI of the Brain (FMRIB) Software Library (FSL), versión 5.0.5 (<http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/>).

El proceso para dicha conversión es el siguiente:

- Se calculan las máscaras binarias del tamaño y la forma del cerebro de cada sujeto haciendo uso de la herramienta de FSL llamada BET (Brain Extraction Tool). Esta máscara que se calcula es importante a la hora de realizar futuras correcciones de la propia imagen.
- Se elaboran los mapas de anisotropía fraccional FA (Fractional Anisotropy) para cada sujeto mediante la función FDT (FMRIB's Diffusion Toolbox). En este punto se obtiene el valor del tensor de difusión para cada vóxel del cerebro. En este paso además se corrigen las distorsiones de los posibles estiramientos surgidos de la realización de las imágenes de referencia.
- Se aplica la función del paquete de FSL llamada TBSS (Tract-Based Spatial Statistics), alineando los datos de los valores de FA de todo el cerebro de cada sujeto. Además se realiza una alineación mediante un registro no lineal con el espacio cerebral representativo de cada sujeto, y finalmente alineando dichos datos al espacio estándar MNI152.
- Finalmente, todos los sujetos al registrarse en un mismo espacio común, se genera la imagen de FA, de la cual se crea el esqueleto promedio. Esta imagen representa los centros de todos los tractos comunes de todos los sujetos. Este proceso se le aplicó un umbral de $FA = 0.20$.

En este punto se tienen alineados todos los valores FA individuales de cada sujeto.

4.2.1.2 Selección de Regiones de Interés (A2)

Para este proceso de selección de las regiones de interés se han seleccionado 2 mapas cerebrales, también llamados Atlas, los cuales contienen las divisiones de todas las zonas cerebrales fisiológicas. Gracias a estos mapas cerebrales, se pueden seleccionar zonas específicas, las cuales intervienen en las dos patologías estudiadas en esta tesis.

4. DISEÑO

Por tanto, las Regiones de Interés analizadas en esta tesis se seleccionaron por su participación conocida en cada una de las dos patologías: Dislexia y Migraña.

Los 2 Atlas seleccionados son:

- JHU White-Matter Tractography Atlas, el cual es específico para los tractos de materia blanca.
- MNI Structural Atlas, en el cual se seleccionaron únicamente las zonas en los que se encontraban los tractos específicos de la materia blanca. Este Atlas también contiene zonas cerebrales que se corresponden a la materia gris del cerebro, la cual queda excluida de esta tesis, ya que requiere otro tipo de procesamiento de imagen no correspondiente al de DTI, como puede ser estudios de Resting.

Para cada uno de estos 2 Atlas, se procede a realizar un ajuste de las zonas de la materia blanca. Para ello, se segmenta el cerebro gracias al paquete de la herramienta FSL llamado FAST (FMRIB Automated Segmentation Tool), permitiendo que las Regiones de Interés se correspondan perfectamente con la propia materia blanca de la cual se quieren extraer los datos.

También ha sido necesaria la creación de nuevas Regiones de Interés, a partir de las originales obtenidas de los Atlas, pudiendo dividir en sub-zonas las de referencia aportadas por dichos Atlas. Para ello se ha hecho uso del paquete de FSL llamado FSLview. Este paquete permite visualizar el esqueleto cerebral creado previamente con las zonas comunes de todos los sujetos, y poder pintar manualmente las Regiones de Interés necesarias. Hay que tener en cuenta, que al estar en un espacio de 3 dimensiones (coronal, vista frontal; axial, vista desde arriba; sagital, vista lateral), es necesario pintar la Región de Interés en dichas 3 dimensiones.



Figura 4.4: Región de interés en las 3 dimensiones.

Siguiendo con el procedimiento, una vez acabado el dibujo manual de la Región de Interés, se procede al guardado de dicha zona, para utilizarlo posteriormente a la hora de la extracción de los valores numéricos de FA.

Esto se realiza gracias a la función de la herramienta FSL llamada FSLMATHS, la cual permite realizar operaciones matemáticas con los datos tales como operaciones binarias, algebraicas, matriciales, de dimensionalidad, estadísticas, etc.

Específicamente se hace uso de esta función para convertir las Regiones de Interés creadas manualmente, para guardarlas en el espacio estándar MNI, como se ha comentado previamente. De esta manera se asegura que se extraerán correctamente los datos de cada uno de los sujetos, ya que el espacio se encuentra estandarizado.

Finalmente aplicando estas Regiones de Interés calculadas (tanto las provenientes de los Atlas, como las creadas de forma manual), se extraen los valores medios de la imagen de FA calculada previamente, obteniendo para cada uno de los sujetos el valor medio de FA en cada una de las Regiones de Interés seleccionadas. Para ello se hace uso del paquete de FSL llamada Featquery.

Esta función Featquery, permite el cálculo de ciertos datos estadísticos de las zonas de interés, haciendo uso del espacio estándar MNI. El resultado de la llamada a esta función, aporta los datos estadísticos necesarios para la posterior investigación de los sujetos. Aporta una tabla con contenido estadístico en la cual se puede obtener el tipo de estadístico utilizado; el número de vóxels de dicha Región de Interés que no tienen valor 0; valores estadísticos como la media, mediana, desviación estándar, mínimo y máximo; y las coordenadas de la Región de Interés analizada.

Para esta tesis, se ha utilizado el valor estadístico Z, el cual describe la posición de una observación relativa a la media en unidades de la desviación estándar, sin aplicar ningún umbral para la obtención de los valores estadísticos.

Por tanto, se le aplica dicha función a cada uno de los sujetos de los 2 grupos estudiados en esta tesis, y se obtiene los datos estadísticos de FA para cada uno de ellos. Para los siguientes pasos de esta tesis, se cogen los valores medios de FA resultantes de esta función.

4. DISEÑO

Featquery output

stats image	# voxels	image statistics						position of max					
		min	10%	mean	median	90%	max	stddev	vox (FMRI space)	min (standard space)	max voxel atlas result		
stats_mstat1	710	0.07998	0.3171	0.63	0.8695	0.852	1	0.1081	81	73	23		

Figura 4.5: Salida de la función Featquery.

4.2.1.3 Características (A3)

En este nuevo bloque, se aúnan todas las características que van a ser utilizadas para la clasificación de los sujetos en los diferentes grupos de cada una de las dos patologías de esta tesis, Dislexia y Migraña.

Gracias a este bloque, se genera una matriz de NxM, en la que la N se corresponde al número de sujetos del estudio, y la M son las diferentes características que serán utilizadas por los clasificadores para la correcta clasificación de dichos sujetos.

Por tanto, este bloque se centra en la unión de las diferentes características, provenientes de los valores FA resultantes de las Regiones de Interés específicas; y de los diversos test específicos para cada una de las patologías.

A continuación, se enumeran las características finales para los sujetos del estudio de Dislexia, y las pertenecientes al estudio de Migraña.

Caso de estudio 1

Regiones de Interés para el estudio de **Dislexia** (nombradas en Latín):

- Arcuate Fasciculus
- Genu (Anterior Corpus Callosum)
- Splenium (Posterior Corpus Callosum)
- Fasciculus Occipitofrontalis Inferior
- Capsula Interna

Respuestas de los test y pruebas específicas de **Dislexia**:

- WISC-IV
- PROLEC-R

- BASC (Behaviour Assessment System for Children)

Caso de estudio 2

Regiones de Interés para el estudio de **Migraña** (nombradas en Latín):

- Corpus Callosum
- Body Corpus Callosum
- Fasciculus Occipitofrontalis Inferior
- Genu (Anterior Corpus Callosum)
- Insula
- Lobus Frontalis

Respuestas de los test y pruebas específicas de **Migraña**:

- SF-36
- MSQoL
- MIDAS
- BDI
- BAI
- MDQ-H

4.2.2 Selección de características (B)

En este segundo bloque del diseño, es el encargado de seleccionar las características más relevantes dentro de todo el conjunto inicial. Gracias a este sistema, se pueden desechar variables innecesarias o repetitivas, permitiendo entregar al modelo de clasificación, lo realmente significativo para que la eficacia sea la mayor posible.

Este bloque es el que aúna diferentes técnicas de selección y reducción de características, de las cuales se obtendrá como resultado las características más relevantes del conjunto inicial de partida.

4. DISEÑO

El uso de estos algoritmos ha sido posible gracias al empleo de la librería del lenguaje Python “Scikit-Learn”, procediendo a modelar dichas técnicas para que la eficiencia sea la mejor posible.

Otra de las características de este conjunto de técnicas es la relacionada sobre el tipo de técnica de la que hacen uso, la supervisada. Esto quiere decir, que a los algoritmos se les debe enseñar un conjunto de muestras previamente clasificadas, para poder ajustar el modelo en función de lo que ha visto. Posteriormente, se les introduce un conjunto de muestras nunca vistas antes, en el que se puede medir la eficiencia del mismo.

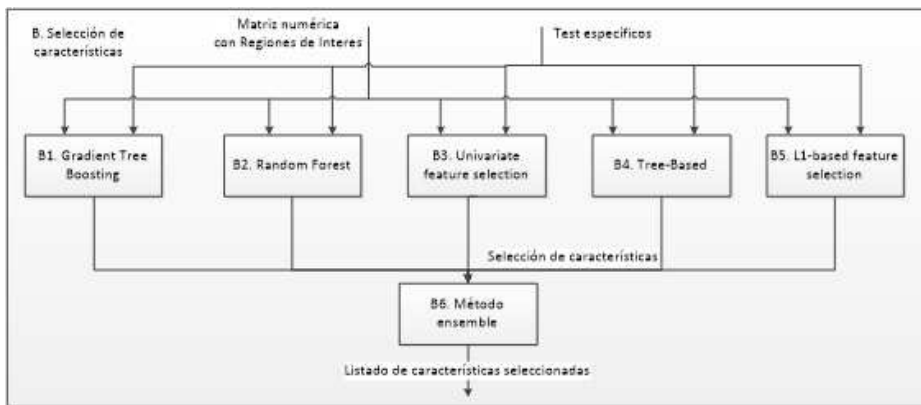


Figura 4.6: Bajo nivel (B).

4.2.2.1 Gradient Tree Boosting (B1)

El primer algoritmo utilizado para seleccionar características es el llamado Gradient Tree Boosting, el cual realiza una clasificación de las muestras que se le introducen para observar el error cometido, y poder seleccionar las características que hacen que el error sea el mínimo posible.

Por tanto, hay que parametrizar el algoritmo de tal manera que dé como resultado el conjunto de características que el propio algoritmo ha considerado como más importantes.

Para ello, se realiza la llamada al algoritmo y el atributo específico que contiene una función de decisión.

Además de hacer la llamada a la función de decisión, también se parametriza el propio algoritmo de Gradient Tree Boosting

Los parámetros son los siguientes:

- Número de estimadores: 100.
- Ratio de aprendizaje: 1.
- Máxima profundidad: 1.
- Estado: 0.

Este algoritmo se ha implementado tanto en el estudio de Dislexia como de Migraña.

4.2.2.2 **Random Forest (B2)**

El siguiente algoritmo utilizado, es Random Forest, el cual hace uso de árboles de decisión para poder seleccionar las características dentro del todo el conjunto inicial.

De nuevo, como este algoritmo se utiliza también para la predicción, se hace uso del atributo de la función de decisión, para que en vez de dar como resultado la clasificación, devuelva las características que ha considerado más relevantes para una óptima clasificación de las muestras.

Los parámetros que se han ajustado para conseguir las características más importantes son:

- Número de estimadores: 100.
- Máxima profundidad: no asignada, por lo que los nodos son expandidos hasta que las hojas son puras.
- Estado: 0.
- Mínimo de muestras: 2.
- Mínimo de muestras en la hoja: 1.

Este algoritmo se ha implementado tanto en el estudio de Dislexia como de Migraña.

4. DISEÑO

4.2.2.3 Univariate feature selection (B3)

El tercer algoritmo es el llamado Univariate feature selection, este algoritmo es específico para este fin, ya que los anteriores también son utilizados como clasificadores. Este algoritmo, analiza cada una de las características por separado, y obtiene un ranking de importancia de las propias características.

La técnica del análisis univariado utilizada ha sido:

- Selección de las K mejores, selecciona las características en función de las puntuaciones de las propias características.

En la parametrización de este algoritmo se realiza mediante los siguientes parámetros:

- K: 10.
- Función de puntuación: Chi cuadrado.

Este algoritmo se ha implementado tanto en el estudio de Dislexia como de Migraña.

4.2.2.4 Tree-Based (B4)

El algoritmo Tree-Based, es otro de los que hacen uso de árboles de decisión, por lo que se puede utilizar tanto para seleccionar características como para la clasificación de muestras.

Los parámetros que se han ajustado para conseguir las características más importantes son:

- Número de estimadores: 100.
- Máxima profundidad: no asignada, por lo que los nodos son expandidos hasta que las hojas son puras.
- Estado: 0.
- Mínimo de muestras: 2.
- Mínimo de muestras en la hoja: 1.

Este algoritmo únicamente se ha implementado en el estudio de Dislexia.

4.2.2.5 L1-based feature selection (B5)

El último algoritmo, L1-based feature selection, es otro de los específicos para este fin de poder seleccionar las mejores características de todo el conjunto inicial.

La parametrización que se ha llevado a cabo para este caso es la siguiente:

- C: 0.01.
- Penalización: L1.
- Dual: Falso.

Este algoritmo únicamente se ha implementado en el estudio de Migraña.

4.2.2.6 Método ensemble (B6)

Una vez se han ejecutado los algoritmos de selección de características, se obtiene como salida por cada uno de ellos el listado de las características más relevantes e importantes. Hay que tener en cuenta que estos algoritmos pueden tener ciertos errores, y en función de cómo se hayan modelado se puede dar más peso a una característica y no a otra.

Por tanto, se ha considerado crear un método de Ensemble por el cual convergen todas las salidas de cada uno de los selectores de características.

El método utilizado es por medio de votos, es decir, en este caso al utilizar 4 métodos de selección de características, se seleccionarán finalmente las características las cuales como mínimo se encuentren en 2 de los 4.

Como resultado de este método, se genera una nueva matriz de dimensiones de N sujetos por M características seleccionadas.

4.2.3 Algoritmos Machine Learning (C)

En este bloque se aúnan los algoritmos de clasificación automática. Los múltiples algoritmos existentes son capaces de aprender conceptos por sí solos, tratando de abstraer reglas analizando los datos, y sacar patrones que les permita clasificar nuevas observaciones nunca vistas antes.

