



TESIS DOCTORAL

**DIAGNOSTICO DEL GLAUCOMA EN ESTADIO
PRECOZ MEDIANTE EL ESTUDIO DE LA CAPA DE
FIBRAS RETINIANAS PERIPAPILARES: UNA
APLICACIÓN DE MINERIA DE DATOS EN MEDICINA**

Autor:

CARLOS SALVADOR FERNANDEZ ESCAMEZ

Director:

NICOLAS TOLEDANO FERNÁNDEZ

Programa de Doctorado en Ciencias de la Salud
Escuela Internacional de Doctorado

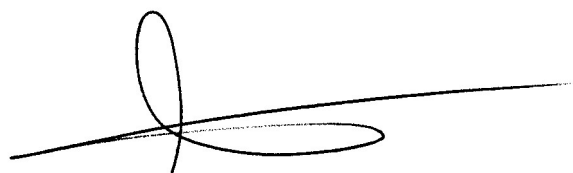
MADRID, MARZO 2021

D. NICOLAS TOLEDANO FERNANDEZ, Profesor Asociado de Oftalmología de la Universidad Rey Juan Carlos de Madrid.

CERTIFICA

Que **D. CARLOS SALVADOR FERNANDEZ ESCAMEZ**, Licenciado en Medicina y Cirugía, ha realizado bajo su supervisión el trabajo “Diagnóstico del glaucoma en estadio precoz mediante el estudio de la capa de fibras retinianas peripapilares: una aplicación de minería de datos en medicina”, y que dicho trabajo reúne las condiciones científicas y académicas necesarias para optar al Grado de Doctor por la Universidad Rey Juan Carlos de Madrid.

En Madrid, a 12 de Febrero de 2021

A handwritten signature in black ink, consisting of a large loop on the left and a long horizontal stroke extending to the right.

Fdo: D. Nicolás Toledano Fernández

AGRADECIMIENTOS

A mis padres por ser el faro que ha dado luz a mi vida y proporcionarme siempre la dirección para llegar a buen puerto.

A mis hermanos, por ser mis acompañantes más abnegados e infatigables en el camino de la vida.

A Cristina, por enseñarme a descubrir y sentir más allá de la ciencia, su gran apoyo y por todo el tiempo robado para escribir esta tesis.

A Nicolás Toledano, por su ayuda continua y confiar en que esta tesis algún día vería la luz.

A mis compañeros de trabajo, por ser excelentes compañeros y amigos, y hacer que el trabajo sea más placer que esfuerzo.

INDICE GENERAL

I) INTRODUCCION

1.EL GLAUCOMA CRONICO DE ANGULO ABIERTO

1.1 Definición de glaucoma.....	1
1.2 Epidemiología del glaucoma.....	2
1.3 Diagnostico del glaucoma.....	4
1.4 Calidad de vida en pacientes con glaucoma.....	5
1.5 El glaucoma en estadio precoz.....	6

2.TOMOGRAFIA DE COHERENCIA OPTICA

2.1 Características generales.....	8
2.2 Estudio de la capa de fibras peripilares mediante OCT.....	11
2.2.1 Importancia del estudio de la CFNR para el diagnóstico precoz del glaucoma.....	11
2.2.2 Representación de resultados del estudio de la capa de fibras nerviosas con OCT	12
2.2.3 Artefactos en las imágenes con OCT.....	13
2.2.4 OCT en el diagnóstico del glaucoma.....	15
2.2.5 Influencia del tamaño discal en la medida de la CFNR.....	18
2.3 Estudio de la cabeza del nervio óptico con OCT.....	19

3. MINERIA DE DATOS

3.1 Introducción a la minería de datos.....	20
3.2 ¿Qué es la minería de datos?.	23
3.3 Minería de datos frente a estadística clásica.....	24
3.4 Descubrimiento del conocimiento en bases de datos.....	25
3.4.1 Fases del descubrimiento de conocimiento en bases de datos.....	27
3.5 Nomenclatura básica utilizada en la minería de datos.....	29
3.6 Clasificación de las técnicas de minería de datos.....	30
3.6.1 Técnicas supervisadas vs técnicas no supervisadas.....	30
3.6.2 Modelos según su grado de complejidad.....	33
3.6.2.1 Árboles de decisión o clasificación.....	36
3.6.2.2 Combinaciones de árboles de decisión.....	38
3.6.2.3 Redes neuronales artificiales y modelos de aprendizaje profundo.....	41

3.7 Interpretabilidad de los modelos de clasificación.....	45
3.8 Minería de datos en oftalmología.....	49
II) JUSTIFICACION DEL ESTUDIO.....	61
III) OBJETIVOS.....	65
IV) MATERIAL Y METODOS.	
1. CARACTERISTICAS GENERALES DEL ESTUDIO.....	69
2. PROTOCOLO DE EXPLORACION OFTALMOLOGICA.....	69
3. SELECCIÓN DE PACIENTES	
3.1 Criterios de inclusión	
3.1.1 Grupo de pacientes con glaucoma en estadio inicial.....	71
3.1.2 Grupo control de sujetos sanos.....	72
3.2 Criterios de exclusión	
3.2.1 Grupo de pacientes con glaucoma en estadio inicial.....	73
3.2.2 Grupo control de sujetos sanos.....	73
4. VARIABLES INCLUIDAS EN EL ESTUDIO. PARAMETROS SELECCIONADOS PARA LA CREACION DEL MODELO	
4.1 Variables epidemiológicas.....	74
4.2 Variables obtenidas a partir de parámetros de OCT.....	74
4.2.1 Identificación de cuadrantes y sectores horarios.....	75
4.2.2 Correspondencia entre ojo derecho y ojo izquierdo.....	76
4.2.3 Control de la calidad de la imagen de OCT.....	76
4.3 Variables obtenidas a partir de parámetros del campo visual.....	77
5. SOFTWARE PARA ANALISIS DE DATOS.....	77
6. ANALISIS ESTADISTICO GENERAL DE LA BASE DE DATOS	
6.1 Análisis estadístico de variables epidemiológicas y de campo visual.....	78
6.2 Análisis estadístico del grosor de la CFNR entre los grupos.....	79
6.3 Comparación del tamaño del disco óptico entre grupos.....	79
7. METODOLOGIA DE MINERIA DE DATOS EMPLEADA PARA EL ANALISIS DE LA BASE DE DATOS.....	79
7.1 Determinación del número óptimo de agrupaciones.....	83
7.2 Detección y eliminación de los casos extremos.....	83

7.3 Creación del modelo de clasificación mediante el algoritmo de potenciación del gradiente basado en árboles.....	84
7.4 Estudio de la importancia de las variables del modelo en el clasificador XGB.....	85
7.5 Cálculo e interpretación de los valores Shapley del modelo XGB.....	85
7.5.1 Análisis del impacto de las variables de la CFNR en la predicción del modelo XGB.....	85
7.5.2 Análisis de la relación entre grosor de la CFNR e impacto en el resultado del clasificador: análisis de dependencia.....	86
7.6 Selección y categorización de los parámetros de mayor impacto en el clasificador	86
7.7 Creación de un clasificador de alta interpretabilidad basado en árboles de decisión.....	87

V) RESULTADOS

1. ANALISIS DESCRIPTIVO DE LA BASE DE DATOS

1.1 Variables poblacionales.....	91
1.2 Variables oculares	
1.2.1 Proporción de ojo derecho/izquierdo.....	92
1.2.2 Valores de grosor de la CFNR peripapilar.....	92
1.2.3 Tamaño del disco óptico.....	95
1.2.4 Variables del campo visual.....	97

2. NUMERO OPTIMO DE AGRUPACIONES. EXCLUSION DE VALORES EXTREMOS DE LA BASE DE DATOS.....

3. MODELO DE CLASIFICACION DE POTENCIACION DEL GRADIENTE BASADO EN ARBOLES DE DECISION.....

3.1 Importancia relativa de las variables del modelo XGB.....	104
---	-----

4. EVALUACION DE LA IMPORTANCIA DE LAS VARIABLES MEDIANTE LOS VALORES SHAPLEY.....

5. ANALISIS DE DEPENDENCIA ENTRE VALOR DEL PARAMETRO DE ENTRADA E IMPACTO EN LA PREDICION.....

5.1 Análisis del grosor de cuadrante inferior.....	108
5.2 Análisis del grosor del sector horario H1.....	110
5.3 Análisis del grosor del sector horario H6.....	111
5.4 Análisis del grosor del sector horario H2.....	112
5.6 Análisis del grosor del cuadrante superior... ..	113

5.5 Análisis del grosor promedio de la CFNR.....	115
6. CONSTRUCCION DEL CLASIFICADOR DE ALTA INTERPRETABILIDAD BASADO EN ARBOLES DE DECISION.....	116
VI) DISCUSION	
1. CONSIDERACIONES GENERALES Y METODOLOGICAS DEL ESTUDIO.....	123
2. CARACTERISTICAS BASALES DE LA BASE DE DATOS.....	127
3. FUNDAMENTOS DE NUESTRO MODELO DE CLASIFICACION.....	129
4. ANALISIS DE AGRUPACIONES DE LOS DATOS DE ENTRADA.....	130
5. ANALISIS DEL MODELO DE CLASIFICACION XGB.....	131
6. ANALISIS DEL MODELO XGB MEDIANTE VALORES SHAPLEY.....	134
7. ANALISIS DE DEPENDENCIA DE LOS PARAMETROS RESPECTO AL VALOR SHAPLEY.....	134
8. CLASIFICADOR DE ALTA INTERPRETABILIDAD: ARBOL DE DECISION.....	137
VII) CONCLUSIONES.....	143
VIII) BIBLIOGRAFIA.....	147
IX) ANEXOS	167

ABREVIATURAS

AD	Análisis discriminante
ALD	Análisis lineal discriminante
ANS	Análisis no supervisado
AP	Aprendizaje profundo
AROC	Area bajo la curva ROC
AUC	Area bajo la curva
CFNR	Capa de fibras nerviosas de la retina
CI	Cuadrante inferior
CN	Cuadrante nasal
CS	Cuadrante superior
CT	Cuadrante temporal
dB	Decibelio
DCBD	Descubrimiento de conocimiento en bases de datos
DE	Desviación estándar
DM	Desviación media
DMAE	Degeneración macular asociada a la edad
EGS	Sociedad Europea de Glaucoma
GPAA	Glaucoma primario de ángulo abierto
Hi	Sector horario "i"
HRT	Tomografía retiniana de Heidelberg
IC	Intervalo de confianza
KM	k-medias
K-S	Kolmogorov-Smirnov
MVS	Máquinas de vectores soporte
OCT	Tomografía de coherencia óptica
PIO	Presión intraocular
PSD	Desviación estándar del patrón
RD	Retinopatía diabética
RF	Random-Forest
RNA	Red neuronal artificial
ROC	Curva Operativa del Receptor (" Receiver Operating Characteristic")
SD-OCT	Tomografía de coherencia óptica de dominio espectral

SHAP	Explicaciones aditivas Shapley ("Shapley additive explanations")
SSI	Índice de intensidad de señal
TD-OCT	Tomografía de coherencia óptica de dominio temporal
XGB	Potenciación extrema del gradiente ("Extreme Gradient Boosting")

INDICE DE TABLAS

Tabla 1: Prevalencia del glaucoma en los distintos estudios.

Tabla 2: Exploraciones necesarias para el diagnóstico del glaucoma y nivel de calidad.

Tabla 3: Guía del diagnóstico del glaucoma.

Tabla 4: Diferencias entre aprendizaje supervisado y no supervisado.

Tabla 5: Comparación de métodos de clasificación según distintos criterios.

Tabla 6: Resumen de estudios de aplicación de aprendizaje automático en el diagnóstico del glaucoma usando datos o imágenes de la OCT

Tabla 7: Estadísticos de la variable edad.

Tabla 8: Estadísticos de la variable sexo.

Tabla 9: Estadísticos de la variable ojo derecho / ojo izquierdo.

Tabla 10: Valores de AUC para cada parámetro de la CFNR y intervalo de confianza

Tabla 11: Valores de corte del grosor de CFNR con especificidad superior a 0.7 y nivel de sensibilidad.

Tabla 12: Parámetros de CFNR ordenados según AROC en orden descendente.

Tabla 13: Estadísticos descriptivos de la variable tamaño del nervio óptico.

Tabla 14: Estadísticos descriptivos de la variable Desviación Media del campo visual..

Tabla 15: Estadísticos descriptivos de la variable Desviación Estándar del Patrón.

Tabla 16: Estadísticos descriptivos de la variable Índice de Campo Visual.

Tabla 17: Frecuencias de la variable Hemicampo afectado en el grupo de glaucoma precoz.

Tabla 18: Valor medio y desviación estándar de las distancias al centroide.

Tabla 19: Número de casos en los datos de entrenamiento y datos de validación y proporción de glaucoma vs sanos.

Tabla 20: Medidas de precisión, AROC, sensibilidad y especificidad del modelo XGB.

Tabla 21: Valores de importancia relativa de las variables de entrada al modelo XGB.

Tabla 22: Categorías para la conversión de las variables a cualitativas ordinales

Tabla 23: Medidas de precisión, AUC y valor de especificidad y sensibilidad del modelo árbol de clasificación.

INDICE DE FIGURAS

Figura 1: Representación gráfica de la progresión del glaucoma

Figura 2: Segmentación de las capas de la retina en una imagen de OCT

Figura 3: Informe de estudio de la CFNR mediante OCT

Figura 4: Variación de la medida grosor de la CFNR peripapilar para distintos valores del índice de intensidad de la señal.

Figura 5: Representación de la pirámide del conocimiento

Figura 6: Niveles del conocimiento en función de su grado de abstracción.

Figura 7: Resumen de las fases que componen el proceso de descubrimiento de conocimiento en bases de datos.

Figura 8: Ejemplo de algoritmo no supervisado para realizar la partición de los datos en agrupamientos homogéneos.

Figura 9: Clasificación de las técnicas de minería de datos según su metodología.

Figura 10: Correlación entre complejidad de un modelo y su precisión.

Figura 11: Ejemplo de estructura de árbol de decisión binario.

Figura 12: Representación del funcionamiento de la técnica de Bagging

Figura 13: Relación entre complejidad e interpretabilidad del modelo.

Figura 14: Diagrama de una neurona artificial

Figura 15. Esquema de una red neuronal artificial con cinco capas de neuronas.

Figura 16: Esquema de una red convolucional.

Figura 17: Esquema de la relación entre datos, modelo y valores Shapley.

Figura 18. Representación de la contribución de cada parámetro de la CFNR en la predicción del modelo en función de su valor Shapley.

Figura 19: Ubicación de los sectores horarios en el ojo derecho.

Figura 20: Correspondencia de los sectores horarios de la CFNR entre ojo derecho e izquierdo

Figura 21: Flujo de trabajo hasta la producción de la base de datos etiquetada depurada.

Figura 22: Flujo de análisis de datos desde la base de datos etiquetada tras eliminar casos extremos.

Figura 23. Curva ROC para cada parámetro individual de la CFNR.

Figura 24: Relación entre porcentaje de varianza total explicada y número de particiones.

Figura 25: Representación gráfica bidimensional de los agrupamientos generados por el algoritmo k-medias.

Figura 26: Representación de los casos en la partición [1] según su distancia a centroide

Figura 27: Representación de los casos en la partición [2] según su distancia a centroide.

Figura 28: Gráfico de la importancia relativa de los parámetros de la CFNR ordenados en orden decreciente

Figura 29: Ordenación de los parámetros de la CFNR según la media del valor Shapley sin signo.

Figura 30: Representación de los valores Shapley según el valor numérico del grosor de la CFNR para cada parámetro.

Figura 31: Gráfico de análisis de dependencia del grosor de CFNR en cuadrante inferior.

Figura 32: Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el gráfico de dependencia del cuadrante inferior.

Figura 33: Gráfico de análisis de dependencia del grosor de CFNR en sector horario H1

Figura 34: Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el cuadrante inferior

Figura 35: Gráfico de análisis de dependencia del grosor de CFNR en sector horario H6

Figura 36. Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el gráfico de dependencia del sector horario H6

Figura 37: Gráfico de análisis de dependencia del grosor de CFNR en el sector horario H2.

Figura 38: Gráfico de análisis de dependencia del grosor de CFNR en el cuadrante superior.

Figura 39: Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el gráfico de dependencia del cuadrante superior.

Figura 40: Gráfico de análisis de dependencia del grosor promedio de CFNR peripapilar

Figura 41: Agrupamientos creados mediante algoritmo no-supervisado para el valor promedio de la CFNR peripapilar

Figura 42: Arbol de clasificación construido con las variables categorizadas

INTRODUCCION

1. EL GLAUCOMA CRONICO SIMPLE.

1.1. DEFINICION DE GLAUCOMA.

El término glaucoma hace referencia a un grupo de enfermedades que cursan con atrofia lenta y progresiva del nervio óptico como consecuencia de la muerte de las células ganglionares de la retina. Se caracteriza por una pérdida de la función visual periférica y un aspecto excavado de la papila óptica. El glaucoma se asocia frecuentemente, aunque no en todos los casos, a una presión intraocular elevada. Según el valor de la presión ocular, se pueden distinguir entre glaucomas hipertensivos y glaucomas normotensivos o de baja tensión.

La Sociedad Europea de Glaucoma (European Glaucoma Society – EGS) define el glaucoma primario de ángulo abierto (GPAA) como el conjunto de neuropatías de carácter crónico y progresivo que tienen en común una serie de cambios característicos en la morfología de la cabeza del nervio óptico y de la capa de fibras nerviosas, en ausencia de otras anomalías congénitas o enfermedades oculares que los justifique, y en presencia de un ángulo abierto en la exploración gonioscópica.(1)

Por su parte, la Academia Americana de Oftalmología describe al GPAA como una neuropatía óptica crónica y progresiva en sujetos adultos, en la que la PIO y otros factores todavía desconocidos contribuyen al daño del nervio óptico y en la que, en ausencia de otras causas identificables, hay una atrofia adquirida del disco óptico como consecuencia de la muerte de células ganglionares y la pérdida de sus axones.(2)

La razón fundamental de la gran variabilidad en la definición del glaucoma deriva del hecho de que su etiología y los mecanismos causantes son objeto de debate hasta el momento actual. Clásicamente, se han descrito dos teorías etiopatogénicas: la teoría mecánica y la teoría vascular. En la teoría de origen mecánico, el daño glaucomatoso es secundario a la lesión de los axones de las células ganglionares por simple estrangulamiento debido a la presión intraocular. La teoría vascular establece que procesos isquémicos en la cabeza del nervio óptico son los causantes del daño glaucomatoso. Otros factores genéticos, mutaciones de diferentes genes (ej. MYOC,

OPTN, WD36) y el estrés oxidativo podrían participar también en la génesis de la enfermedad.(3)

1.2. EPIDEMIOLOGIA DEL GLAUCOMA.

El glaucoma tiene una gran relevancia en la salud pública, constituyendo la segunda causa de ceguera tanto a nivel mundial como en Europa Occidental.(4–10).