Dentro de esta tesis, se han utilizado algoritmos tanto de aprendizaje supervisado, como por refuerzo, debido a que ya se tenían las muestras previamente

4. DISEÑO

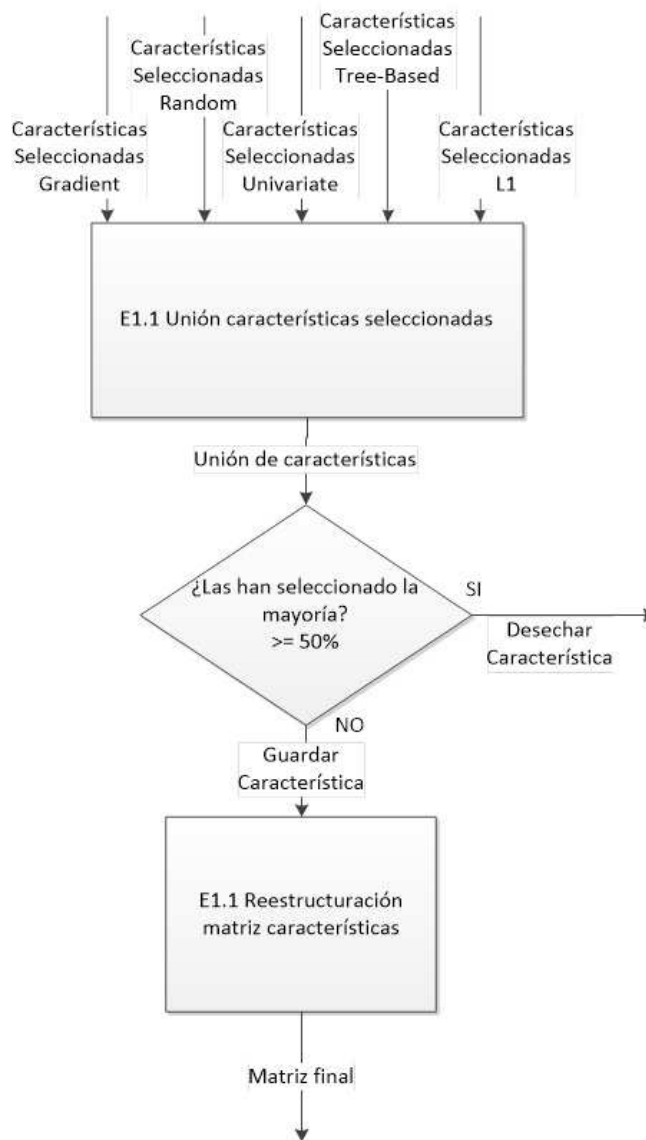


Figura 4.7: Método ensemble (B6).

clasificadas, y no se requería obtener relaciones del conjunto de datos. Sino, que se quiere conseguir un modelo capaz de clasificar correctamente a los sujetos de las dos patologías que se tratan en esta tesis: Dislexia y Migraña.

Otro punto a tener en cuenta, es que en todos los algoritmos de Machine Learning, se han implementado con validación cruzada o cross-validation, útil para evaluar los resultados de la clasificación, garantizando que son independientes entre los datos de entrenamiento y test.

Por tanto, en todos los algoritmos ejecutados, se les entrena con la validación cruzada, introduciéndoles el conjunto de entrenamiento y sus etiquetas asociadas a cada uno de los grupos. Una vez finaliza esta fase, se procede a la validación, en la que se le muestra el conjunto de test, que como condición es que no puede haber ninguna muestra que se haya incluido en el entrenamiento. El motivo de que no puedan coincidir las muestras en los conjuntos de entrenamiento y validación, es que no se puede influenciar al algoritmo, ya que evidentemente algo que ya ha visto en el entrenamiento y sabe su etiqueta, si lo ve en el conjunto de validación, lo va a clasificar correctamente.

4.2.3.1 SVM (C1)

El algoritmo de clasificación SVM, como ya se ha mencionado previamente, es de tipo supervisado. Para esta presente tesis, específicamente se ha hecho uso de la variante del algoritmo C-SVM.

La validación cruzada seleccionada para garantizar que los resultados de clasificación son correctos, se ha hecho uso del algoritmo Stratified ShuffleSplit, el cual devuelve conjuntos de entrenamiento y test aleatorios, conservando el porcentaje de la muestra para cada clase de los grupos a clasificar.



Figura 4.8: Algoritmo SVM (C1).

4. DISEÑO

La parametrización de la validación cruzada es la siguiente:

- Número de división: 10.
- Tamaño de test: 35
- Estado: 0.

Por tanto, una vez generada la validación cruzada, se ejecutará el algoritmo SVM las veces que se indique en la validación cruzada. Obteniendo diferentes resultados de las diferentes combinaciones respecto al conjunto de test frente al conjunto con el que se ha entrenado el modelo.

La parametrización para el algoritmo de aprendizaje supervisado SVM es la siguiente:

- Kernel: lineal, Support Vector Classifier.
- C: 1.

El resultado que proporciona este algoritmo, una vez se ha modelado durante el entrenamiento, y se procede posteriormente a enfrentar el conjunto de test, son los resultados de las clasificaciones de cada una de las muestras introducidas en la fase de validación. Estos resultados son numéricos, desde 0 hasta el número de grupos que haya para clasificar, y coincide con el etiquetado previo que se ha introducido en la fase de entrenamiento.

4.2.3.2 Adaboost (C2)

Otro de los algoritmos utilizados para esta tesis, es el llamado Adaboost, el cual pertenece al tipo de supervisados, pero se caracteriza por tener el aprendizaje por refuerzo, que mezcla los aprendizajes supervisados y no supervisados.

De nuevo, se ha realizado una validación cruzada, la cual permite asegurar que el conjunto de entrenamiento y de test son independientes, y generando unos resultados totalmente confiables.

La parametrización para el algoritmo Adaboost es la siguiente:

- Número de estimadores: 100.



Figura 4.9: Algoritmo Adaboost (C2).

- Estimador base: SAMME, el cual se adapta en base a los errores de las predicciones.
- Algoritmo: DecisionTreeClassifier (máxima profundidad: 1).

En esta ocasión, el estimador base que se ha utilizado es de tipo árboles de decisión. Mientras que el algoritmo escogido es SAMME, el cual se adapta en base a los errores de las predicciones.

Una vez realizada la fase de entrenamiento, posteriormente en la fase de validación, el algoritmo da como resultado las muestras etiquetadas del conjunto de test.

4.2.3.3 Naïve Bayes (C3)

El último algoritmo utilizado es el llamado Naïve Bayes, también perteneciente a los algoritmos supervisados.



Figura 4.10: Algoritmo Naïve Bayes (C3).

Para esta tesis, se ha utilizado la función Gaussiana.

4. DISEÑO

De nuevo se ha añadido la validación cruzada para poder determinar que los resultados de la clasificación son totalmente fiables.

Por último, los resultados que se obtienen una vez finalizada la fase de validación, son las muestras del conjunto de test clasificadas.

4.2.4 Comité (D)

En este último bloque, se describe el método por el cual, una vez teniendo los resultados preliminares de cada uno de los algoritmos entrenados e introduciéndoles posteriormente el conjunto de test, se procede a una técnica para obtener un único resultado de clasificación de todos los clasificadores.

Esta técnica también es llamada Comité, y el motivo de implementación es que cada uno de los algoritmos está parametrizado de una manera, siendo uno mejor que otro en clasificar cierto grupo. Lo que se quiere conseguir, es una unión de estos resultados, generando una única clasificación por sujeto.

4.2.4.1 Método de Ensemble algoritmos de Machine Learning(D1)

El proceso para la implementación del comité, también llamado método ensemble, es la de coger cada una de las clasificaciones del conjunto de validación de cada uno de los algoritmos ejecutados, y crear una nueva salida única.

Por lo tanto, el comité es formado para no tener una única opinión de un clasificador, porque los demás resultados también existen, y dependiendo de cómo se encuentren parametrizados los algoritmos, dependerá mucho del resultado final.

El método, por lo tanto, es seleccionar el resultado más repetido.

Este ensemble únicamente se ha aplicado a los algoritmos de Machine Learning, ya que su naturaleza es lineal, no abstrayéndose de las características que observan y generando unas nuevas, únicamente generan patrones de lo que ven durante la fase de entrenamiento. En cambio con las técnicas de Deep Learning, este paso no es necesario; los algoritmos tienen intrínsecamente selección de características por sí mismos, debido a su aprendizaje profundo.

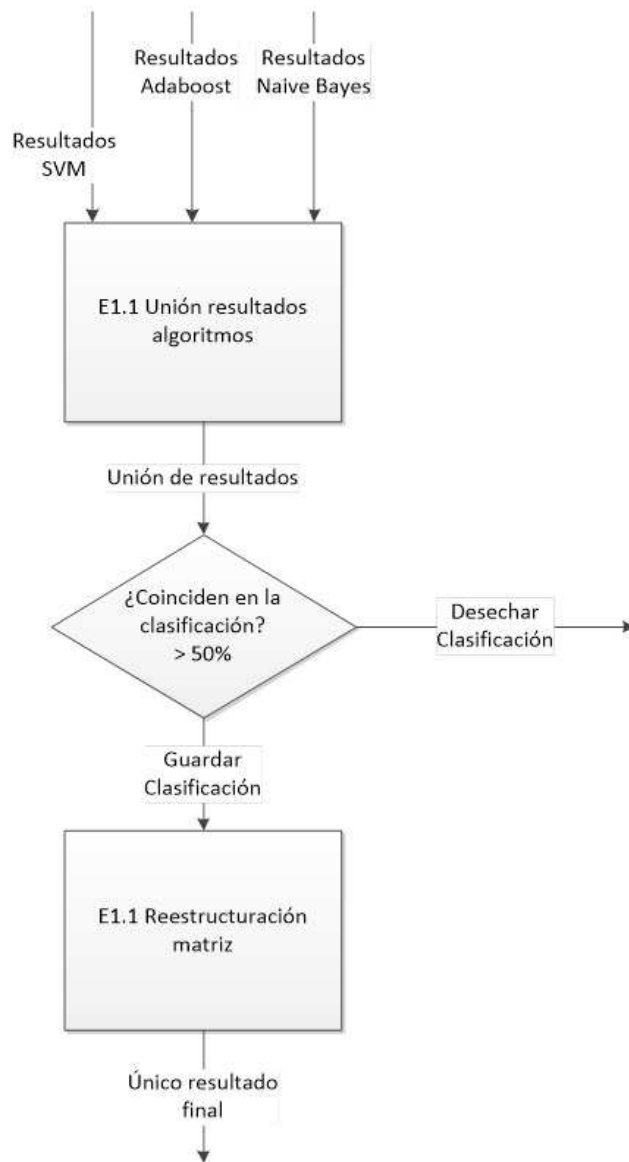


Figura 4.11: Método ensemble de Machine Learning (D1).

4. DISEÑO

4.2.5 Algoritmos Deep Learning (E)

En este nuevo bloque, se explica el algoritmo implementado perteneciente a la técnica de Deep Learning. Al igual que en el bloque de los algoritmos de Machine Learning, también son algoritmos de clasificación automática, los cuales pueden aprender y extraer patrones por sí solos, permitiendo que puedan clasificar muestras nunca vistas con unos ratios de clasificación muy buenos.

En esta tesis se ha implementado una red neuronal de tipo convolucional, es decir, Convolutional Neural Network.

4.2.5.1 Convolutional Neural Network (E1)

El algoritmo de Deep Learning utilizados en esta tesis, es el llamado Convolutional Neural Network, basado en la arquitectura de Redes Neuronales.

Estos algoritmos están preparados para la detección y clasificación de imágenes, por lo que hay que procesar los datos de tal manera que puedan introducirse en el algoritmo, pese a no ser imágenes.

Para ello se debe redimensionar el conjunto de datos de la forma que espera el modelo recibir dichos datos. Para ello se indica el número de muestras que contiene el conjunto de entrenamiento y la dimensión de la muestra (1 muestra por N características).

Posteriormente es necesario codificar la variable de la etiqueta que describe a qué clase pertenece cada muestra. Por este motivo, se crea una columna categorizando las muestras.

Una vez los datos están en el formato debido, se construye el modelo. El modelo que se ha utilizado es de tipo secuencial, el cual permite crear dicho modelo capa por capa. Por lo que se tiene que ir añadiendo y especificando cada capa que se quiera incluir.

Las 2 primeras capas son de tipo convolucional 2D, las cuales tratan los datos de entrada, y son las encargadas de aprender patrones locales. En este paso se selecciona el número de neuronas internas que tendrá cada una de las dos capas, en este caso 33 y 64 respectivamente. También se indica el tamaño del kernel o núcleo, específicamente 1x1 en las dos capas, siendo el tamaño del filtro para la convolución. Y por último, se incluye la función de activación de la capa, en este

caso se ha determinado la función “Relu” también llamada activación lineal rectificada, también en las dos capas, demostrándose que es de las que mejor funciona en las redes neuronales.

La tercera capa añadida es de tipo Max Pooling 2D, siendo un proceso de discretización y se consigue reducir la dimensionalidad del conjunto de entrenamiento dando una representación abstracta de los datos. Gracias a esta capa se reduce el costo computacional ya que se reducen la cantidad de parámetros.

La cuarta capa es la llamada Dropout, en la cual se seleccionan un número de neuronas al azar y se ignoran en las siguientes capas de entrenamiento, por lo que la contribución de estas neuronas en pasos posteriores se elimina temporalmente en la siguiente capa. Con esta capa se quiere evitar que el algoritmo se especialice en exceso en las características mostradas durante el entrenamiento, pudiendo conseguir peores resultados en conjuntos de validación no vistos. La parametrización ha sido de 0.25.

La quinta capa es de tipo “Flatten”, que únicamente sirve como conexión entre las capas de convolución y las capas densas.

La sexta capa es de tipo “Dense”, o capa densa, es la encargada de aprender patrones globales del conjunto de datos que recibe. En esta capa también se ha indicado la función “Relu” y 128 neuronas internas.

En la séptima capa se ha añadido otra capa de Dropout, para volver a desvincular ciertas neuronas y no especializar el modelo en exceso, en este caso parametrizada a 0.5.

Y por último, la octava capa de nuevo es una capa “Dense”, pero en esta ocasión se ha seleccionado una función “Softmax” y 3 neuronas internas, que permite que la salida de la clasificación sea en función de probabilidades, es decir, arrojará una probabilidad de la muestra introducida de cada uno de los grupos que conste la clasificación.

Una vez se han creado las capas, se procede a compilar el modelo mediante tres parámetros:

- Optimizador: es el que controla la tasa de aprendizaje. Parametrizada con la función Adadelta.

4. DISEÑO

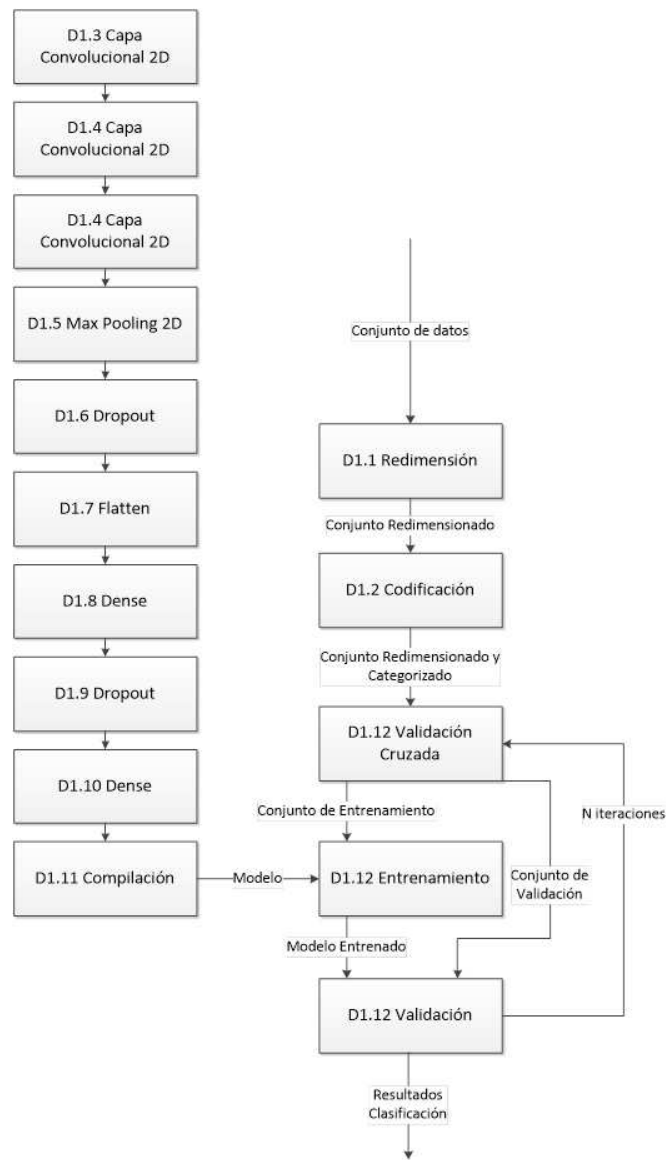


Figura 4.12: Diseño de la red Convolutional Neural Network (E1).

- Función de pérdida: mostrando los datos de error durante la fase de entrenamiento. Parametrizada a Categorical Crossentropy.
- Métrica: es el valor con el cual se quiere ver la precisión tanto del conjunto de entrenamiento como del de validación. Parametrizada a Accuracy.

Por último, se procede a entrenar el modelo, el cual también se ha aplicado, como en el caso de los algoritmos de Machine Learning, una validación cruzada.

Para el entrenamiento se ha seguido la siguiente parametrización:

- Épocas: 200.
- Tamaño batch: 1.
- Recuperación del mejor modelo: Verdadero.

Y finalmente se procede a la validación.

El resultado que arroja el algoritmo es la probabilidad de la muestra introducida de que pertenezca a cada uno de los grupos que compone el estudio.

4.3 Sumario

En este capítulo se ha detallado el diseño empleado para la realización de esta tesis.

Para su mejor comprensión, las fases llevadas a cabo durante esta tesis se han dividido en 5 partes, las cuales se corresponden cada una de ellas a los 5 grandes bloques señalados previamente.

Estos bloques, as su vez, se han dividido en sub-bloques los cuales permiten mostrar con exactitud los pasos a seguir para la realización de todo el proceso de la tesis. Además también se han señalado parámetros y técnicas necesarias para la ejecución de los métodos.

Resultados

En este siguiente capítulo se detallan los resultados obtenidos a partir del diseño explicado previamente. Se encuentra organizado de la siguiente manera. La sección 5.1 se describe el experimento que se ha llevado a cabo. La sección 5.2 enumera el conjunto de resultados para cada uno de los bloques que compone el diseño de esta tesis: Pre-procesado de imágenes DTI, Selección de características, Algoritmos de Machine Learning, Comité y Algoritmos de Deep Learning. La sección 5.3, realiza una comparativa en cuanto a diversos factores de Machine Learning y Deep Learning. Por último, la sección 5.4 se proporciona el sumario de este capítulo.

5.1 Definición del experimento

- Caso de estudio 1: Dislexia La muestra actual para este caso de estudio es de 57 sujetos (19 controles, 20 sujetos con dislexia y 18 con visión monocular).
- Caso de estudio 2: Migraña La muestra actual para este caso de estudio es de 52 sujetos (15 controles, 19 sujetos con migraña esporádica y 18 con migraña crónica).

La especificación de las muestras se encuentra recogida en el capítulo de Base de datos.

5. RESULTADOS

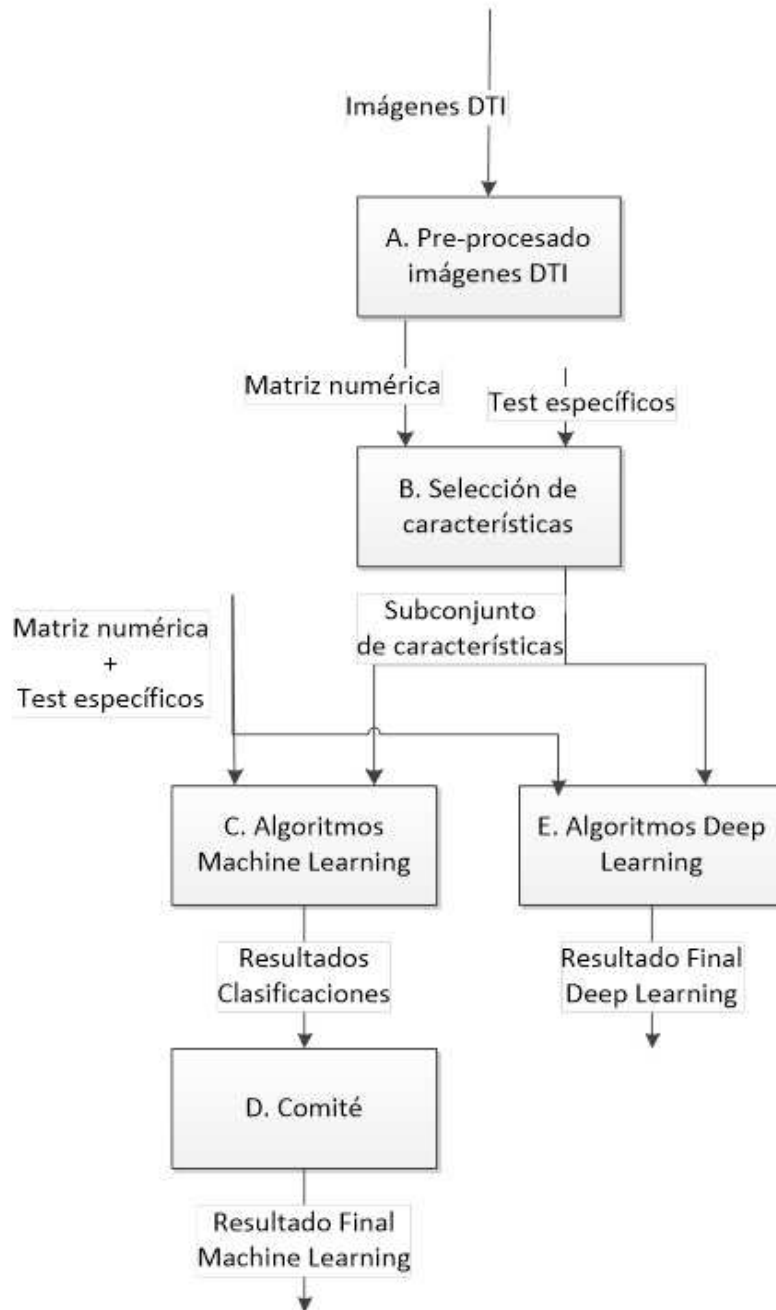


Figura 5.1: Diseño de alto nivel.

5. RESULTADOS

Para la obtención de los datos de las zonas cerebrales se seleccionaron Regiones de interés. Estas regiones son específicas para cada caso de estudio, siendo las siguientes:

Caso de estudio 1: Dislexia

- Arcuate Fasciculus

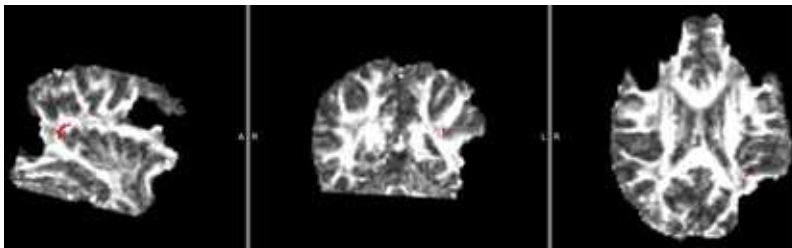


Figura 5.3: Arcuate Fasciculus.

La entrada para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del fascículo arcuato de la materia blanca.

- Genu (Anterior Corpus Callosum)

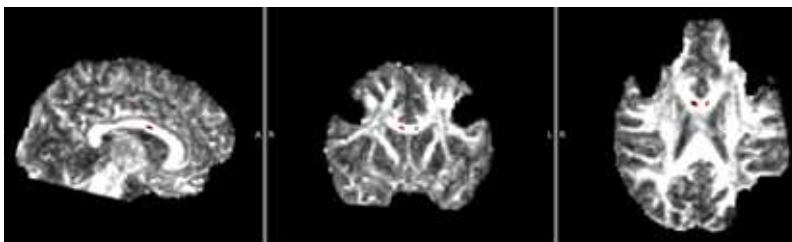


Figura 5.4: Genu (Anterior Corpus Callosum).

La entrada para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del genu de la materia blanca.

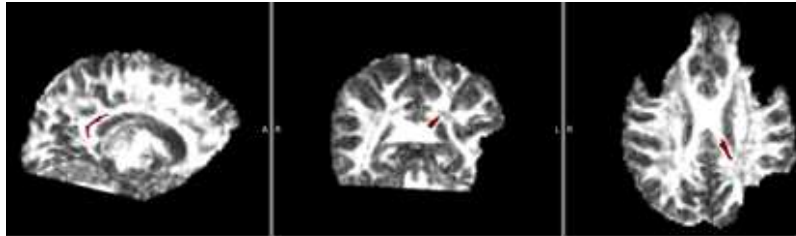


Figura 5.5: Splenium (Posterior Corpus Callosum).

- Splenium (Posterior Corpus Callosum)

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del splenium de la materia blanca.

- Fasciculus Occipitofrontalis Inferior

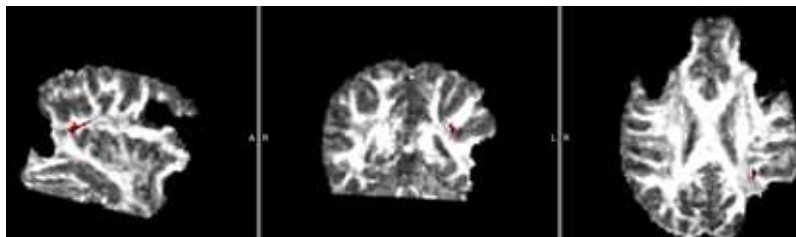


Figura 5.6: Fasciculus Occipitofrontalis Inferior.

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del fascículo occipitofrontal inferior de la materia blanca.

- Cápsula Interna

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona de la cápsula interna de la materia blanca.

Caso de estudio 2: Migraña

5. RESULTADOS

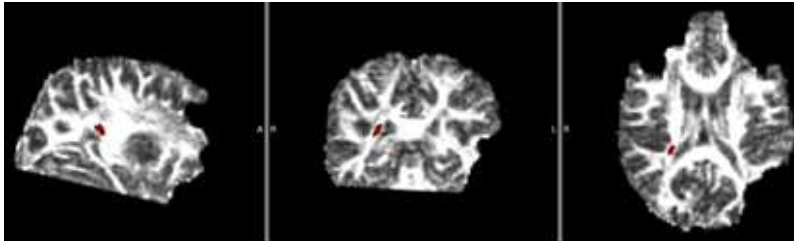


Figura 5.7: Cápsula Interna.



Figura 5.8: Corpus Callosum.

- Corpus Callosum

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del cuerpo calloso de la materia blanca.

- Body Corpus Callosum

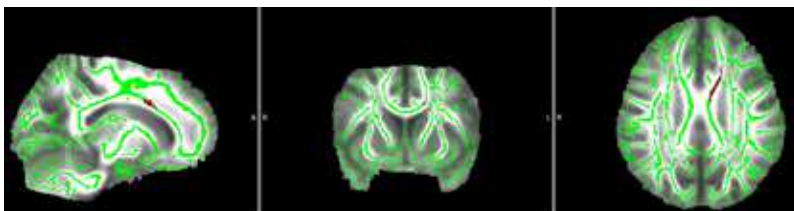


Figura 5.9: Body Corpus Callosum.

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del cuerpo calloso de la materia blanca.

- Fasciculus Occipitofrontalis Inferior

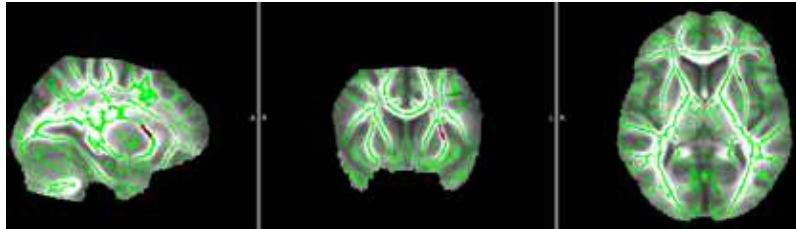


Figura 5.10: Fasciculus Occipitofrontalis Inferior.

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del fascículo occipitofrontal inferior de la materia blanca.

- Genu (Anterior Corpus Callosum)

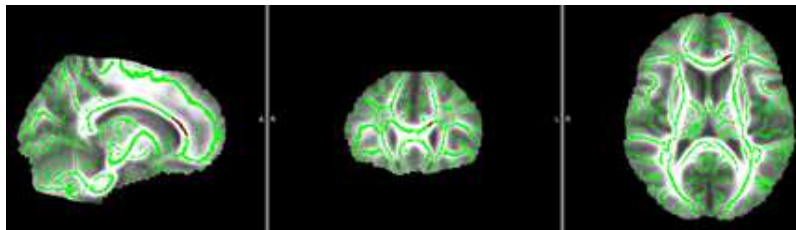


Figura 5.11: Genu (Anterior Corpus Callosum).