Se estima que en el año 2020 hubo en el mundo unos 64,3 millones de personas que padecieron glaucoma, cifra que se calcula que se elevará hasta 111.8 millones de enfermos hacia el año 2040. De este computo total, 52 millones corresponden a pacientes con GPAA en el año 2020, proyectándose hasta 79.76 millones de enfermos de GPAA en el año 2040 (11).

Una revisión sistemática y meta-análisis de las publicaciones sobre prevalencia del glaucoma ha estimado que, en el momento actual, la prevalencia global del glaucoma en la población entre 40-80 años es de 3.54 % (95% IC: 2.09 - 5.82) (11).

En nuestro entorno, se estima que el GPAA afecta al 2% de adultos mayores de 40 años de edad (10,12,13). En España, estudios epidemiológicos han estimado una prevalencia de glaucoma del 2.1% en sujetos entre 40 y 79 años de edad. (14)

En general, la prevalencia del glaucoma se va incrementando con la edad, de modo que a partir de los 70 años se calcula que un 4-6 % de la población sufrirá la enfermedad (15). El envejecimiento progresivo de la población conllevará, por ello, a que el glaucoma suponga un problema sociosanitario de aún mayores proporciones en las próximas décadas.

El riesgo y subtipo de glaucoma varía entre razas y regiones geográficas. El GPAA es la forma predominante de glaucoma a nivel mundial, aunque determinadas razas como la raza mongol y birmana muestran mayor afectación por glaucoma primario de ángulo cerrado que de glaucoma de ángulo abierto.(16,17)

La prevalencia de GPAA encontrada en los distintos estudios está vinculada a los criterios diagnósticos utilizados en el mismo. No obstante, hay una alta consistencia en

las estimaciones de prevalencia obtenidas en estudios llevados a cabo en poblaciones tanto asiática como caucásica (tabla 1).

Tabla 1
Prevalencia del glaucoma en los distintos estudios

Tipo de glaucoma	Localización	Año	Edad	Prevalencia (%)
GCAA y otros tipos de glaucoma	Rotterdam, Netherlands	1990-1999	65 – 74	3.5
GCAA	Líbano	2007	> 40	7.30
GCAA y otros tipos de glaucoma	Francia	2003	> 18	3.40
GCAA precoz	Austria	2006	Adultos	2.9
GCAA	España	2004	40-79	2.1
GCAA	Italia	1997	>40	2.51
GCAA	Austria	2007	>60	6.9
GCAA	Thessaloniki, Grecia	1999	>60	3.8

Un rasgo común, en muchos de los estudios publicados, es el alto porcentaje de sujetos que no habían sido diagnosticados previamente. La prevalencia de glaucoma no diagnosticado en una población de Estados Unidos ha sido estimada en un 2.9 %, incrementándose con la edad hasta el 6.6% en el grupo de individuos de más de 70 años.(18)

En los países industrializados, cerca del 50% de los pacientes con GPAA desconocen que padecen la enfermedad (19–22). Este porcentaje se eleva hasta el 75-90% en países subdesarrollados o en subgrupos minoritarios (23,24).

1.3. DIAGNOSTICO DEL GLAUCOMA

El diagnóstico de glaucoma está basado en la evaluación clínica de las estructuras oculares y las pruebas funcionales afectadas por la enfermedad: el nervio óptico, la capa de fibras nerviosas de la retina y el campo visual. No obstante, no existe una definición universalmente aceptada de qué criterios diagnósticos deben emplearse para definir un glaucoma (25).

La Sociedad Finlandesa de Oftalmología ha publicado una “Guía de prácticas clínicas en el paciente glaucomatoso” en la que se establecen unos niveles de calidad diagnóstica, estableciendo las pruebas que deben llevarse a cabo en cada uno de los niveles diagnósticos (26). Se establecen cuatro niveles de calidad, incluyendo la toma de la presión intraocular en todos ellos. En el “nivel bueno” de diagnóstico, deberán incluirse, además de la medida de la PIO, la realización de gonioscopia, campo visual y la exploración del disco óptico y/o capa de fibras retinianas peripapilares (tabla 2).

Tabla 2

Exploraciones necesarias para el diagnóstico del glaucoma y nivel de calidad.

Nivel	Exploraciones				
Muy Bueno	PIO	+ Gonioscopia	+ Campo visual	+ Imagen de Nervio Optico	+ Imagen de CFNR
Bueno	PIO	+ Gonioscopia	+ Campo visual	+ Imagen de Nervio Optico ó CFNR	
Suficiente	PIO	+ Gonioscopia	+ Campo visual		
Insuficiente	PIO				

La comprobación de la agudeza visual y el fondo de ojo son exploraciones básicas. Al determinar el resto de exploraciones oculares se debe tener en cuenta la edad del paciente, la severidad del glaucoma y otras posibles enfermedades oculares.

Fuente: Tuulonen A. The Finnish evidence-based guideline for open-angle glaucoma (26).

Dentro de estas recomendaciones de la Sociedad Finlandesa, se estipula la realización de una curva de PIO diurna cuando sea necesaria, así como el calibrado regular del tonómetro. Acerca del campo visual, indica que la opción preferible es la práctica de

dos exploraciones de campo visual mediante una estrategia umbral para la determinación de estado basal del campo visual. La perimetría azul-amarillo, el estudio de los 10 grados centrales del campo visual y el análisis cuantitativo de la cabeza del nervio óptico podrían proporcionar información adicional de gran utilidad en ciertos casos.

Las mayores dificultades diagnósticas se presentan en las fases iniciales de la enfermedad ya que ninguna de las exploraciones que se realiza posee la sensibilidad y especificidad suficiente para confirmar de forma inequívoca la presencia de glaucoma.

El diagnóstico del GPAA debe realizarse, por tanto, considerando todos los datos de la exploración y la historia clínica, pero debe basarse de modo fundamental en los siguientes puntos:

- Lesión anatómica del nervio óptico.
- Lesión funcional del campo visual.
- Aumento de la presión intraocular.

1.4. CALIDAD DE VIDA EN PACIENTES CON GLAUCOMA.

La Sociedad Europea de Glaucoma establece que los objetivos del tratamiento en el glaucoma deben ser (1):

- Calidad de vida.
- Calidad visual.
- Control de los costes.

La finalidad última del tratamiento es la preservación de la función visual del paciente y de su calidad de vida relacionada con la visión, con un coste aceptable, entendiendo por coste tanto los efectos secundarios derivados del tratamiento como los desembolsos económicos que conlleva su adquisición.

La calidad de vida está estrechamente ligada a la función visual. Los pacientes con glaucoma precoz o moderado por daño glaucomatoso tienen una buena función visual pero una reducción moderada de su calidad de vida.

La principal expresión clínica del glaucoma es la pérdida visual como una reducción progresiva de la amplitud del campo visual. En los casos más avanzados, esta pérdida campimétrica provoca dificultades en la vida personal y relacional del sujeto afecto de glaucoma. Gran cantidad de estudios han demostrado la existencia de limitaciones en la realización de las actividades cotidianas en los pacientes con glaucoma (27–30).

El grado de discapacidad secundaria a la patología glaucomatosa en un paciente es difícil de cuantificar (31). Se han utilizado encuestas de autovaloración, con la subjetividad que éstas conllevan, así como tests que miden la capacidad del individuo para la ejecución de distintas tareas en condiciones estandarizadas. En la medición de esta discapacidad también pueden influir factores no oculares, como son la edad, enfermedades concomitantes, nivel educativo, capacidad cognitiva o la presencia de depresión (32), por lo que la validez de estos tests es cuestionable.

Estudios basados en encuestas indican que las principales tareas afectadas en el glaucoma son la dificultad para subir o bajar escaleras (33) y el choque involuntario con objetos al andar (34). También se ha observado que pacientes con afectación campimétrica muestran peor equilibrio (35), mayor frecuencia de caídas y más dificultad en la conducción de vehículos (36).

La conducción de vehículos motorizados es una de las actividades más afectadas en pacientes afectados de glaucoma, especialmente en los casos de enfermedad bilateral. Se ha confirmado una mayor incidencia de colisiones en pacientes con glaucoma que practican la conducción de vehículos (36).

De manera resumida, es indiscutible la existencia de un cierto grado de discapacidad en pacientes con glaucoma secundario a la pérdida funcional visual que provoca la enfermedad, aunque es complicado determinar su magnitud.