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del genu de la materia blanca.

- Insula

La entrada para para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona de la ínsula de la materia blanca.

5. RESULTADOS

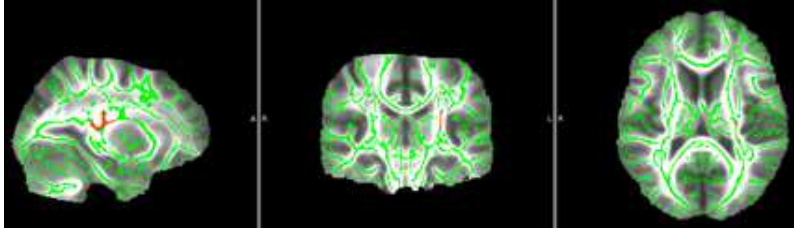


Figura 5.12: Insula.

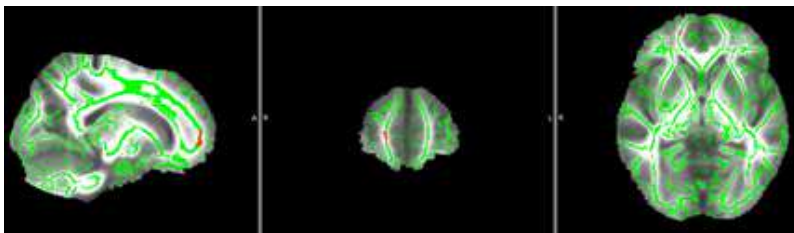


Figura 5.13: Lobus Frontalis.

- Lobus Frontalis

La entrada para aplicar la zona de interés es la imagen DTI completa del cerebro, siendo la salida el dato numérico del valor FA de la zona del lóbulo frontal de la materia blanca.

5.2.2 Selección de características (B)

La selección de características es muy importante en el ciclo de la investigación, ya que gracias a si la propia selección es correcta por parte de los algoritmos utilizados, proporcionarán unos mayores ratios de aciertos en la clasificación posterior.

Los resultados previamente a la selección de características y tras la aplicación de los algoritmos son los siguientes para cada uno de los casos de estudio de esta tesis:

Caso de estudio 1: Dislexia

Se parte de un total de 50 características entre las que se encuentran diferentes zonas del cerebro y datos de test específicos.

5.2 Evaluación de los resultados

Tabla 5.1: Número y porcentaje de reducción de características (Dislexia)

Algoritmo	Número de características	Porcentaje de reducción
Gradient Tree-Boosting	9	82 %
Random Forest	9	82 %
Tree Based	12	76 %
Univariate	10	80 %

Por tanto, la entrada tanto para la selección de características como para los algoritmos de clasificación automática es la correspondiente con los valores numéricos de FA de las imágenes DTI correspondientes a las regiones de interés; y los valores numéricos provenientes de los test específicos de los que se obtiene la puntuación por cada una de las preguntas realizadas.

10 de las 50 características se corresponden con zonas cerebrales relativas a:

- Arcuate Fasciculus
- Genu (Anterior Corpus Callosum)
- Splenium (Posterior Corpus Callosum)
- Fasciculus Occipitofrontalis Inferior
- Capsula Interna

Y las 40 características restantes se corresponden a preguntas de los siguientes test específicos:

- WISC-IV
- PROLEC-R
- BASC (Behaviour Assessment System for Children)

Se aplica la selección de características con 4 algoritmos: Gradient Tree-Boosting, Random Forest, Tree Based y Univariate, los cuales llegan a una **reducción de características del 82 %** en el caso de Gradient Tree-Boosting y Random Forest.

5. RESULTADOS

Tabla 5.2: Características seleccionadas (Dislexia)

Característica	Frecuencia (Todas = 4)
PROIPprecpsseudPD	4
PROIPrecpalPD	4
PROIVpalPDL	3
PROIVpsseudPD	2
spleniumTempR	2
postFIFOL	2
LIFR	2

De estos resultados, se deduce que las características que no ha seleccionado, o bien son repetitivas en cuanto al valor que aportan, o bien porque no son diferenciadoras en base a los grupos para clasificar.

Una vez obtenidas las características por cada uno de los algoritmos, se aplica el método de Comité para obtener una selección final de características, ya que no todos los algoritmos tienen que seleccionar las mismas características. Por lo que se ha tenido en cuenta para que una característica se tenga en cuenta, que por lo menos 2 de los algoritmos la hayan seleccionado.

De esta selección, se obtienen un total de **7 características finales** (3 de ellas de zonas cerebrales), las cuales son significativas a la hora de la posterior clasificación de los sujetos. En esta selección no se ha tenido en cuenta la opinión de los profesionales médicos, ya que no se quería entrenar un algoritmo de clasificación automática con características que los profesionales creen como significativa para ellos mismos. Es el ciclo de vida de esta tesis, el que **desde principio a fin sean los algoritmos los que decidan tanto las características importantes, como la clasificación de los sujetos.**

Por tanto, como se puede comprobar las características seleccionadas provenientes de las imágenes DTI (spleniumTempR, postFIFOL y LIFR), son las relacionadas con la parte cerebral dedicada a la lectura, por lo que refuerza el hecho de que se puede realizar una correcta clasificación de los 3 grupos sometidos al estudio [GZGCS⁺16].

Caso de estudio 2: Migraña

Se parte de un total de 41 características entre las que se encuentran diferentes zonas del cerebro y datos de test específicos.

La entrada tanto para la selección de características como para los algoritmos de clasificación automática es la correspondiente con los valores numéricos de FA de las imágenes DTI correspondientes a las regiones de interés; y los valores numéricos provenientes de los test específicos de los que se obtiene la puntuación por cada una de las preguntas realizadas.

10 de las 41 características se corresponden con zonas cerebrales relativas a:

- Corpus Callosum
- Body Corpus Callosum
- Fasciculus Occipitofrontalis Inferior
- Genu (Anterior Corpus Callosum)
- Insula
- Lobus Frontalis

Y las 31 características restantes se corresponden a preguntas de los siguientes test específicos:

- SF-36
- MSQoL
- MIDAS
- BDI
- BAI
- MDQ-H

5. RESULTADOS

Tabla 5.3: Número y porcentaje de reducción de características (Migraña)

Algoritmo	Número de características	Porcentaje de reducción
Gradient Tree-Boosting	4	90 %
L1-based	21	48 %
Random Forest	12	70 %
Univariate	10	75 %

Tabla 5.4: Características seleccionadas (Migraña)

Característica	Frecuencia (Todas = 4)
Días totales de dolor	4
Número de analgésicos	4
Puntuación MSQoL	3
Left Uncinate	3
Left Cingulate Gyrus	2
Puntuación MDQ-H	2
Días de dolor en el Mes 1	2

Se aplica la selección de características con 4 algoritmos: Gradient Tree-Boosting, L1-based feature selection, Random Forest y Univariate, los cuales llegan a una **reducción de características del 90 %** en el caso de Gradient Tree-Boosting.

De estos resultados, se deduce que las características que no ha seleccionado, o bien son repetitivas en cuanto al valor que aportan, o bien porque no son diferenciadoras en base a los grupos para clasificar.

En este caso de estudio, se puede observar cómo existe una mayor diferencia (90 % el de mayor reducción, y 48 % el de menor reducción) en cuanto al número de características seleccionadas por los algoritmos.

Una vez obtenidas las características por cada uno de los algoritmos, se aplica de nuevo el método de Comité para obtener una selección final de características.

De esta selección, se obtienen un total de **7 características finales** (2 de ellas de zonas cerebrales), las cuales son significativas a la hora de la posterior clasificación de los sujetos. De nuevo, en esta selección no se ha tenido en cuenta la opinión de los profesionales médicos, ya que no se quería entrenar un algoritmo de clasifica-

Tabla 5.5: Porcentaje de acierto en algoritmos con todo el conjunto de datos (Dislexia)

Algoritmo	Accuray	Precision	Recall	F1 Score
SVM	65 %	67.4 %	65 %	64.8 %
Adaboost	52.5 %	55 %	52.5 %	52.3 %
Naive Bayes	52.5 %	54.7 %	52.5 %	51.6 %

ción automática con características que los profesionales creen como significativa para ellos mismos.

En conclusión, atendiendo a las zonas de interés del cerebro seleccionadas por los algoritmos, se encuentran relacionadas con Left Uncinate, Left Cingulate Gyrus, siendo zonas implicadas en la modulación del dolor, la emoción y la percepción. Los cambios son más pronunciados en el lado derecho, y pueden contribuir al dolor la cronificación. Por tanto, se puede concluir que los algoritmos de selección de características han escogido correctamente las características, además reforzando el hecho de que el clasificador va a tener un buen ratio de clasificación bajo los tres grupos del estudio [GBOZ⁺16] .

5.2.3 Algoritmos Machine Learning (C)

Una vez se han obtenido las características significativas, se procede a detallar los resultados de la clasificación de los algoritmos de Machine Learning.

Para la evaluación de los mismos, se reportan los resultados de los ratios de acierto de cada uno de ellos en 2 situaciones: previamente a la selección de características, es decir, teniendo en cuenta la totalidad; y posteriormente a la selección de características.

Caso de estudio 1: Dislexia

Para el caso de Dislexia, se han utilizado 3 algoritmos de Machine Learning, entre los que se encuentran SVM, Adaboost y Näive Bayes. El mejor ratio de clasificación con la matriz de todas las características (50) es de un **65 % de ratio de acierto en la clasificación (SVM)** de los sujetos en los 3 grupos de esta patología (controles, sujetos con dislexia y sujetos con visión monocular).

De estos resultados, se deduce que las características seleccionadas no son muy diferenciadoras para poder obtener un buen ratio de clasificación. Esto puede ser

5. RESULTADOS

Tabla 5.6: Porcentaje de acierto en algoritmos con la selección de características (Dislexia)

Algoritmo	Accuray	Precission	Recall	F1 Score
SVM	64.5 %	68.6 %	64.5 %	64.1 %
Adaboost	63 %	67.3 %	63 %	63.5 %
Naive Bayes	52.5 %	54.7 %	52.5 %	51.6 %

Tabla 5.7: Porcentaje de acierto en algoritmos con todo el conjunto de datos (Migraña)

Algoritmo	Accuray	Precission	Recall	F1 Score
SVM	89.47 %	90.4 %	89.3 %	89.3 %
Adaboost	94.73 %	95.4 %	94.9 %	94.9 %
Naive Bayes	69.47 %	71 %	69.5 %	67.6 %

debido a que las características no son las adecuadas, o que el conjunto de las mismas están confundiendo a los algoritmos y no permiten el aprendizaje correcto.

En la segunda fase, se obtienen los mismos resultados, pero en esta ocasión introduciendo las características seleccionadas, las 7 enunciadas previamente que se componen de 3 relativas a zonas cerebrales, y 4 de ellas de test específicos.

El mejor ratio de clasificación con la matriz de las 7 características seleccionadas es de un **64.5 % (SVM)** de los sujetos en los 3 grupos.

Para este caso de estudio, el entrenar a los algoritmos con todas las características, o las significativamente importantes no tiene ninguna mejora importante.

Pese a no ser unos ratios de clasificación muy buenos, es un logro el poder conseguir estos ratios, ya que es una patología que apenas tienen daños en la materia blanca del cerebro.

Caso de estudio 2: Migraña

Para el caso de Migraña, se han utilizado 3 algoritmos de Machine Learning, entre los que se encuentran SVM, Adaboost y Nãive Bayes.

El mejor ratio de clasificación es de un 94.73 % (Adaboost) con la matriz de todas las características obtenidas inicialmente (41), de los 3 sujetos de esta patología (controles, migraña esporádica y migraña crónica).

Tabla 5.8: Porcentaje de acierto en algoritmos con la selección de características (Migraña)

Algoritmo	Accuray	Precision	Recall	F1 Score
SVM	95.26 %	96.1 %	95.3 %	95.2 %
Adaboost	95.79 %	96.9 %	95.9 %	95.7 %
Naive Bayes	92.63 %	94 %	92.3 %	92.4 %

En esta ocasión, se puede deducir que las características obtenidas en un inicio son las correctas para conseguir buenos ratios de clasificación en 2 de los 3 algoritmos, teniendo peores resultados en Näive Bayes.

En la segunda fase para esta caso de estudio, se vuelve a realizar la clasificación pero únicamente teniendo en cuenta las 7 características significativas obtenidas de la fase de Selección de características.

El mejor ratio de clasificación para esta nueva matriz es de **95.79 % (Adaboost)** con los 3 sujetos de esta patología.

Los resultados obtenidos, deducen que puede llegar a tener vital importancia el realizar una selección de características, sobre todo en el caso del algoritmo de Näive Bayes, llegando a una **mejora de 23.16 puntos porcentuales**.

Bien es cierto, que las características en esta patología son más deterministas para conseguir un buen ratio de clasificación en todos los algoritmos parametrizados.

5.2.4 Comité (D)

Una vez se tienen los resultados de los diferentes algoritmos de Machine Learning, se aplica el bloque del Comité, gracias al cual se tiene en cuenta la opinión de los 3 algoritmos, y seleccionando como resultado final la clasificación que coincida con al menos 2 de los 3 algoritmos.

Tomando como referencia los resultados de la matriz resultante de aplicar el método de Selección de características, los resultados son los siguientes.

Caso de estudio 1: Dislexia

Para el caso de Dislexia aplicando el método de Comité, el ratio de clasificación con la matriz 7 características es de un **72.6 % de ratio de acierto en la clasifica-**

5. RESULTADOS

Tabla 5.9: Porcentaje de acierto con el comité (Dislexia)

Algoritmo	Accuray	Precision	Recall	F1 Score
Comité	72.6 %	69.2 %	68.4 %	65.7 %

ción de los sujetos en los 3 grupos de esta patología (controles, sujetos con dislexia y sujetos con visión monocular).

Partiendo del set inicial con todos los sujetos y todas las características tanto de los cuestionarios como valores numéricos obtenidos del procesado de imágenes DTI, se ha obtenido un porcentaje de acierto de clasificación entre los tres grupos del 65 % con el clasificador SVM que es el que mejor ratio ha aportado de los tres sometidos al estudio.

El clasificador SVM ofrece un método eficiente para la clasificación de las muestras, pese a no haber obtenido un porcentaje alto ya que el introducir al clasificador todas las características sin ningún filtro, pueden confundir al clasificador. Pese a ello, dicho clasificador tiene una base sólida en cuanto al aprendizaje estadístico, permitiendo la optimización de la función de decisión en el proceso de entrenamiento [GPAT03] .