1.5 EL GLAUCOMA EN ESTADIO PRECOZ.

El glaucoma es una enfermedad neurodegenerativa del nervio óptico de carácter adquirido y progresivo. Su presentación clínica puede darse en distintos estadios dentro de un espectro continuo que conduce desde la normalidad hasta la ceguera. Esta

progresión se deriva de la secuencia de acontecimientos cuyo comienzo es la lesión y muerte de células ganglionares de la retina. La degeneración y pérdida de los axones de estas células ganglionares se manifiesta a nivel macroscópico con un aumento de la excavación del nervio óptico. Desde el punto de vista funcional, la destrucción de las células ganglionares se revela con la aparición de escotomas en el campo visual del paciente (37). En la figura 1 se representa este abanico de posibles situaciones en las que puede situarse un paciente con glaucoma.

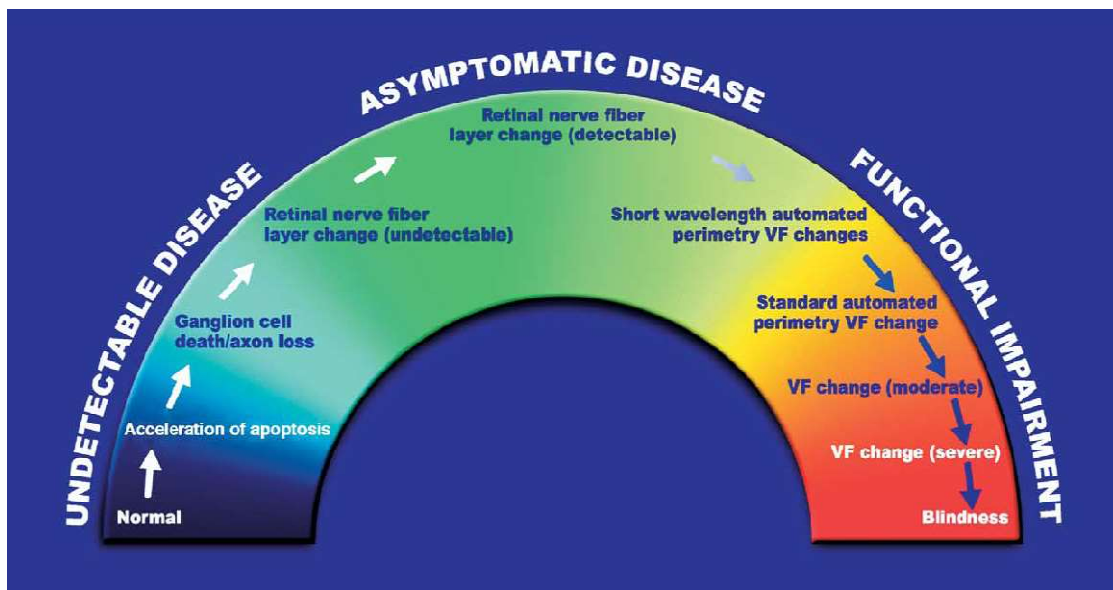


Figura 1: Representación gráfica de la progresión del glaucoma desde los estadios iniciales hasta la ceguera, especificando los cambios microscópicos y funcionales que acontecen en cada fase.

Fuente: Weinreb RN, Khaw PT. Primary open-angle glaucoma (37)

Los cambios iniciales en la retina y el nervio óptico en sujetos con glaucoma son sutiles y con frecuencia indetectables con los test diagnósticos de los que disponemos en la actualidad. Asimismo, no hay un acuerdo global de los criterios para establecer el diagnóstico del daño precoz que antecede a los cambios en la perimetría acromática estándar.

Un número considerable de los pacientes se presentan en estadios incipientes de la enfermedad. La dificultad de diagnóstico en esta fase se deriva de la superposición de hallazgos en el nervio óptico y en la capa de fibras nerviosas que ocurre entre los sujetos normales y los pacientes con signos precoces de lesión glaucomatosa.

La demora diagnóstica esperando la aparición de signos evidentes de la enfermedad involucra aceptar un cierto grado de daño irreversible y una probable progresión. El objetivo primario del tratamiento es evitar, detener o retrasar el avance a lo largo de los estadios de la enfermedad, desde la sospecha de glaucoma hasta el daño precoz del nervio óptico e incluso la pérdida visual significativa.

El carácter progresivo que caracteriza al glaucoma da lugar a que, dentro de la enfermedad, podamos distinguir distintos cuadros clínicos en función de la alteración estructural o funcional detectable en un determinado momento. La Sociedad Finlandesa ha establecido una clasificación de los posibles diagnósticos dentro del concepto general de glaucoma, en función de los hallazgos anormales encontrados en la exploración (tabla 3) (26).

2. TOMOGRAFIA DE COHERENCIA OPTICA.

2.1. CARACTERISTICAS GENERALES

La tomografía de coherencia óptica (OCT) fue introducida en 1991 como un método de imagen no invasivo basado en el principio de la interferometría de baja coherencia para la obtención de imágenes in vivo de la retina (38). Su uso en oftalmología ha posibilitado la visualización de la estructura así como la medición del grosor de las distintas capas de la retina, convirtiéndose en la última década en una herramienta esencial para el estudio de las enfermedades que afectan a la retina, y en el caso del glaucoma, para la medición del grosor de la capa de fibras nerviosas. La OCT nos permite identificar y cuantificar el edema macular (39), evaluar los cambios en el grosor retiniano en respuesta a tratamiento (40), y medir el grosor de la capa de fibras nerviosas de la retina (41).

El fundamento físico de la OCT es semejante al de los instrumentos de ecografía convencional, utilizando ondas de luz en lugar de ondas acústicas. La OCT proyecta en la retina un haz luminoso emitido desde un láser diodo, y analiza las características de la luz reflejada para determinar las dimensiones de las estructuras retinianas.

Tabla 3

Guía del diagnóstico del glaucoma .

Hallazgos anormales	Hallazgos normales	Diagnostico	Comentarios
CFNR Nervio óptico Campo visual		Glaucoma	Diagnostico definitivo
CFNR Campo visual	Nervio óptico	Glaucoma	Posiblemente un disco óptico pequeño
CFNR Nervio óptico	Campo visual	Glaucoma preperimétrico	Perimetría azul-amarillo podría ser anormal
Nervio óptico Campo visual	CFNR	Otro diagnóstico distinto a glaucoma (por ej neurológico)*	
CFNR	Nervio óptico Campo visual	¿Glaucoma preperimétrico?	Seguimiento para buscar progresión.
Nervio óptico	CFNR Campo visual	Sospecha de glaucoma	Posible megalopapila
Campo visual	CFNR Nervio óptico	Sospecha de glaucoma	Repetir campo visual.

*En la literatura no hay descritos casos de este tipo de glaucoma. Al ordenar las exploraciones a realizar el oftalmólogo debe tener en consideración la edad del paciente, el grado de daño glaucomatoso y la existencia de otras posibles enfermedades oculares.

Fuente: Tuulonen A. The Finnish evidence-based guideline for open-angle glaucoma (26).

Los primeros modelos de OCT utilizaban tecnología de dominio tiempo (time-domain OCT o TD-OCT) para la adquisición de las imágenes (42). Estos OCT proyectan un haz de luz cercano al espectro de la luz infrarroja hacia la retina y hacia un espejo de referencia posicionado a una distancia conocida de cada capa de la retina. Cuando la luz reflejada de la retina se combina con la luz reflejada del espejo de referencia, se forma un patrón de interferencia que es interpretado por el aparato como una señal eco-A. El espejo de referencia se va moviendo a distintas distancias de la retina. El aparato compara los retardos de los ecos de la luz reflejada con los obtenidos al reflejar el mismo haz luminoso en el espejo de referencia situado a una distancia conocida.

Utilizando algoritmos internos, se procesan los datos obtenidos para generar una serie de imágenes que representan cortes anatómicos transversales de la retina.

Estas imágenes de los cortes axiales de la retina se codifican con una escala de colores en función de la reflectividad de la estructura representada. Los colores del espectro blanco y rojo corresponden a tejidos de mayor reflectividad, mientras que el color negro o azul representa escasa o nula resistencia a la transferencia de la luz (figura 2).

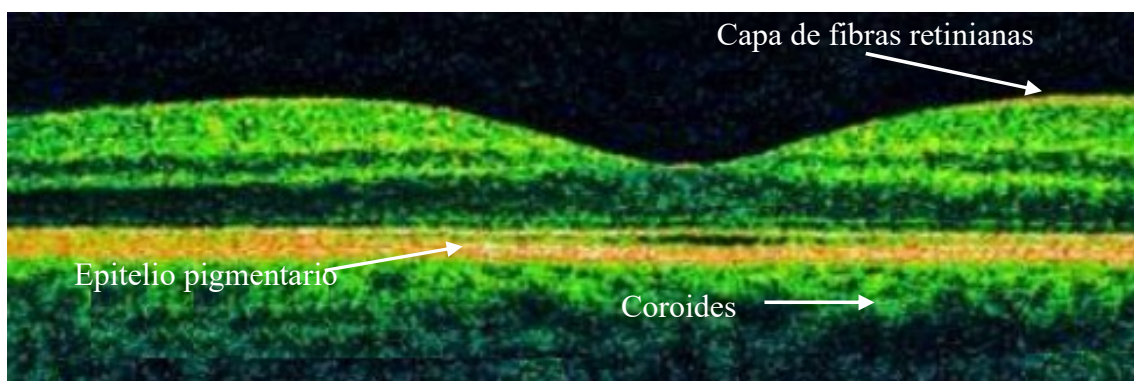


Figura 2: Segmentación de las capas de la retina en una imagen de OCT, indicando la reflectividad en código de color.