En el estudio de Hoeft [HMB⁺11] , obtienen un 72 % de acierto, aunque en dicho estudio existe una diferencia significativa en la materia blanca, mientras que en este estudio además se ha incluido al grupo de Visión monocular, complicando la labor al clasificador.

A la hora de aplicar los métodos de selección de características, se ha podido observar que en todos los casos exceptuando con los métodos Gradient Tree-Boosting y Tree Based aplicados al clasificador SVM, existe un ligero decrecimiento del ratio de acierto. Pese a ello aplicando Adaboost y Naive Bayes existe una mejoría significativa aplicando los métodos de selección de características.

Respecto a las características seleccionadas por dichos métodos, gracias al comité de selección de características diseñado se puede observar una mezcla de características pertenecientes a los cuestionarios realizados y a los valores numéricos de zonas específicas de dislexia obtenidas de las imágenes DTI. Por otro lado, no se ha encontrado ningún artículo en el cual se aplique este tipo de métodos de selección de características para poder contrastar este resultado. Por otro lado en

Tabla 5.10: Porcentaje de acierto con el comité (Migraña)

Algoritmo	Accuray	Precission	Recall	F1 Score
Comité	96 %	95.9 %	95.3 %	96.7 %

otro estudio [IKG⁺10] , la extracción de características implica una mejoría en los ratios de clasificación. Si las características que “describen” a los sujetos son independientes de las demás, el porcentaje de acierto de clasificación será mayor una vez se haya reducido el set de características. Al igual que en los estudios [ETPZ09, JUA05] , se escogen varios métodos y se compara con la matriz de características inicial, mejorando en casi todos los casos la clasificación al mejorar dichos métodos. Bien es cierto que los porcentajes de mejora no son excesivamente altos, pero al estar estudiando una patología en la cual no existe una degeneración de la materia blanca, y al no contar con una gran base de datos, se podría llegar a extrapolar que con un mayor número de muestras, los porcentajes también ascenden.

Por último, si se aplica también el comité de clasificadores, teniendo en cuenta las clasificaciones de las muestras de cada uno de ellos, se ha conseguido un porcentaje de acierto del 72 %, lo cual es un aumento significativo del porcentaje de acierto inicial.

Caso de estudio 2: Migraña

Para el caso de Migraña aplicando el método de Comité, el ratio de clasificación con la matriz 7 características es de un **96 % de ratio de acierto en la clasificación** de los sujetos en los 3 grupos de esta patología (controles, migraña esporádica y migraña crónica).

En esta ocasión, se observa cómo aplicando el comité, apenas existe mejora en el ratio de clasificación. Esto se debe a que ya en esta patología, los ratios de clasificación de los 3 algoritmos cada uno por separado, ya eran unos ratios muy buenos. Por lo que moviéndose en dichos ratios de acierto, apenas se nota la mejora.

Aun así, existe una mejora, la cual queda validada además en el caso de estudio de Dislexia, donde existía una mejora significativa aplicando este método.

Distinguir pacientes con migraña esporádica, la migraña crónica, y los pacientes con riesgo de medicación en el uso excesivo es posible mediante técnicas de

5. RESULTADOS

selección de características y el aprendizaje automático. Se obtuvo una clasificación con más de 93 % de precisión en el caso de los tres clasificadores incluidos en este artículo (SVM, Boosting and Näive Bayes). Además, el conjunto inicial de datos que contiene 41 características (cuestionarios e imágenes DTI) se redujo cuando se llevó a cabo la función selección de las 7 características más importantes con un combinación de imágenes de DTI y cuestionarios relacionados a la emoción y a la cognición. La clasificación mejoró en un 28 % en el caso de la clasificación de Näive Bayes debido a esta selección de características.

Así, este método puede clasificar a los pacientes gracias a un pequeño conjunto de características de cuestionarios específicos relacionados con la emoción y la cognición, combinado con características obtenidas de imágenes DTI resultantes de la selección previa de las más características importantes.

Esta selección de características es efectiva cuando se ignoran las que se consideran no relevantes dentro del conjunto, o las que se consideran redundantes. De esta manera, la precisión del clasificador mejora al aplicar estos tipos de métodos, obteniendo un sistema más robusto y preciso, y evitando cualquier confusión debido a las características sin importancia [GCGZGB⁺17].

Se observa cómo aplicando el comité, se mejora sustancialmente el ratio de acierto para este caso en el que los porcentajes no eran muy altos. Por tanto, se valida este método, ya que hace que mejore la robustez de los algoritmos cada uno por separado.

Las características fueron seleccionadas en este estudio para llevar a cabo una mejor clasificación al diferenciar entre el sujeto que sufren de patologías de migraña esporádicas y crónicas y el uso excesivo de medicamentos.

Desde un conjunto completo inicial que comprende todos los sujetos y todas las características pertenecientes a ambos cuestionarios psicológicos, datos relativos a los días con dolor, cantidad de los analgésicos y los valores obtenidos de las imágenes del DTI, nosotros obtuvo una precisión de clasificación del 90 % entre los tres grupos sujetos a estudio en los tres algoritmos estudiados de clasificación automática. Aun así, el clasificador SVM ofrece una eficiente método para clasificar las muestras, ya que tiene fundamentos sólidos en términos de aprendizaje estadístico, permitiendo así optimizar la función de decisión en el proceso de entrenamiento [GPAT03].

5.2 Evaluación de los resultados

No hay ningún estudio en el que las técnicas de aprendizaje de la máquina han sido aplicado para clasificar la patología de la migraña, aunque aplicado en otros estudios a otras patologías, por ejemplo, en la estudio realizado por Dyrba, en el que la máquina usada técnicas de aprendizaje específicamente el clasificador SVM para clasificación de los sujetos a través de imágenes DTI pertenecientes a Enfermedad de Alzheimer [DEW⁺13] . En esta ocasión, se obtuvo 80,3 % de precisión, logrando un mejor porcentaje que en este artículo en particular, ya que no sólo se hizo la clasificación con las características obtenidas de las imágenes DTI, sino que se añadieron cuestionarios específicos para la migraña. El clasificador SVM fue usado en otro estudio de Ingalhalikar [IKG⁺10] , aunque en esta ocasión para la clasificación de la esquizofrenia - también a través de imágenes DTI. Los autores obtuvieron una precisión del 90,62 % tras un proceso de selección de características.

De lo que se puede observar obteniendo características del DTI imágenes es un buen método para hacer la clasificación, facilitando el trabajo del especialista sin la necesidad de que visualmente revisar las imágenes de todos los sujetos para determinar si que sufren de la patología o que están sanos. Cualquiera innecesario De esta manera se evitan los costos. Además, y como en el estudio, la selección de las características implica una mejora en la precisión de los clasificadores [IKG⁺10] .

Si las características que describen a los sujetos son independientes de en términos de la patología de la migraña, luego la clasificación posterior tendrá unos mayores porcentajes de acierto.

A pesar de estos buenos resultados, cabe señalar que el estudio no tiene un gran número de muestras en la totalidad con menos de 20 sujetos por grupo. Por lo tanto, el problema de la sobrecarga puede estar dado por la baja número de muestras para el entrenamiento, y la clasificación el algoritmo puede ser ajustado a características realmente específicas de la datos de entrenamiento. Por lo tanto, en el futuro las líneas incluirán la hecho de aumentar el número de muestras para evitar estos problemas y el objetivo principal será apostar el porcentaje de la precisión del clasificador.

Los autores han utilizado los mismos cuestionarios y evaluaciones clínicas de la máquina aprendiendo a usarla antes para estableciendo el diagnóstico clínico.

5. RESULTADOS

Este hecho podría tal vez introducir una cierta circularidad que los algoritmos ciertamente encontrar estas características para diferenciar mejor el tema grupos.

Debido al pequeño número de participantes, las pruebas se llevó a cabo en el mismo conjunto de datos utilizado para el entrenamiento y no se usó ningún equipo de prueba de retención. Por esta razón, el Las medidas de rendimiento pueden no generalizarse a los datos no vistos. Por consiguiente, también se puede llegar a la conclusión de que hay características que influyen en los clasificadores, es decir que la precisión puede ser menor cuando se introducen todos los en el clasificador. Llevamos a cabo una selección de características utilizando una gama de métodos para llevar a cabo una robusta estudio. Como en el caso de los estudios [JUA05, ETPZ09] , varios métodos y se hizo una comparación con el conjunto de datos que contiene las características originales, mejorando la clasificación en casi todos los casos al aplicar esos métodos.

Cuando se crea el comité de selección de características y no se tiene en cuenta aquellos rasgos que no alcanzaron el umbral, mejorando la precisión en el caso de los tres clasificadores. Por lo tanto, se puede demostrar que las características que existían a lo largo de el conjunto de datos que o bien no tenía la importancia necesaria para el propósito de hacer una clasificación adecuada o porque se consideraron redundantes. Aunque la precisión obtenido a través del comité no puede ser el más alto y se obtuvieron específicamente por un algoritmo de selección de características como la potenciación de los árboles de gradiente, el comité de selección de características se considera que logra un sistema más robusto como el las mejores características de selección de entre cuatro características de selección algoritmos - sin tener en cuenta la opinión de sólo uno de ellos.

También hay que tener en cuenta que la combinación de características derivadas de las imágenes DTI y características obtenido de cuestionarios específicos sobre esta patología proporciona una clasificación robusta y fiable. En otras palabras, las imágenes DTI y los cuestionarios complementan entre sí, el hecho de que las dos partes se consideren clínicamente vital para poder establecer un resultado en cuanto a si el sujeto sufre de migraña o no.

Esto ocurre en el estudio realizado por Yan Q. [YYH⁺09] , en que hace uso de imágenes de resonancia magnética y datos de los cuestionarios para predecir la tráquea asistida por la clínica intubación usando múltiples características.

Tabla 5.11: Porcentaje de acierto algoritmo Deep Learning con todas las características (Dislexia)

Algoritmo	Accuray	Precission	Recall	F1 Score
CNN	80 %	72 %	65 %	65 %

Tabla 5.12: Porcentaje de acierto algoritmo Deep Learning con selección de características (Dislexia)

Algoritmo	Accuray	Precission	Recall	F1 Score
CNN	80 %	76 %	75 %	74 %

5.2.5 Algoritmos Deep Learning (E)

En los resultados de Deep Learning, se hace el mismo estudio que con los algoritmos de Deep Learning. Pero en esta ocasión, únicamente validaremos el algoritmo diseñado con este tipo de técnica, Convolutional Neural Network.

Caso de estudio 1: Dislexia

Para el caso de Dislexia y haciendo uso del algoritmo CNN, el ratio de clasificación con la matriz con todas las características (50) es de un **80 % de ratio de acierto en la clasificación** de los sujetos en los 3 grupos de esta patología (controles, sujetos con dislexia y sujetos con visión monocular).

Se observa cómo haciendo uso de técnicas de Deep Learning, se mejora sustancialmente el ratio de acierto para este caso respecto a los ratios de acierto que han emitido los algoritmos de Machine Learning. Por tanto, se puede decir que las características seleccionadas son las correctas, aunque con la técnica de Machine Learning no lo pareciera en exceso. Se consigue un ratio de acierto muy bueno para una patología en la que apenas existe daño en la materia blanca del cerebro.

En la segunda fase, en la que se introducen al algoritmo la matriz de características tras la Selección de características, el ratio de clasificación con dicha matriz (7) es de un **80 % de ratio de acierto en la clasificación** de los sujetos en los 3 grupos de esta patología.

Realmente no ha habido mejora en cuanto al ratio de acierto general haciendo uso de las características seleccionadas, pero se puede observar una mejora sustancial en las otras métricas estadísticas. Lo cual quiere decir, que dicha selección es la

5. RESULTADOS

Tabla 5.13: Porcentaje de acierto algoritmo Deep Learning con todas las características (Migraña)

Algoritmo	Accuray	Precission	Recall	F1 Score
CNN	94.73 %	89 %	84 %	84 %

Tabla 5.14: Porcentaje de acierto algoritmo Deep Learning con selección de características (Migraña)

Algoritmo	Accuray	Precission	Recall	F1 Score
CNN	98.45 %	93.8 %	92.4 %	89.7 %

correcta para que el algoritmo aprenda mejor y no emita resultados de clasificación erróneos.

Caso de estudio 2: Migraña

Para el caso de Migraña y haciendo uso del algoritmo CNN, el ratio de clasificación con la matriz con todas las características (41) es de un **94.73 % de ratio de acierto en la clasificación** de los sujetos en los 3 grupos de esta patología (controles, migraña esporádica y migraña crónica).

Se observa cómo haciendo uso de técnicas de Deep Learning, el ratio de acierto es muy parecido a los que emitieron los algoritmos de Machine Learning. Pese a ello, son unos resultados formidables para la clasificación de esta patología.

En la segunda fase, en la que se introducen al algoritmo la matriz de características tras la Selección de características, el ratio de clasificación con dicha matriz (7) es de un **98.45 % de ratio de acierto en la clasificación** de los sujetos en los 3 grupos de esta patología.

Vuelve a existir una mejora en cuanto se le introducen al algoritmo las características tras la fase de selección. Por tanto, se vuelve a validar esta metodología para mejorar la clasificación de los sujetos.

5.3 Comparativa de Machine Learning vs. Deep Learning

Tabla 5.15: Comparativa de porcentajes de acierto Machine Learning vs. Deep Learning

	Accuracy	Precision	Recall	F1 Score
Dislexia				
Comité (ML)	72.6 %	69.2 %	68.4 %	65.7 %
CNN (DL)	80.1 %	76.5 %	75.2 %	74.7 %
Migraña				
Comité (ML)	96 %	95.9 %	95.3 %	96.7 %
CNN (DL)	98.45 %	93.8 %	92.4 %	89.7 %

5.3 Comparativa de Machine Learning vs. Deep Learning

Tras la obtención de todos los resultados, tanto de los de selección de características, algoritmos de Machine Learning y algoritmos de Deep Learning, se procede a su comparativa. Obteniendo de esta manera unas conclusiones respecto a qué técnica es la adecuada.

5.3.1 Ratios de acierto

En primer lugar, si se tiene en cuenta el entrenar el algoritmos con la matriz completa de todos los algoritmos, o por el contrario únicamente enseñarle con la matriz de características seleccionadas, se puede concluir que los resultados son significativamente mejores cuando se aplica esta metodología de selección.

Se llega a una **mejora de alrededor de 20 puntos porcentuales** en el caso del caso de estudio de la Migraña.

Si se atiende a los resultados obtenidos por los algoritmos de clasificación automática, teniendo en cuenta los ratios de acierto resultantes de entrenar la matriz con las características seleccionadas, se consigue un ratio de acierto superior al **95 % de acierto** para los dos casos de estudio analizados en esta tesis.

Por lo que atendiendo a los resultados de clasificación, las dos técnicas, tanto la tradicional de Machine Learning, como la novedosa y actual Deep Learning, para estos casos de estudio son bastante parejas.

5. RESULTADOS

Pese a ello, se podría decantar en aplicar técnicas de Deep Learning, ya que el estado del arte relacionado, expone unas mejoras significativas en el uso de esta técnica respecto a la tradicional.

5.3.2 Recursos Hardware

Una vez vista la comparativa en cuanto a los ratios de acierto de las dos técnicas utilizadas en esta tesis, Machine y Deep Learning, se procede a realizar una comparativa en cuanto a los recursos Hardware necesarios para su implementación y ejecución.

Para esta tesis y debido a que el número de sujetos que se han obtenido para los 2 casos de estudio (no más de 60 sujetos), para la ejecución de los algoritmos de Machine Learning y de Deep Learning ha sido suficiente mediante la CPU del ordenador.

Para los algoritmos de Machine Learning, si el número de sujetos a entrenar no es muy excesivo, bastándose con la CPU del ordenador es correcto. Aunque a medida que crezcan las muestras de la base de datos utilizada, irá aumentando linealmente el tiempo necesario para el entrenamiento de los algoritmos.

Por el contrario, si se hace uso de algoritmos de Deep Learning, debido a sus características de uso de redes neuronales, requieren cálculos más complejos en los que una CPU empieza a ser más lenta. En estos casos, es recomendable ejecutarlos a través de la GPU del ordenador, pudiendo ser entre 2 y 3 veces más rápida en solucionar los cálculos necesarios que una CPU. Además, respecto al tamaño de la base de datos con la que se quiera entrenar, en el caso de que sea un número alto de muestras, a través de la CPU será imposible realizar el entrenamiento de los algoritmos.

Por tanto, es recomendable para entrenar modelos, bien sean de Machine o Deep Learning tener una tarjeta gráfica con la que hacer el procesado de los cálculos, pero al ser excesivamente caras, se puede hacer uso de algoritmos de Machine Learning para poder realizar este tipo de clasificaciones automáticas.

5.3 Comparativa de Machine Learning vs. Deep Learning

Tabla 5.16: Comparativa de tiempo de entrenamiento Machine Learning vs. Deep Learning

	Tiempo de entrenamiento (aproximadamente)
CPU	
Machine Learning	4 horas
Deep Learning	8 horas
GPU	
Machine Learning	2 horas
Deep Learning	4 horas

5.3.3 Tiempos de procesamiento

Derivado del punto anterior, también se puede realizar una comparativa en cuanto a los tiempos de procesamiento de los algoritmos.

Como ya se ha mencionado en el estado del arte, los tiempos computacionales al aplicar algoritmos de Deep Learning son mucho mayores que los que requieren los algoritmos de Machine Learning, todo ello debido a los cálculos internos que realizan los algoritmos. Mucho tiene que ver en que los algoritmos de Deep Learning hacen uso de estructuras de red de neuronas, las cuales son muy parecidas a las que tiene el ser humano. Para esta tesis, los tiempos han sido los siguientes:

Se puede observar que haciendo uso de la CPU del ordenador, y únicamente para alrededor de 60 muestras para el entrenamiento, se tiene una diferencia de 4 horas si se aplican técnicas de Machine Learning o de Deep Learning.

Por otro lado también existe una diferencia significativa de unos 2 horas haciendo uso de una GPU.

Si únicamente se comparan los tiempos de técnicas de Deep Learning, se puede reducir hasta la mitad si se utiliza una GPU que una CPU. Además, los tiempos de entrenamiento de las técnicas de Deep Learning tienen su diferencia en función de lo compleja que sea la red que tenga el algoritmo. Es decir, no es lo mismo entrenar un modelo el cual contenga 4 capas ocultas, que si por el contrario tiene 20 capas ocultas. Evidentemente los tiempos computacionales se dispararían debido a la complejidad de los cálculos que se tendrían que realizar. Y todos estos tiempos

5. RESULTADOS

aumentarían también al incrementar el número de muestras utilizadas (pudiendo hablar de millones).

5.3.4 Conclusión

Para esta tesis, la conclusión obtenida teniendo en cuenta todas las comparativas realizadas en cuanto al ratio de acierto, hardware necesario y tiempos de entrenamiento, es que con los algoritmos de Machine Learning es suficiente para conseguir buenos resultados en poco tiempo, y sin la necesidad de un coste alto.

Pese a ello, los avances que están teniendo los algoritmos de Deep Learning, incitan a utilizarlos y dejar atrás la técnica tradicional de Machine Learning. Teniendo en cuenta además que empresas están invirtiendo en implantar Deep Learning en todo su conocimiento, y dada la cantidad de datos disponibles para poder entrenar y realizar predicciones, hay que mirar al futuro e implantar estas técnicas desde el inicio de los estudios de investigación.

5.4 Sumario

En este capítulo se ha podido observar los resultados obtenidos a partir del diseño descrito en el capítulo 4. Los resultados se han dividido en 5 grandes bloques, cada uno de ellos con un propósito diferente para la realización de esta tesis.

En primer lugar se han obtenido los resultados del pre-procesamiento de las imágenes DTI utilizadas. En segundo lugar, se han obtenido los datos relacionados con la extracción de características de cada uno de los 2 casos de estudio analizados. Posteriormente, se han extraído los ratios de acierto de la clasificación de los algoritmos de Machine Learning tanto para el caso de estudio de Dislexia como para el de Migraña. Seguidamente se han detallado los resultados de la técnica llamada Comité que se ha aplicado a este tipo de técnica. Seguidamente, se precisan los ratios de acierto de la clasificación haciendo uso de la técnica de Deep Learning, también para cada uno de los 2 casos de estudio analizados. Y por último se ha descrito una comparativa entre las técnicas de Machine Learning y Deep Learning en cuanto a los ratios de acierto, los recursos Hardware necesarios, y

los tiempos de procesamiento tanto para el entrenamiento como para el test de los propios algoritmos.

Conclusiones

En este capítulo se detallan las conclusiones finales de la tesis a partir de la hipótesis enunciada de partida, junto con el objetivo principal y los objetivos específicos asociados, los cuales fueron indicados en el capítulo de Introducción. La hipótesis de partida de esta tesis era la siguiente:

Es posible desarrollar un método de clasificación de patologías cerebrales a través del análisis de imágenes de resonancia magnética basado en tractografía, mediante métodos de Machine Learning y Deep Learning.

Durante el desarrollo de esta tesis, se ha validado la hipótesis gracias a los capítulos de esta tesis de Diseño y Resultados, donde se describen y detallan tanto los bloques necesarios para la validación de la propia hipótesis, como los resultados finales que se obtuvieron en su demostración.

Gracias a las técnicas de Inteligencia Artificial, específicamente las de selección de características, Machine Learning y Deep Learning, junto con las imágenes DTI, **se ha creado un nuevo método capaz de clasificar sujetos a partir de la materia blanca del cerebro** que se han analizado en esta tesis. Además, los métodos utilizados para la validación de la hipótesis se encuentran validados.

Los capítulos de Diseño y Resultados de esta tesis, han sido clave en la investigación, donde se han podido validar gracias a los 2 casos de estudios de diferentes

6. CONCLUSIONES

patologías estudiados, comprobando la robustez de las características utilizadas, y la validación de los métodos utilizados.

El objetivo principal descrito en el capítulo de Introducción ha sido el siguiente:

Clasificar patologías cerebrales mediante el análisis de imágenes de resonancia magnética a través de algoritmos de Machine Learning y Deep Learning.

Como se enuncia, el objetivo principal de esta investigación era la clasificación de patologías cerebrales a partir de imágenes de resonancia magnética DTI, pero remarcando que ha sido a través de un método válido que proporciona resultados consistentes.

Este objetivo se ha tenido en cuenta a lo largo de esta tesis, ya que era necesario que los resultados obtenidos fueran fiables gracias a los bloques diseñados relacionados con el **ratio de acierto en la clasificación** a través de técnicas de Machine Learning y Deep Learning. Además para cada una de las técnicas utilizadas, se han utilizado diferentes algoritmos para comprobar la fiabilidad de los mismos.

A partir del objetivo principal, en esta tesis se enunciaban un conjunto de objetivos secundarios, siendo los siguientes:

- OS1: Extraer características de las imágenes de resonancia magnética y de test específicos para cada patología. El primer bloque del diseño (Bloque A: Pre-procesado de imágenes DTI), es el encargado de la obtención de las diferentes características (zonas cerebrales y test específicos de cada uno de los 2 casos de estudio), las cuales serán utilizadas por los bloques posteriores. Sin estas características, no podríamos validar la hipótesis.
- OS2: Diseñar un método que extraiga las características más relevantes a partir de las iniciales obtenidas de las imágenes de resonancia magnética y de los test específicos. El segundo bloque (Bloque B: Selección de características), es el que permite obtener un conjunto de características, obtenidas a partir del bloque previo, las cuales son significativas para la correcta clasificación de los sujetos. Este bloque, puede ser utilizado por cualquier patología.
- OS3: Diseñar un método que provea la clasificación de sujetos en función de su patología. El segundo, tercer y quinto bloque (Bloque B: Selección

de características, Bloque C: Algoritmos de Machine Learning y Bloque E: Algoritmos de Deep Learning), se han centrado en la selección de diversos algoritmos capaces de obtener gracias a su parametrización unos resultados fiables en cuanto al ratio de acierto.

- OS4: Validar los algoritmos de Machine Learning y Deep Learning aplicados a las características obtenidas de las imágenes de resonancia magnética. El bloque C (Algoritmos de Machine Learning) y bloque E (Algoritmos de Deep Learning), son los responsables de la validación de los algoritmos de selección de características, ya que el ratio de acierto de cada uno de los mismos, indica si dicha selección de características aportada por el bloque B (Selección de características) es la correcta.
- OS5: Validar la confiabilidad de los algoritmos de Machine Learning y Deep Learning tras la selección de características más relevantes a partir de este tipo de algoritmo. El bloque C (Algoritmos de Machine Learning) y bloque E (Algoritmos de Deep Learning), son los encargados de la validación de la confiabilidad. Al hacer uso de diversos algoritmos, la confiabilidad se mide a través de la variación de los resultados de cada uno de ellos. Esta variación debe ser mínima para que la confiabilidad en los propios algoritmos sea alta.
- OS6: Validar la confiabilidad de los algoritmos de Machine Learning y Deep Learning aplicando el método “ensemble” a partir de los resultados obtenidos por cada uno de los algoritmos. El bloque D (Comité) es el que proporciona una robustez en cuanto a los resultados obtenidos de los diferentes algoritmos de Machine Learning. Por tanto, se obtiene un único resultado a partir del resultado de clasificación de cada uno de los algoritmos que componen el Comité.

6.1 Conclusiones técnicas

Las conclusiones técnicas derivadas de la realización de esta tesis, se ha contribuido a la realización de un **nuevo método, el cual permite unificar técnicas de Machine Learning y Deep Learning para la clasificación de sujetos a través de**

6. CONCLUSIONES

imágenes DTI y test específicos. Esto permite dotar de un mayor rigor y confiabilidad a los resultados emitidos.

Además de la aportación descrita, se ha enunciado una metodología para la clasificación de sujetos de diferentes patologías, gracias a los 2 casos de estudio analizados. Se han observado los problemas de otros estudios, en los que los algoritmos se encuentran parametrizados para la correcta clasificación de una única patología. Pero a través de algoritmos de selección de características, y la validación de resultados de Machine Learning y Deep Learning, se puede obtener un método de clasificación para diferentes patologías. **Esta confiabilidad conseguida, permite tener un método robusto frente a diferentes patologías,** no sólo adecuando el método para una única.

6.2 Conclusiones médicas

Gracias al diseño y desarrollo del software final implementado a través de los métodos descritos en esta tesis, los médicos especialistas, neurólogos o radiólogos, **reciben un diagnóstico sobre la clasificación del sujeto** en uno de los grupos descritos para cada uno de los 2 casos de estudio.

Esta clasificación, supone una **ayuda al diagnóstico** a los médicos, permitiendo concretar si los sujetos se encuentran afectados con la patología; y permitiendo acelerar el diagnóstico final, **evitando largos análisis realizados por los profesionales para tener un diagnóstico fiable.**

6.3 Impacto científico

A continuación, se muestra el impacto científico derivado de la realización de esta tesis.

El impacto se diferencia en los artículos publicados en revistas científicas, comunicaciones en congresos y proyectos de investigación realizados y enmarcados con la temática de la tesis.

Tabla 6.1: Revistas científicas I

Título	Automatic classification of dyslexic children by applying machine learning to fMRI images
Autores	García-Chimeno Y, García-Zapirain B, Saralegui-Prieto I, Fernández-Ruano B
Revista	Bio-Medical Materials and Engineering
IF (2018)	1.05 (Q3)
Fecha de publicación	2014
DOI	10.3233/BME-141120
Enlace web	https://content.iospress.com/articles/bio-medical-materials-and-engineering/bme1120
Citas	1

Tabla 6.2: Revistas científicas II

Título	Differences in effective connectivity between children with dyslexia, monocular vision and typically developing readers: A DTI study
Autores	García-Zapirain B, García-Chimeno Y, Saralegui-Prieto I, Fernández-Ruano B, Martínez R
Revista	Biomedical Signal Processing and Control
IF (2018)	3.83 (Q2)
Fecha de publicación	2016
DOI	10.1016/j.bspc.2015.07.009
Enlace web	https://www.sciencedirect.com/science/article/abs/pii/S1746809415001342
Citas	1

6.3.1 Artículos en revistas científicas

El listado de artículos se compone con los publicados en las revistas científicas internacionales en la fecha de presentación de esta tesis, es el siguiente.

6.3.2 Comunicación en congresos

El listado de congresos internacionales se compone de las comunicaciones realizadas en la fecha de presentación de esta tesis, es el siguiente.

6.3.3 Proyectos relacionados

El listado de proyectos relacionados directamente con esta tesis y concluidos con anterioridad a la fecha de presentación de esta tesis, es el siguiente.