En el año 2008 se desarrollaron los primeros modelos de OCT de dominio espectral (Spectral Domain OCT o SD-OCT). Al contrario que en la OCT de dominio temporal, estos aparatos se caracterizan por utilizar un espejo de referencia estático y analizar la luz reflejada de la retina en un único punto. La formación de la imagen se basa en el espectro de frecuencias de la luz reflejada desde las distintas capas de la retina. A través de un análisis basado en la transformada de Fourier, estudia y subdivide la onda reflejada en las distintas frecuencias que la componen, elaborando el mapa de grosor retiniano a partir de estos datos.

La evolución del dominio temporal al dominio espectral en la OCT ha ido acompañada de una mejora en la resolución de las imágenes obtenidas, al no depender del movimiento de un espejo para el análisis del haz de luz reflejado. Así, la OCT de

dominio temporal alcanza un resolución de aproximadamente 12-15 μm , mientras que OCT de dominio espectral llega a una resolución de 5 μm (43).

2.2. ESTUDIO DE LA CAPA DE FIBRAS PERIPAPILAR MEDIANTE TOMOGRAGIA DE COHERENCIA OPTICA.

2.2.1. IMPORTANCIA DEL ESTUDIO DE LA CAPA DE FIBRAS NERVIOSAS PARA EL DIAGNOSTICO PRECOZ DEL GLAUCOMA

El glaucoma es una neuropatía óptica caracterizada por la pérdida de las células ganglionares de la retina y la degeneración de sus axones (44). Su diagnóstico está fundamentado actualmente en la apariencia del nervio óptico y el estudio de campo visual (2). Dada la gran variabilidad en la apariencia del nervio óptico, el diagnóstico del glaucoma se apoya primordialmente en un cambio en el tamaño o morfología de la excavación discal a lo largo del tiempo (45). No obstante, un cambio en la excavación discal detectado en oftalmoscopia o estereofotografía seriada representa la pérdida irrecuperable de miles de axones de las células ganglionares.

Diversos estudios han mostrado que existe una destrucción de hasta una 25 % a 40 % de las células ganglionares previo a la aparición de anomalías detectadas en el campo visual (44,46-48), lo que acentúa la necesidad de anticipar la detección de la enfermedad antes de la aparición de defectos campimétricos.

En estudios postmortem, Quigley et al. encontraron una reducción del 20 % de células ganglionares en áreas retinianas que mostraban una reducción luminosa en el campo visual de 5 dB, y un 40% de pérdida de células ganglionares en zonas con una reducción de sensibilidad de 10 dB (44). La evaluación de la relación estructura-función usando como parámetro estructural la medida del grosor de la capa de fibras nerviosas mediante OCT, ha mostrado una fuerte relación entre el grosor medio de la CFNR y la sensibilidad luminosa en el campo visual en el grupo de pacientes con desviación estándar del patrón (PSD) mayor de 1.9 dB y grosor medio de la CFNR inferior a 70 micras (49).

Los cambios estructurales pueden preceder en 5 o 6 años a la pérdida funcional detectable en el campo visual (50). Hood et al. han reportado que los cambios estructurales en el glaucoma precoz son mucho mayores que los cambios en la sensibilidad de la retina (51).

A modo de resumen, podemos concluir que la detección directa de una reducción de la CFNR permitiría diagnosticar la enfermedad o detectar su progresión de manera temprana, proporcionando la oportunidad de realizar una intervención terapéutica precoz capaz de prevenir el deterioro del campo visual, y por ende de la calidad de vida del paciente.

2.2.2. REPRESENTACION DE RESULTADOS DEL ESTUDIO DE LA CAPA DE FIBRAS NERVIOSAS CON OCT

La alta resolución de la imagen obtenida con OCT así como el desarrollo de algoritmos de segmentación ha permitido aislar la CFNR de otras capas retinianas, pudiendo medir y cuantificar su grosor (52). Los valores obtenidos se muestran en un círculo de 3.4 mm de diámetro centrado en el nervio óptico, indicando los grosores medios en los distintos cuadrantes y sectores horarios peripapilares. Cada uno de estos segmentos se codifica con un color, de acuerdo al grosor relativo de ese área comparado con una base normalizada de sujetos sanos y ajustada para la edad. Los colores verde, amarillo y rojo indican un 5% a 95%, 1% a 5% y menos de de 1% respectivamente de probabilidad de que el grosor calculado se encuentre dentro del valor normal para una población de esa edad. El color blanco indica una CFNR de mayor grosor que el 95% de los controles normales.

El diámetro más apropiado para realizar la medida del grosor de la CFNR circunpapilar ha sido estudiado por Heindl et al (53), comparando la capacidad de discriminar entre sujetos sanos y pacientes con glaucoma del valor promedio de la CFNR peripapilar medida en los diámetros de 3.5, 4.1 y 4.7 mm centrados en el nervio óptico. El autor concluye que el diámetro más pequeño, con 3.5 mm, es el que mejor diferencia entre los grupos control y glaucoma.

En la figura 3 se representa el formato estándar de un informe de análisis de la CFNR con OCT.

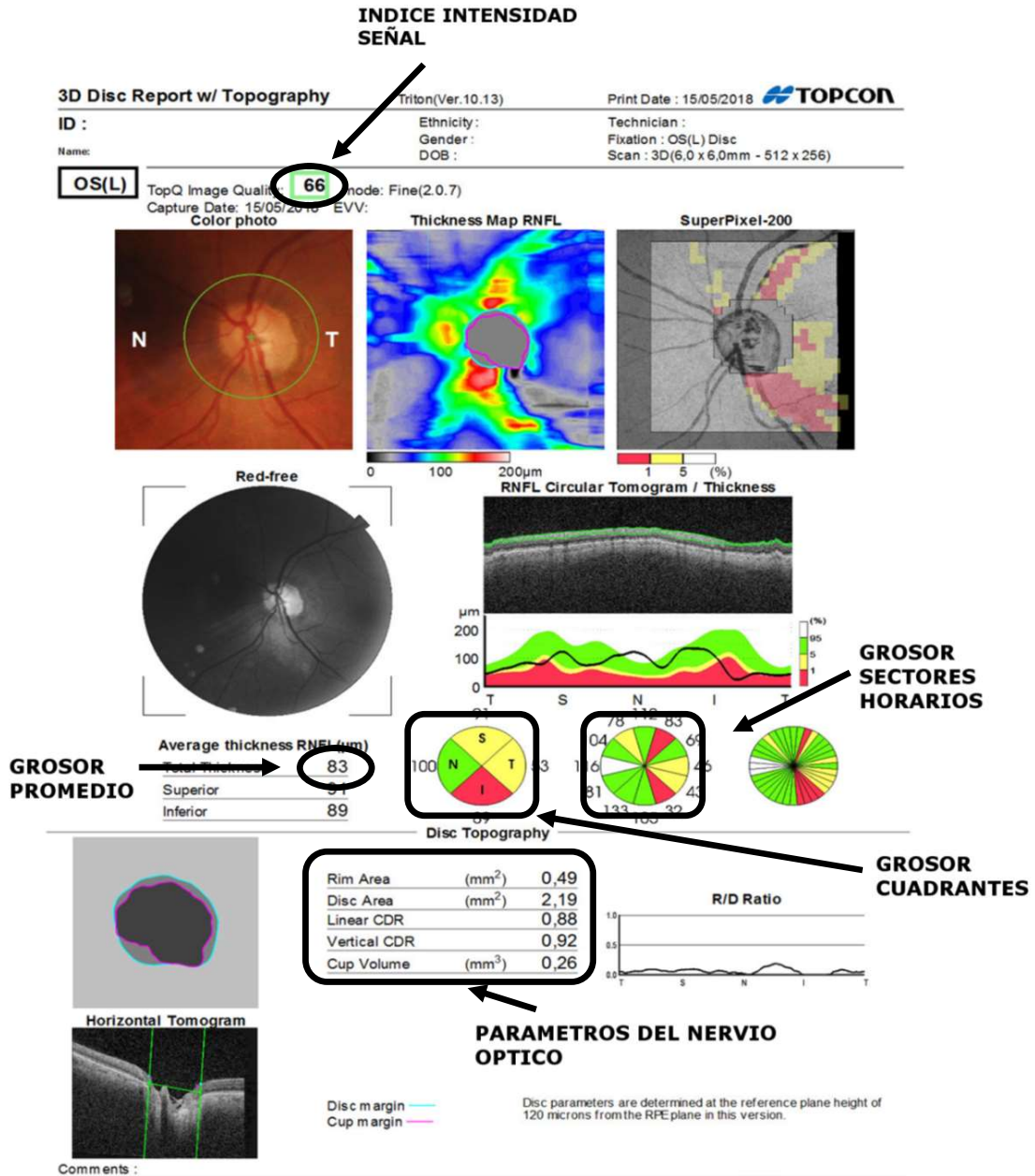


Figura 3: Informe estándar de estudio de la CFNR mediante OCT

2.2.3. ARTEFACTOS EN LAS IMÁGENES CON OCT

La presencia de artefactos en las imágenes obtenidas mediante OCT puede dar lugar a una interpretación incorrecta de la misma. De manera genérica, la causa de los artefactos puede responder a dos motivos:

- Problemas en la adquisición de la imagen, debido a opacidad de medios, movimiento del paciente, etc.
- Error en el algoritmo software de segmentación, situando inadecuadamente las líneas de segmentación en la retina. Este fallo da lugar a un cálculo incorrecto del grosor de las distintas capas de la retina.