Tabla 6.3: Revistas científicas III

Título	Right fronto-insular white matter tracts link cognitive reserve and pain in migraine patients
Autores	Gomez-Beldarrain M, Oros I, García-Zapirain B, Fernández-Ruano B, García-Chimeno Y, Cabrera A, Anton-Ladislao A, Aguirre-Larracochea U, García-Monco JC
Revista	The Journal of Headache and Pain
IF (2018)	3.83 (Q2)
Fecha de publicación	2016
DOI	10.1186/s10194-016-0593-1
Enlace web	https://thejournalofheadacheandpain.biomedcentral.com/articles/10.1186/s10194-016-0593-1
Citas	18

6. CONCLUSIONES

Tabla 6.4: Revistas científicas IV

Título	Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data
Autores	Garcia-Chimeno Y, Garcia-Zapirain B, Gomez-Beldarrain M, Fernandez-Ruanova B, Garcia-Monco JC
Revista	BMC Medical Informatics and Decision Making
IF (2018)	2.38 (Q2)
Fecha de publicación	2017
DOI	10.1186/s12911-017-0434-4
Enlace web	https://bmcmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-017-0434-4
Citas	14

Tabla 6.5: Comunicación en congreso I

Título	Dyslexia classification using feature selection and Machine Learning Techniques
Autores	Garcia-Chimeno Y, Garcia-Zapirain B, Saralegui-Prieto I, Fernandez-Ruanova B
Congreso	7th International Work-Conference on Bioinformatics and Biomedical Engineering (IWBBIO 2019)
IF (2018)	2.38 (Q2)
Fecha de publicación	2019
Enlace web	http://iwbbio.ugr.es/ProceedingsExtendedAbstractIWBBIO2019.pdf

6.4 Líneas futuras

Esta investigación propone un conjunto de métodos, pero esta queda abierta para su replicabilidad y mejora.

En esta tesis se han analizado 2 casos de estudio diferentes, pero en el futuro consiste en la reutilización en más patologías diferentes. En el caso de que la replicabilidad en otras patologías con mayor deterioro en la materia blanca del cerebro (como por ejemplo el Alzheimer), ayudaría significativamente en mejorar los resultados de acierto de la clasificación de los sujetos. En esta tesis se han estudiado 2 patologías que no suponen un daño tan generalizado en la materia blanca, como pudiera ser con otras patologías con daños mayores, como por ejemplo la esclerosis múltiple.

Además, con la ampliación de esta tesis a diferentes patologías, se pueden mejorar los algoritmos de selección de características, ya que únicamente se han validado con 2 patologías diferentes. El poder analizar más patologías, ayudaría a enriquecer dichos algoritmos, y ajustar su parametrización.

Tabla 6.6: Proyecto MIGREIN

Proyecto	Disfunción Orbito Frontal en Pacientes con Migraña Crónica y Abuso de Analgésicos. Estudio con Resonancia Magnética Funcional (fMRI)
Autores	Garcia-Chimeno Y, Garcia-Zapirain B, Saralegui-Prieto I, Fernandez-Ruanova B
Acronimo	MIGREIN
Ámbito del proyecto	Nacional
Calidad en la que se ha participado	Investigador posterior a la finalización de la subvención
Entidades financiadoras	Instituto de Salud Carlos III (ISCIII). Programa FIS
Fecha	01/01/2012 - 31/12/2014

Por otro lado, esta investigación puede ampliarse en cuanto a la mejora de los parámetros de los algoritmos utilizados, incluso la valoración de otros algoritmos más novedosos actualmente. Bien se sabe, que las técnicas de Inteligencia Artificial se están especificando más en los algoritmos de Deep Learning, ya que gracias a su aprendizaje más profundo, se observa que pueden llegar a tener mejores resultados en cuanto a la clasificación. Aunque también hay que tener en cuenta, que el gasto computacional, tanto de recursos hardware y de tiempo, es mucho mayor.

Otra vía totalmente abierta de investigación es la unión de imágenes cerebrales pertenecientes a la materia gris y a la blanca (estudiada en esta tesis), es decir, una combinación de las técnicas de fMRI y de DTI, pudiendo mejorar en la precisión de la clasificación.

Por último, mencionar que debido a las patologías que se han analizado en esta tesis, el número de sujetos que comprenden el estudio, no ha sido excesivamente alto, y aun así se han podido obtener buenos resultados. Pese a ello, es de vital importancia aumentar el número de sujetos con los que se entrenan al algoritmo, para que puedan mejorar aún más en su ratio de acierto.

Bibliografía

- [Anz12] Yuichiro Anzai. *Pattern recognition and machine learning*. Elsevier, 2012. [21](#)
- [AOH⁺17] U Rajendra Acharya, Shu Lih Oh, Yuki Hagiwara, Jen Hong Tan, Muhammad Adam, Arkadiusz Gertych, and Ru San Tan. A deep convolutional neural network model to classify heartbeats. *Computers in biology and medicine*, 89:389–396, 2017. [25](#)
- [Arc] Juan Ignacio Barrios Arce. Matriz de confusión y sus métricas. [13](#)
- [BČL13] M Catherine Bushnell, Marta Čeko, and Lucie A Low. Cognitive and emotional control of pain and its disruption in chronic pain. *Nature Reviews Neuroscience*, 14(7):502–511, 2013. [28](#)
- [BJSJ19] Thierry Bouwmans, Sajid Javed, Maryam Sultana, and Soon Ki Jung. Deep neural network concepts for background subtraction: A systematic review and comparative evaluation. *Neural Networks*, 117:8–66, 2019. [24](#)
- [BRH04] James F Bowring, James M Rehg, and Mary Jean Harrold. Active learning for automatic classification of software behavior. *ACM SIGSOFT Software Engineering Notes*, 29(4):195–205, 2004. [13](#)
- [Bru16] Verónica Bruno. Introducción a la neurología global. *Neurología Argentina*, 8(3):227–228, 2016. [2](#)

BIBLIOGRAFÍA

- [CCR⁺20] Gloria Castellazzi, Maria Giovanna Cuzzoni, Matteo Cotta Ramusino, Daniele Martinelli, Federica Denaro, Antonio Ricciardi, Paolo Vitali, Nicoletta Anzalone, Sara Bernini, Fulvia Palesi, et al. A machine learning approach for the differential diagnosis of alzheimer and vascular dementia fed by mri selected features. *Frontiers in Neuroinformatics*, 14, 2020. [29](#)
- [CNM06] Rich Caruana and Alexandru Niculescu-Mizil. An empirical comparison of supervised learning algorithms. In *Proceedings of the 23rd international conference on Machine learning*, pages 161–168, 2006. [21](#)
- [CWQ⁺20] Qi Chen, Yao Wang, Yage Qiu, Xiaowei Wu, Yan Zhou, and Guangtao Zhai. A deep learning-based model for classification of different subtypes of subcortical vascular cognitive impairment with flair. *Frontiers in Neuroscience*, 14:557, 2020. [29](#)
- [CWZ⁺18] Yunliang Cai, Shaoju Wu, Wei Zhao, Zhigang Li, Zheyang Wu, and Songbai Ji. Concussion classification via deep learning using whole-brain white matter fiber strains. *PloS one*, 13(5):e0197992, 2018. [29](#)
- [DEW⁺13] Martin Dyrba, Michael Ewers, Martin Wegrzyn, Ingo Kilimann, Claudia Plant, Annahita Oswald, Thomas Meindl, Michela Pievani, Arun LW Bokde, Andreas Fellgiebel, et al. Robust automated detection of microstructural white matter degeneration in alzheimer’s disease using machine learning classification of multicenter dti data. *PloS one*, 8(5):e64925, 2013. [89](#)
- [ETPZ09] Pablo A Estévez, Michel Tesmer, Claudio A Perez, and Jacek M Zurada. Normalized mutual information feature selection. *IEEE Transactions on neural networks*, 20(2):189–201, 2009. [87](#), [90](#)
- [FP98] Julie A Fiez and Steven E Petersen. Neuroimaging studies of word reading. *Proceedings of the National Academy of Sciences*, 95(3):914–921, 1998. [27](#)

- [GBOZ⁺16] Marian Gomez-Beldarrain, Isabel Oroz, Begoña Garcia Zapirain, Begoña Fernandez Ruanova, Yolanda Garcia Fernandez, Alberto Cabrera, Ane Anton-Ladislao, Urko Aguirre-Larracoechea, and Juan Carlos Garcia-Monco. Right fronto-insular white matter tracts link cognitive reserve and pain in migraine patients. *The journal of headache and pain*, 17(1):4, 2016. [83](#)
- [GCGZGB⁺17] Yolanda Garcia-Chimeno, Begonya Garcia-Zapirain, Marian Gomez-Beldarrain, Begonya Fernandez-Ruanova, and Juan Carlos Garcia-Monco. Automatic migraine classification via feature selection committee and machine learning techniques over imaging and questionnaire data. *BMC medical informatics and decision making*, 17(1):1–10, 2017. [88](#)
- [Gér19] Aurélien Géron. *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems*. O'Reilly Media, 2019. [24](#)
- [GG05] Cyril Goutte and Eric Gaussier. A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *European conference on information retrieval*, pages 345–359. Springer, 2005. [13](#)
- [GPAT03] Deon Garrett, David A Peterson, Charles W Anderson, and Michael H Thaut. Comparison of linear, nonlinear, and feature selection methods for eeg signal classification. *IEEE Transactions on neural systems and rehabilitation engineering*, 11(2):141–144, 2003. [86](#), [88](#)
- [GZGCS⁺16] Begonya Garcia-Zapirain, Yolanda Garcia-Chimeno, Ibone Saralegui, Begonya Fernandez-Ruanova, and Ricardo Martinez. Differences in effective connectivity between children with dyslexia, monocular vision and typically developing readers: a dti study. *Biomedical Signal Processing and Control*, 23:19–27, 2016. [80](#)

BIBLIOGRAFÍA

- [Hab00] Michel Habib. The neurological basis of developmental dyslexia: an overview and working hypothesis. *Brain*, 123(12):2373–2399, 2000. [26](#)
- [HAU⁺19] Mohammad Mehedi Hassan, Md Golam Rabiul Alam, Md Zia Uddin, Shamsul Huda, Ahmad Almogren, and Giancarlo Fortino. Human emotion recognition using deep belief network architecture. *Information Fusion*, 51:10–18, 2019. [24](#)
- [HHW⁺20] Saeed Hamouda, Amr Hassan, Mohammed E Wahed, Mohammed Ail, and Osama Farouk. Tuning to optimize svm approach for breast cancer diagnosis with blood analysis data. *Available at SSRN 3537067*, 2020. [22](#)
- [HLM⁺19] Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon, Asad Malik, Tanvir Ahmad, Amjad Ali, Shah Nazir, Ijaz Ahad, Muhammad Shahid, et al. Feature selection based on l1-norm support vector machine and effective recognition system for parkinson’s disease using voice recordings. *IEEE access*, 7:37718–37734, 2019. [20](#)
- [HMB⁺11] Fumiko Hoeft, Bruce D McCandliss, Jessica M Black, Alexander Gantman, Nahal Zakerani, Charles Hulme, Heikki Lyytinen, Susan Whitfield-Gabrieli, Gary H Glover, Allan L Reiss, et al. Neural systems predicting long-term outcome in dyslexia. *Proceedings of the National Academy of Sciences*, 108(1):361–366, 2011. [86](#)
- [HZNF07] Pengfei Han, Xiuzhen Zhang, Raymond S Norton, and Zhiping Feng. Reducing overfitting in predicting intrinsically unstructured proteins. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 515–522. Springer, 2007. [16](#)
- [IKG⁺10] Madhura Ingalhalikar, Stathis Kanterakis, Ruben Gur, Timothy PL Roberts, and Ragini Verma. Dti based diagnostic prediction of a disease via pattern classification. In *International*

- Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 558–565. Springer, 2010. [87](#), [89](#)
- [Jor19] Michael I Jordan. Artificial intelligence—the revolution hasn’t happened yet. *Harvard Data Science Review*, 1(1), 2019. [13](#)
- [JUA05] Thanyaluk Jirapech-Umpai and Stuart Aitken. Feature selection and classification for microarray data analysis: Evolutionary methods for identifying predictive genes. *BMC bioinformatics*, 6(1):148, 2005. [87](#), [90](#)
- [L⁺92] Kurt Lewin et al. La investigación-acción y los problemas de las minorías. AA. VV., *La investigación-acción participativa. Inicio y desarrollo, Biblioteca de Educación de Adultos*, 6:13–25, 1992. [6](#)
- [Lat03] A Latorre. Investigación acción. graó, 2003. [6](#)
- [LBCR⁺13] Diana López-Barroso, Marco Catani, Pablo Ripollés, Flavio Dell’Ácqua, Antoni Rodríguez-Fornells, and Ruth de Diego-Balaguer. Word learning is mediated by the left arcuate fasciculus. *Proceedings of the National Academy of Sciences*, 110(32):13168–13173, 2013. [27](#)
- [LBMP⁺01] Denis Le Bihan, Jean-François Mangin, Cyril Poupon, Chris A Clark, Sabina Pappata, Nicolas Molko, and Hughes Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 13(4):534–546, 2001. [26](#)
- [LFL⁺19] Xi Leng, Peng Fang, Huan Lin, Chunhong Qin, Xin Tan, Yi Liang, Chi Zhang, Hongzhuo Wang, Jie An, Donglin Wu, et al. Application of a machine learning method to whole brain white matter injury after radiotherapy for nasopharyngeal carcinoma. *Cancer Imaging*, 19(1):1–10, 2019. [29](#)

BIBLIOGRAFÍA

- [LGF20] Luca Luceri, Silvia Giordano, and Emilio Ferrara. Detecting troll behavior via inverse reinforcement learning: A case study of russian trolls in the 2016 us election. In *Proceedings of the International AAI Conference on Web and Social Media*, volume 14, pages 417–427, 2020. [23](#)
- [LKN19] Ingoo Lee, Jongsoo Keum, and Hojung Nam. Deepconv-dti: Prediction of drug-target interactions via deep learning with convolution on protein sequences. *PLoS computational biology*, 15(6):e1007129, 2019. [29](#)
- [LRDG91] Margaret S Livingstone, Glenn D Rosen, Frank W Drislane, and Albert M Galaburda. Physiological and anatomical evidence for a magnocellular defect in developmental dyslexia. *Proceedings of the National Academy of Sciences*, 88(18):7943–7947, 1991. [27](#)
- [LWC⁺14] Wangchao Lou, Xiaoqing Wang, Fan Chen, Yixiao Chen, Bo Jiang, and Hua Zhang. Sequence based prediction of dna-binding proteins based on hybrid feature selection using random forest and gaussian naive bayes. *PloS one*, 9(1):e86703, 2014. [16](#)
- [LYZ⁺17] Yuan Luo, Wenbin Ye, Xiaojin Zhao, Xiaofang Pan, and Yuan Cao. Classification of data from electronic nose using gradient tree boosting algorithm. *Sensors*, 17(10):2376, 2017. [15](#)
- [MDSCM⁺15] Ana Martins Da Silva, Sara Cavaco, Inês Moreira, Andreia Bettencourt, Ernestina Santos, Cláudia Pinto, Alexandra Gonçalves, Ester Coutinho, Raquel Samões, Cláudia C Dias, et al. Cognitive reserve in multiple sclerosis: protective effects of education. *Multiple Sclerosis Journal*, 21(10):1312–1321, 2015. [28](#)
- [MG09] M Mohammadi and GB Gharehpetian. Application of core vector machines for on-line voltage security assessment using a decision-tree-based feature selection algorithm. *IET generation, transmission & distribution*, 3(8):701–712, 2009. [19](#)