La frecuencia y tipos de artefactos en las imágenes ha sido estudiado por Han et al (54). El artefacto más frecuente fue la segmentación incompleta de las capas de la retina, presente en cerca del 80% de las imágenes. No obstante, el autor indica que únicamente un 5 a 8% de los casos tuvieron significación clínica.

Al precisar menor tiempo para la adquisición de las imágenes, la OCT de dominio espectral ha mostrado menor frecuencia de artefactos en comparación con la OCT de dominio temporal (55). La rapidez en la adquisición de la imagen minimizaría los artefactos derivados del movimiento del paciente. Por otra parte, la mayor resolución de la SD-OCT daría lugar a menor frecuencia de errores de segmentación de las capas de la retina (56).

La calidad de la imagen puede ser estimada por el índice de intensidad de la señal o “signal-strength index (SSI)”. Sadda et al (57) demostraron que valores bajos de este índice en TD-OCT estaban asociados a errores de mayor magnitud en la segmentación. De manera general, las imágenes con mayor intensidad de la señal muestran una mayor claridad, mejor segmentación de las capas de la retina y mejor reproducibilidad de las medidas (58). No obstante, la intensidad de la señal no se correlaciona necesariamente con la calidad de la imagen. Así, Han et al (54) han observado imágenes sin ningún tipo de artefacto con valores de este índice de 4 (sobre un valor máximo de 10), e imágenes artefactadas con valores SSI de 10 sobre 10.

El valor calculado por OCT del grosor de la CFNR peripapilar puede oscilar de manera significativa según la calidad de la imagen, obteniendo medidas menores a peor intensidad de la señal (59).

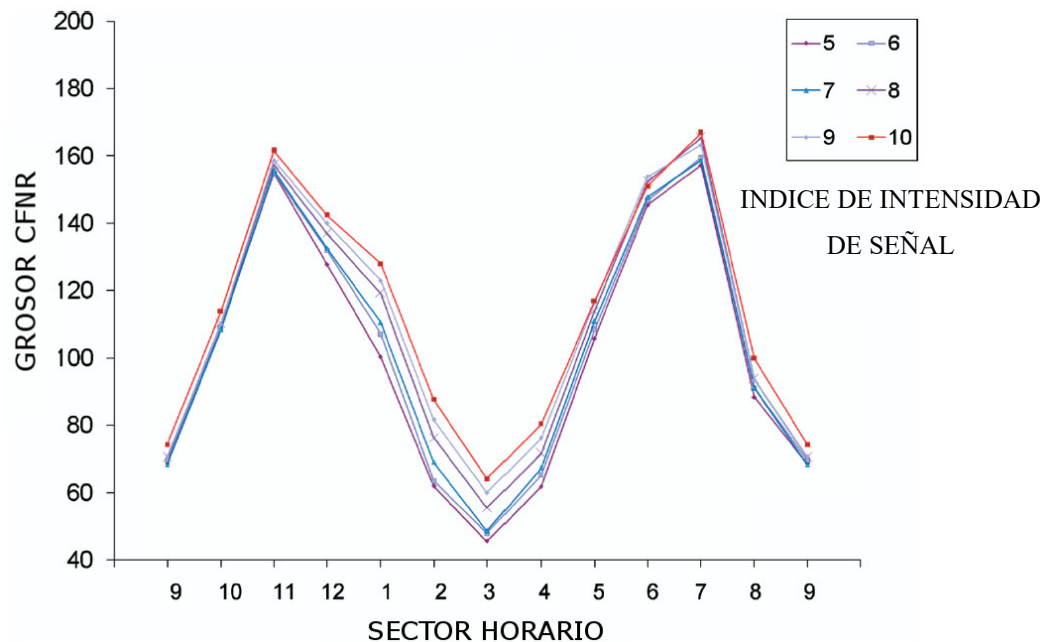


Figura 4. Variación de la medida grosor de la CFNR peripapilar para distintos valores del índice de intensidad de la señal (de 5 a 10)

Fuente: Cheung et al. Relationship between retinal nerve fiber layer measurement and signal strength in optical coherence tomography (59)

La alta frecuencia de artefactos en las imágenes obtenidas ha dado lugar a que se recomiende revisar en todos los casos si la imagen es correcta o presenta errores que pueden dar lugar a una interpretación inadecuada de los resultados. En el caso de la determinación del grosor de la CFNR, se recomienda que las imágenes tengan el mayor índice de intensidad de señal posible.

2.2.4. OCT EN EL DIAGNOSTICO DEL GLAUCOMA.

El uso efectivo de instrumentos para el diagnóstico médico requiere que estos tengan una alta precisión diagnóstica. Pero igual de importante es que posean una alta reproducibilidad en sus resultados. Esto es especialmente cierto en el manejo y

seguimiento del glaucoma, ya que una alta reproducibilidad es un factor crítico para monitorizar el progreso de la enfermedad.

La medida de la CFNR peripapilar con SD-OCT muestra una reproducibilidad excelente (60–62). Esta reproducibilidad es mayor conforme aumenta el índice de la intensidad de la señal de la imagen (58). De manera global, parece que el grosor medio de la CFNR circunpapilar es una de las medidas más reproducibles, lo cual podría sugerir que podría ser el mejor parámetro para monitorizar la progresión de la enfermedad.

La sensibilidad y especificidad del estudio de la CFNR con OCT para discriminar entre pacientes con glaucoma y sujetos sanos ha sido estudiada ampliamente.

Budenz et al (63) ha demostrado que el examen de la CFNR utilizando Stratus-OCT puede ser de utilidad para diferenciar entre individuos normales y pacientes con glaucoma perimétrico. Toma para su análisis la probabilidad de normalidad del grosor medio, de cada cuadrante, y de cada sector horario, comparándolo con la base normalizada de la OCT ajustada a la edad. Este ajuste para la edad es importante, ya que existe un adelgazamiento progresivo de la CFNR conforme aumenta la edad, con una tasa de reducción de 0.2% por año (64). Los mejores parámetros que encuentra para discriminar los grupos estudiados son:

- Uno o más de un cuadrante anormales con probabilidad $< 5\%$.
- Uno o más sectores horarios anormales con probabilidad $< 5\%$.

Begum et al (65) determina la sensibilidad y especificidad diagnóstica del examen de la CFNR en pacientes con glaucoma perimétrico y con glaucoma preperimétrico. En pacientes con glaucoma y presencia de defecto campimétrico, el parámetro con mayor sensibilidad es el grosor medio de la CFNR, seguido por la afectación del cuadrante inferior y del sector horario a las 6 horas. En pacientes con glaucoma y campo visual normal, el valor de la sensibilidad es muy bajo, alcanzando su máximo en la afectación del cuadrante inferior.

Mwanza et al (66) aporta datos de sensibilidad y especificidad para discriminar entre pacientes con glaucoma en estadio inicial e individuos normales. Encuentra que el

parámetro con mayor sensibilidad y especificidad es el grosor del cuadrante inferior, seguido del valor medio global de toda la CFNR.

En una publicación posterior, el mismo autor coincide con los datos anteriores al determinar el valor de sensibilidad de los parámetros de la OCT de CFNR en pacientes con glaucoma precoz. El parámetro con mayor sensibilidad es el cuadrante inferior, seguido del valor medio global de la capa de fibras peripapilar (67). Nouri-Mahdavi et al (68) encuentran una buena capacidad de clasificación con el estudio con OCT en sujetos con glaucoma precoz.

La combinación de los distintos parámetros obtenidos mediante OCT ha mostrado una alta tasa de éxito para discernir entre sujetos sanos y pacientes con glaucoma en estadio precoz. Se han propuesto múltiples modelos multivariable combinando parámetros de la CFNR, de la cabeza del nervio óptico y de la capa de células ganglionares con el fin de mejorar la habilidad diagnóstica de la OCT.

Un modelo de regresión logística multivariable que incluía los valores de la capa de células ganglionares, los parámetros de la cabeza del nervio óptico y un subgrupo de parámetros de la CFNR peripapilar fue altamente competente en la predicción de glaucoma precoz, mejorando a los modelos univariables en términos de área bajo la curva operativa del receptor (ROC) y tasa de clasificación correcta (69).

Larrosa JM et al (70) proponen un modelo predictivo multivariable para la detección del glaucoma utilizando una combinación de datos obtenidos mediante SD-OCT de la CFNR, capa de células ganglionares y parámetros de la cabeza del nervio óptico. Este trabajo diseña una calculadora diagnóstica cuyos resultados de clasificación para la detección del glaucoma mejoraba los obtenidos utilizando los mejores parámetros de modo aislado.