- [MG13] Farooq H Maniyar and Peter J Goadsby. Functional imaging in chronic migraine. *Current pain and headache reports*, 17(5):333, 2013. [27](#)
- [MHS⁺18] P Mikolas, J Hlinka, A Skoch, Z Pitra, E Bakstein, T Frodl, F Spaniel, and T Hajek. 07-connectivity of the anterior insula differentiates participants with first-episode schizophrenia spectrum disorders from controls: A machine-learning study. *Clinical Neurophysiology*, 129(4):e8, 2018. [29](#)
- [NJF19] Bin Ning, Wu Junwei, and Hu Feng. Spam message classification based on the naïve bayes classification algorithm. *IAENG International Journal of Computer Science*, 46(1):46–53, 2019. [23](#)
- [OBK⁺13] Jes Olesen, André Bes, Robert Kunkel, James W Lance, Giuseppe Nappi, Volker Pfaffenrath, Frank Clifford Rose, Bruce S Schoenberg, Dieter Soyka, Peer Tfelt-Hansen, et al. The international classification of headache disorders, (beta version). *Cephalalgia*, 33(9):629–808, 2013. [37](#)
- [OLM⁺12] Laurence O’Dwyer, Franck Lambertson, Silke Matura, Monika Scheibe, Julia Miller, Dan Rujescu, David Prvulovic, and Harald Hampel. White matter differences between healthy young apoe4 carriers and non-carriers identified with tractography and support vector machines. *PLoS One*, 7(4):e36024, 2012. [2](#)
- [ONWK03] Pernille J Olesen, Zoltan Nagy, Helena Westerberg, and Torkel Klingberg. Combined analysis of dti and fmri data reveals a joint maturation of white and grey matter in a fronto-parietal network. *Cognitive Brain Research*, 18(1):48–57, 2003. [12](#)
- [PDF⁺01] Eraldo Paulesu, J-F Démonet, Ferruccio Fazio, Eamon McCrory, Valerie Chanoine, Nicky Brunswick, Stefano F Cappa, Giuseppe Cossu, Michel Habib, Chris D Frith, et al. Dyslexia: cultural diversity and biological unity. *Science*, 291(5511):2165–2167, 2001. [27](#)

BIBLIOGRAFÍA

- [PLD17] Yuliang Pan, Diwei Liu, and Lei Deng. Accurate prediction of functional effects for variants by combining gradient tree boosting with optimal neighborhood properties. *PloS one*, 12(6):e0179314, 2017. [15](#)
- [PPO⁺19a] Seyedmehdi Payabvash, Eva M Palacios, Julia P Owen, Maxwell B Wang, Teresa Tavassoli, Molly Gerdes, Anne Brandes-Aitken, Elysa J Marco, and Pratik Mukherjee. Diffusion tensor tractography in children with sensory processing disorder: Potentials for devising machine learning classifiers. *NeuroImage: Clinical*, 23:101831, 2019. [29](#)
- [PPO⁺19b] Seyedmehdi Payabvash, Eva M Palacios, Julia P Owen, Maxwell B Wang, Teresa Tavassoli, Molly Gerdes, Annie Brandes-Aitken, Pratik Mukherjee, and Elysa J Marco. White matter connectome correlates of auditory over-responsivity: Edge density imaging and machine-learning classifiers. *Frontiers in integrative neuroscience*, 13:10, 2019. [29](#)
- [PPRMdL⁺14] Eduardo J Pedrero-Perez, Gloria Rojo-Mota, José M Ruiz-Sánchez de León, Laura M Fernandez-Mendez, Sara Morales-Alonso, and Ana Prieto-Hidalgo. Cognitive reserve in substance addicts in treatment: relation to cognitive performance and activities of daily living. *Revista de neurologia*, 59(11):481–489, 2014. [28](#)
- [PWB19] Philipp Probst, Marvin N Wright, and Anne-Laure Boulesteix. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 9(3):e1301, 2019. [17](#)
- [PWP⁺15] Sonia Pujol, William Wells, Carlo Pierpaoli, Caroline Brun, James Gee, Guang Cheng, Baba Vemuri, Olivier Commowick, Sylvain Prima, Aymeric Stamm, et al. The dti challenge: toward standardized evaluation of diffusion tensor imaging tractography for neurosurgery. *Journal of Neuroimaging*, 25(6):875–882, 2015. [12](#)

- [QLC⁺18] Jianping Qiao, Yingru Lv, Chongfeng Cao, Zhishun Wang, and Anning Li. Multivariate deep learning classification of alzheimer's disease based on hierarchical partner matching independent component analysis. *Frontiers in aging neuroscience*, 10:417, 2018. [29](#)
- [QWC⁺19] Xiaoxia Qu, Qian Wang, Weiwei Chen, Ting Li, Jian Guo, Huaizhou Wang, Xun Zhang, Ying Wang, Ningli Wang, and Junfang Xian. Combined machine learning and diffusion tensor imaging reveals altered anatomic fiber connectivity of the brain in primary open-angle glaucoma. *Brain research*, 1718:83–90, 2019. [29](#)
- [RCDS⁺13] Françoise Radat, Sandra Chanraud, Georges Di Scala, Virginie Dousset, and Michèle Allard. Psychological and neuropsychological correlates of dependence-related behaviour in medication overuse headaches: a one year follow-up study. *The journal of headache and pain*, 14(1):59, 2013. [28](#)
- [RMM⁺18] Erika Rovini, Carlo Maremmani, Alessandra Moschetti, Dario Esposito, and Filippo Cavallo. Comparative motor pre-clinical assessment in parkinson's disease using supervised machine learning approaches. *Annals of biomedical engineering*, 46(12):2057–2068, 2018. [29](#)
- [Rob14] Ian H Robertson. A right hemisphere role in cognitive reserve. *Neurobiology of aging*, 35(6):1375–1385, 2014. [28](#)
- [RSR⁺19] Haidi Rao, Xianzhang Shi, Ahoussou Kouassi Rodrigue, Juanjuan Feng, Yingchun Xia, Mohamed Elhoseny, Xiaohui Yuan, and Lichuan Gu. Feature selection based on artificial bee colony and gradient boosting decision tree. *Applied Soft Computing*, 74:634–642, 2019. [15](#)

BIBLIOGRAFÍA

- [SBH⁺14] Jason Steffener, Daniel Barulli, Christian Habeck, Deirdre O'Shea, Qolamreza Razlighi, and Yaakov Stern. The role of education and verbal abilities in altering the effect of age-related gray matter differences on cognition. *PLoS One*, 9(3):e91196, 2014. [28](#)
- [SBJ⁺15] Timothy J Steiner, Gretchen L Birbeck, Rigmor H Jensen, Zaza Katsarava, Lars J Stovner, and Paolo Martelletti. Headache disorders are third cause of disability worldwide, 2015. [27](#)
- [SFJI20] MNAH Sha'abani, N Fuad, Norezmi Jamal, and MF Ismail. knn and svm classification for eeg: A review. In *InECCE2019*, pages 555–565. Springer, 2020. [21](#)
- [SH11] Margaret J Snowling and Charles Hulme. Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology*, 81(1):1–23, 2011. [27](#)
- [SHS⁺17] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dhharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv:1712.01815*, 2017. [21](#)
- [SSP⁺98] Sally E Shaywitz, Bennett A Shaywitz, Kenneth R Pugh, Robert K Fulbright, R Todd Constable, W Einar Mencl, Donald P Shankweiler, Alvin M Liberman, Pawel Skudlarski, Jack M Fletcher, et al. Functional disruption in the organization of the brain for reading in dyslexia. *Proceedings of the National Academy of Sciences*, 95(5):2636–2641, 1998. [26](#)
- [SYL⁺05] Sheng-Kwei Song, Jun Yoshino, Tuan Q Le, Shiow-Jiuan Lin, Shu-Wei Sun, Anne H Cross, and Regina C Armstrong. Demyelination increases radial diffusivity in corpus callosum of mouse brain. *Neuroimage*, 26(1):132–140, 2005. [26](#)

- [TBF⁺15] Gretchen E Tietjen, Dawn C Buse, Kristina M Fanning, Daniel Serrano, Michael L Reed, and Richard B Lipton. Recalled maltreatment, migraine, and tension-type headache: results of the ampp study. *Neurology*, 84(2):132–140, 2015. [28](#)
- [Ten] TensorFlow. Tensorflow. [24](#)
- [TMA14] Sina Tabakhi, Parham Moradi, and Fardin Akhlaghian. An unsupervised feature selection algorithm based on ant colony optimization. *Engineering Applications of Artificial Intelligence*, 32:112–123, 2014. [18](#)
- [TMK⁺17] Toshiaki Taoka, Yoshitaka Masutani, Hisashi Kawai, Toshiki Nakane, Kiwamu Matsuoka, Fumihiko Yasuno, Toshifumi Kishimoto, and Shinji Naganawa. Evaluation of glymphatic system activity with the diffusion mr technique: diffusion tensor image analysis along the perivascular space (dti-als) in alzheimer’s disease cases. *Japanese journal of radiology*, 35(4):172–178, 2017. [12](#)
- [TTT⁺13] Claudio Tana, Emmanuele Tafuri, Marco Tana, Paolo Martelletti, Andrea Negro, Giannapia Affaitati, Alessandra Fabrizio, Raffaele Costantini, Andrea Mezzetti, and Maria Adele Giamberardino. New insights into the cardiovascular risk of migraine and the role of white matter hyperintensities: is gold all that glitters? *The journal of headache and pain*, 14(1):1–9, 2013. [28](#)
- [TYO⁺17] Fengbin Tu, Shouyi Yin, Peng Ouyang, Shibin Tang, Leibo Liu, and Shaojun Wei. Deep convolutional neural network architecture with reconfigurable computation patterns. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 25(8):2220–2233, 2017. [25](#)
- [VP10] Trichur R Vidyasagar and Kristen Pammer. Dyslexia: a deficit in visuo-spatial attention, not in phonological processing. *Trends in cognitive sciences*, 14(2):57–63, 2010. [27](#)

BIBLIOGRAFÍA

- [Wec03] David Wechsler. Wechsler intelligence scale for children—fourth edition (wisc-iv). *San Antonio, TX: The Psychological Corporation*, 2003. [33](#)
- [Wec05] David Wechsler. *Escala Wechsler de inteligencia para niños-IV*. Manual moderno, 2005. [40](#)
- [Weh12] Louis A Wehenkel. *Automatic learning techniques in power systems*. Springer Science & Business Media, 2012. [13](#)
- [WJWS19] Feng Wang, Dingde Jiang, Hong Wen, and Houbing Song. Adaboost-based security level classification of mobile intelligent terminals. *The Journal of Supercomputing*, 75(11):7460–7478, 2019. [22](#)
- [WPS⁺18] Shui-Hua Wang, Preetha Phillips, Yuxiu Sui, Bin Liu, Ming Yang, and Hong Cheng. Classification of alzheimer’s disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. *Journal of medical systems*, 42(5):85, 2018. [25](#), [29](#)
- [WZZ⁺16] Xiaodong Wang, Xu Zhang, Zhiqiang Zeng, Qun Wu, and Jian Zhang. Unsupervised spectral feature selection with l1-norm graph. *Neurocomputing*, 200:47–54, 2016. [20](#)
- [YGW⁺20] Zezhong Ye, Ajit George, Anthony T Wu, Xuan Niu, Joshua Lin, Gautam Adusumilli, Robert T Naismith, Anne H Cross, Peng Sun, and Sheng-Kwei Song. Deep learning with diffusion basis spectrum imaging for classification of multiple sclerosis lesions. *Annals of Clinical and Translational Neurology*, 7(5):695–706, 2020. [29](#)
- [YYH⁺09] Qing Yan, Hongmei Yan, Fei Han, Xinchuan Wei, and Tao Zhu. Svm-based decision support system for clinic aided tracheal intubation predication with multiple features. *Expert Systems with Applications*, 36(3):6588–6592, 2009. [90](#)

- [YZW19] Jianbo Yu, Xiaoyun Zheng, and Shijin Wang. A deep autoencoder feature learning method for process pattern recognition. *Journal of Process Control*, 79:1–15, 2019. [24](#)
- [ZBF⁺12] Marco Zorzi, Chiara Barbiero, Andrea Facchetti, Isabella Lonciari, Marco Carrozzi, Marcella Montico, Laura Bravar, Florence George, Catherine Pech-Georgel, and Johannes C Ziegler. Extra-large letter spacing improves reading in dyslexia. *Proceedings of the National Academy of Sciences*, 109(28):11455–11459, 2012. [27](#)
- [Zho20] Ding-Xuan Zhou. Universality of deep convolutional neural networks. *Applied and computational harmonic analysis*, 48(2):787–794, 2020. [24](#)
- [Zhu05] Xiaojin Jerry Zhu. Semi-supervised learning literature survey. Technical report, University of Wisconsin-Madison Department of Computer Sciences, 2005. [21](#)
- [ZL18] Ying-Teng Zhang and Shen-Quan Liu. Individual identification using multi-metric of dti in alzheimer’s disease and mild cognitive impairment. *Chinese Physics B*, 27(8):088702, 2018. [29](#)
- [ZP15] Zisheng Zhang and Keshab K Parhi. Seizure detection using regression tree based feature selection and polynomial svm classification. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6578–6581. IEEE, 2015. [19](#)
- [ZYST19] Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. Deep learning based recommender system: A survey and new perspectives. *ACM Computing Surveys (CSUR)*, 52(1):1–38, 2019. [23](#)
- [ZZH08] Jiong Zhang, Mohammad Zulkernine, and Anwar Haque. Random-forests-based network intrusion detection systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 38(5):649–659, 2008. [16](#)

Esta tesis se terminó de escribir en Bilbao el viernes, 3 de julio de 2020