La aplicación de algoritmos matemáticos basados en análisis discriminante, redes neuronales, lógica borrosa, algoritmos genéticos y clasificadores automáticos en el diagnóstico del glaucoma es un área en pleno desarrollo. Una función discriminante lineal que incluía múltiples valores seleccionados de la CFNR y del nervio óptico fue capaz de clasificar a ojos con glaucoma y ojos normales con un valor del área bajo la

curva ROC (AROC) que superaba el obtenido con cualquier parámetro individualmente (71).

En resumen, hay una base amplia en la literatura que sostiene que la combinación de parámetros en el estudio de imagen de la CFNR y nervio óptico supera al uso aislado de estos en la tarea de clasificación entre ojos normales y ojos con glaucoma.

2.2.5. INFLUENCIA DEL TAMAÑO DISCAL EN LA MEDIDA DE LA CAPA DE FIBRAS NERVIOSAS PERIPAPILAR.

El tamaño grande del disco óptico así como el grosor corneal central han sido identificados como factores de confusión en el diagnóstico del glaucoma (72).

La medida y representación gráfica del grosor de la CFNR de la retina se hace de modo estandarizado en todas las OCT. La determinación del grosor se efectúa a una distancia fija de 1.7 mm del centro geométrico del nervio óptico. Los resultados se muestran representados en una circunferencia centrada en la papila óptica, indicando los distintos grosores medios de la CFNR para cada cuadrante y sector horario.

Budenz et al (73) describe la presencia de variaciones significativas del grosor de la CFNR con la edad, raza, longitud axial y área del nervio óptico. Los ojos con un nervio óptico de mayor tamaño tenían una CFNR más gruesa. Savini et al (74) encuentra un resultado similar en sujetos caucásicos tomando en cuenta la edad y diversos parámetros del nervio óptico. Este hallazgo podría justificarse si un nervio óptico de mayores dimensiones tuviera un número superior de axones, tal y como encuentra Quigley en estudios histológicos en monos (75). Este hallazgo, sin embargo, no se ha confirmado en estudios en ojos humanos (76).

Un estudio histológico practicado en ojos humanos de sujetos sanos mostró que el grosor de la CFNR disminuye conforme nos alejamos del borde del disco óptico (77). Dado que la determinación del grosor de la CFNR se hace a una distancia fija del centro del nervio óptico, en papilas grandes la medida se efectúa más cerca del borde del disco óptico, donde es de esperar que haya mayor espesor de fibras.

Independientemente de la explicación, la relación existente entre grosor de la CFNR y tamaño papilar sugiere que la potencia diagnóstica de la OCT podría mejorarse si se tuviera en consideración el área discal (73), ya que la precisión en la medida de la CFNR podría estar limitada por un tamaño discal fuera del rango de normalidad poblacional (74).

2.3. ESTUDIO DE LA CABEZA DEL NERVIO OPTICO CON OCT.

La neuropatía óptica glaucomatosa es un proceso crónico donde acontece una pérdida progresiva del tejido de la capa de fibras nerviosas de la retina. Este suceso, sin embargo, es difícil de detectar en la exploración clínica. Por ello, la exploración oftalmológica se orienta con mayor frecuencia a la detección de signos de daño glaucomatoso en la cabeza del nervio óptico, buscando un adelgazamiento del rodete neuroretiniano, ya sea focal o generalizado. No obstante, esta valoración clínica de la papila tiene un alto componente de subjetividad y es altamente dependiente de la experiencia del médico que la realiza.

La tecnología de la OCT nos ha facilitado un método para el análisis del disco óptico de manera objetiva. Para ello, el software automatizado define el margen del nervio óptico tomando como referencia la terminación de la membrana de Bruch (78). Con esa referencia inicial, la OCT calcula los parámetros específicos para ese nervio óptico, siendo los más habituales:

- Área de disco.
- Área de excavación discal.
- Área de anillo neuroretiniano.
- Relación vertical u horizontal de E/D.

Se ha demostrado que existe correlación entre los parámetros estructurales de la papila, determinados con OCT, y los índices principales de la perimetría automatizada en pacientes con glaucoma y defectos campimétricos, aunque esta correlación fue más débil que la obtenida con los parámetros basados en el estudio de la CFNR (79). Jeoung et al (80) han observado que el área del anillo neuroretiniano es el parámetro que mejor discrimina entre ojos normales y ojos con glaucoma, tanto en estadio precoz como

moderado-avanzado. En su trabajo, Anton et al (81) encuentra resultados similares, identificando al área de rodete neuroretiniano como el parámetro con mayor área bajo la curva ROC, seguido por el valor de la E/D vertical.

En pacientes con glaucoma preperimétrico, Na et al (82) encuentran que los valores que mayor sensibilidad diagnóstica son la relación horizontal de E/D, el área de excavación discal y el área de anillo neuroretiniano.

3. MINERIA DE DATOS.

3.1. INTRODUCCION A LA MINERIA DE DATOS.

Elegir la decisión más apropiada se ha convertido en un factor clave en cualquier área de trabajo. La información reduce nuestra incertidumbre sobre algún aspecto de la realidad, y por tanto, nos permite tomar mejores decisiones. Hay tantos modos para encontrar esta decisión como personas distintas participen en ello, ya que cada una elaborará su propio plan personal de resolución. Sin embargo, la metodología subyacente es la misma en todos los casos: la resolución habitualmente se establece a partir de:

- El análisis de la información obtenida a partir de experiencias personales previas de casos similares ya resueltos.
- La información contenida en estudios científicos recientes.
- Un juicio particular de cada profesional.

El número de casos resueltos, de nuevos estudios y de datos disponibles crece de modo continuo y con un ritmo de crecimiento enormemente acelerado. Podría pensarse con ello que el proceso de toma de decisiones se está facilitando gradualmente a lo largo del tiempo, haciéndose cada vez más sencillo y fiable. No obstante, ocurre justo el fenómeno contrario: se está complicando progresivamente debido a que el profesional es incapaz de procesar adecuadamente la ingente cantidad de datos de los que dispone.

Desde la antigüedad, los sistemas que realizan la recopilación y organización de información han resultado de gran utilidad para asistir en la toma de decisiones. En las últimas décadas, el almacenamiento, organización y recuperación de la información

guardada se ha automatizado gracias a los sistemas de bases de datos. Cada día generamos una gran cantidad de información, a veces conscientes de que lo hacemos y otras veces de modo inconsciente. En la actualidad, prácticamente cualquier acto llevado a cabo mediante tecnología digital lleva aparejado la producción de una información que es almacenada. Podríamos decir que hoy día no existe casi ninguna faceta de la realidad de la que no se disponga de información de manera electrónica, ya sea estructurada, en forma de bases de datos, o no estructurada, en forma textual o hipertextual (83).

Pero ¿qué nos ha permitido poder generar y almacenar tal cantidad de información? En los últimos años, debido al extraordinario desarrollo tecnológico, tanto en el área de cómputo como en el de transmisión de datos, ha sido posible mejorar y optimizar el manejo y almacenamiento de la información. Molina et al especifican tres factores importantes que nos han llevado a este hecho (84):

- Abaratamiento de los sistemas de almacenamiento de la información, tanto en soportes temporales como de modo permanente.
- Incremento de las velocidades de cómputo de los procesadores.
- El desarrollo de sistemas para el manejo de bases de datos de gran potencia, que han posibilitado la administración de bases de datos de tamaño inimaginable hasta ahora.

La cantidad de información que nos llega cada día es tan inmensa que nos resulta difícil asimilarla. La minería de datos surge como una tecnología que intenta ayudar a comprender el contenido de una base de datos modelando el conocimiento que los datos encierran implícitamente. La minería de datos no aparece por el desarrollo de tecnologías esencialmente diferentes a las anteriores, sino que se crea, en realidad, por la aparición de nuevas necesidades y, especialmente, por el reconocimiento de un nuevo potencial: el valor, hasta ahora generalmente infrautilizado, de la gran cantidad de datos almacenados en formato digital (83).

Los datos pasan de ser un “producto” o resultado de un almacenamiento digital, a ser una “materia prima” que hay que explotar para obtener el verdadero “producto elaborado” o conocimiento.

De modo general, los datos son el elemento principal y origen de la minería de datos. En el momento en que se les atribuye algún significado especial por algún usuario pasan a convertirse en información. Posteriormente, un grupo de analistas o especialistas elaboran un modelo que pueda explicar o representar esa información, pasando a referirnos a conocimiento. La jerarquía que existe entre estos conceptos queda ilustrada en la figura 5, apreciándose la estrecha unión entre dato e información, no así entre la información y el conocimiento (85). El paso de información a conocimiento no es directo, sino que requiere de técnicas de análisis como las proporcionadas por la minería de datos.

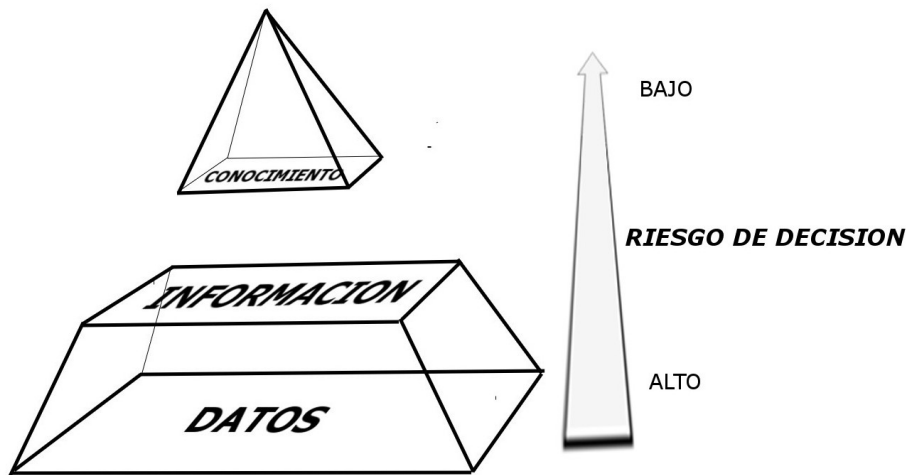


Figura 5. Representación de la pirámide del conocimiento. La imagen muestra el volumen que presenta cada nivel y el nivel de riesgo en la toma de decisiones basadas en cada nivel jerárquico.

La minería de datos trabaja en el nivel superior de la pirámide del conocimiento, buscando patrones, comportamientos, agrupaciones, tendencias o asociaciones que puedan recogerse en forma de modelos. A su vez, estos modelos nos permitirán comprender mejor el dominio de conocimiento así como asistir al profesional en la tarea de toma de decisiones en el campo de estudio.

3.2. QUE ES LA MINERIA DE DATOS?

El campo de la minería de datos es relativamente nuevo y está en estado de evolución. Es un área de estudio que ha surgido a través de la convergencia de otras disciplinas. Utiliza métodos de la inteligencia artificial, aprendizaje automático, estadística y sistemas de bases de datos.

La minería de datos emergió durante los últimos años de la década de 1980, evolucionando enormemente en los años 90, y continuando su fuerte crecimiento en el nuevo milenio. El primer impulso importante para el desarrollo de esta disciplina se dio en “The First International Conference on Knowledge Discovery and Data Mining dentro de la Fourteenth International Joint Conference on Artificial Intelligence” coordinada por Usama Fayad y Ramasamy Uthurusamy en 1995 (86). Hay una gran variedad de definiciones para el concepto de minería de datos, entre las que encontramos (87):

- Proceso de exploración y análisis, por medios automáticos o semiautomáticos, de grandes cantidades de datos con el fin de descubrir patrones significativos y reglas.
- Proceso de descubrir nuevas correlaciones significativas, patrones y tendencias por indagación a través de grandes cantidades de datos almacenados en repositorios, usando tecnologías así como matemáticas y estadísticas.
- Proceso de extraer conocimiento útil y comprensible, previamente desconocido, desde grandes cantidades de datos almacenados en distintos formatos.

De las definiciones se desprende que la tarea fundamental de la minería de datos es encontrar modelos inteligibles, a partir de los datos, mediante procesos automáticos o semi-automáticos, de manera que el uso de los patrones descubiertos ayude en la toma de decisiones más seguras. Entendemos por modelo a un algoritmo establecido a partir de las relaciones y patrones entre los datos y que puede usarse para hacer predicciones.

El nacimiento de la minería de datos tiene lugar como resultado de la convergencia de diversos factores, entre los que podemos destacar:

- Exigencia de disponer de nuevas técnicas de análisis capaces de obtener conocimiento a partir de los enormes volúmenes de información disponibles, en forma de bases de datos estructuradas, almacenados en algún tipo de repositorio.
- La difusión generalizada de herramientas informáticas, habilitando su uso generalizado por los grupos de científicos.

Su germen es, por tanto, la necesidad de concebir nuevos métodos y técnicas especializadas y específicamente orientadas al análisis de bases de datos de gran tamaño.

El concepto de minería de datos es frecuentemente utilizado de modo inapropiado para referirse a cualquier almacén de datos a gran escala o de procesamiento de la información (recolección, extracción, almacenamiento, análisis y estadísticas). El término, sin embargo, debe ser empleado únicamente para referirse al análisis de grandes cantidades de datos para extraer patrones originales o novedosos, hasta entonces desconocidos. El concepto clave que involucra la minería de datos es el descubrimiento, comúnmente definido como "la detección de algo nuevo".

3.3. MINERIA DE DATOS FRENTE A ESTADISTICA CLASICA.

La estadística ha sido la primera ciencia que ha considerado a los datos como su materia prima, aportando un conjunto de herramientas analíticas que son empleadas en su análisis. Algunos paquetes estadísticos son capaces de generar modelos al inferir patrones a partir de los datos, utilizando modelización estadística paramétrica o no paramétrica. No obstante presentan una serie de problemas y limitaciones tales como (83):

- Generalmente no funcionan bien para las bases de datos actuales, con tamaño de cientos de tablas o millones de registros y una alta dimensionalidad.
- Los resultados resultan especialmente crípticos para los no estadísticos.

- Los modelos que establece surgen como una parametrización de un modelo preestablecido o intuitivo por el usuario, no siendo un modelo novedoso y original, como ocurre en la minería de datos.

Todos estos obstáculos e inconvenientes de las aproximaciones estadísticas convencionales han hecho surgir una nueva clase de herramientas y técnicas para sustentar la extracción de conocimiento aprovechable y novedoso desde la información disponible en grandes bancos de datos, y que confluyen en el concepto de la minería de datos.

3.4. DESCUBRIMIENTO DEL CONOCIMIENTO EN BASES DE DATOS.

En la era digital en la que nos encontramos, la información en los diversos campos de investigación es recogida y almacenada con un ritmo de crecimiento exponencial. Se han creado bases de datos de dimensiones excepcionales. Como consecuencia natural de ello, ha surgido la necesidad de impulsar una nueva generación de teorías e instrumentos computacionales que nos asistan en la extracción de conocimiento útil de ese extenso volumen de información. Estas teorías e instrumentos son el objeto de estudio de un área emergente denominada Descubrimiento del Conocimiento en Bases de Datos (DCBD) (88).

Para decidir cuál es la técnica más adecuada para la extracción del conocimiento en cada situación es necesario distinguir el tipo de conocimiento que se desea obtener de los datos.

Según su nivel de abstracción, el conocimiento contenido en los datos puede clasificarse en distintas categorías (figura 6):

- Conocimiento evidente. Se obtiene mediante una consulta simple y sencilla al sistema.
- Conocimiento multidimensional. Considera los datos con una cierta estructura n-dimensional, representándolos a modo de tabla con n atributos independientes. Permite detectar algunas regularidades difíciles de observar con la representación monodimensional clásica.

- Conocimiento oculto. Es información no evidente, desconocida a priori y potencialmente útil. Es de gran valor, ya que no es conocida previamente y se trata de un descubrimiento real de nuevo conocimiento, abriendo la posibilidad de una nueva visión del problema planteado.

Las técnicas de minería de datos van dirigidas a recuperar el conocimiento en el nivel de abstracción de conocimiento oculto, buscando regularidades que a los seres humanos nos resulta difícil de ver.

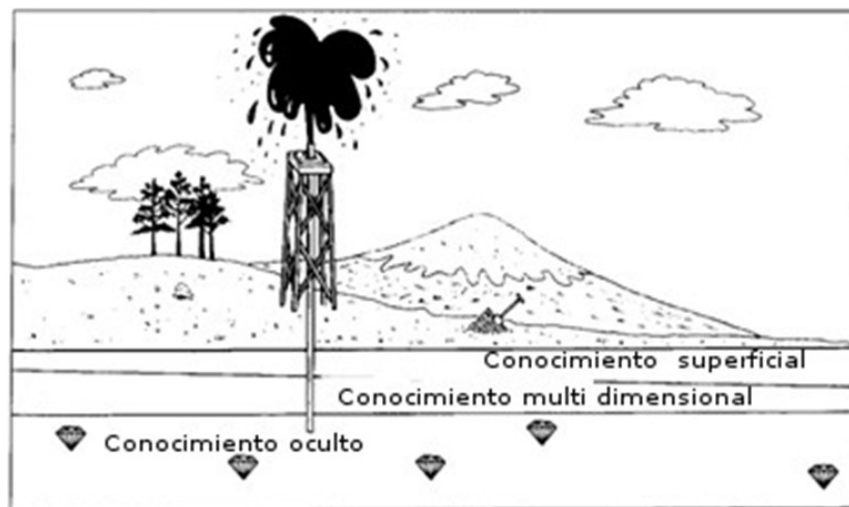


Figura 6: Niveles del conocimiento en función de su grado de abstracción.
Fuente: P.Adriaans. Addison-Wesley. Data mining. 1996 .(89)

Fayyad et al (88) definen el Descubrimiento de conocimiento en Bases de Datos como el proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y, en última instancia, comprensibles a partir de los datos.

En esta definición se resumen cuáles deben ser las propiedades deseables del conocimiento extraído. Este conocimiento debe ser (90):

1. Válido: hace referencia a que los patrones deben seguir siendo precisos para datos nuevos (con un cierto grado de certidumbre), y no sólo para aquellos que han sido usados en su obtención.
2. Novedoso: que aporte algo desconocido para el usuario.
3. Potencialmente útil: la información debe conducir a acciones que reporten algún tipo de beneficio para el usuario.

4. Comprensible: la extracción de patrones no comprensibles dificulta o imposibilita su interpretación, revisión, validación y uso en la toma de decisiones. De hecho, una información incomprensible no proporciona conocimiento (al menos desde el punto de vista de su utilidad).

El DCBD es la convergencia del aprendizaje automático, la estadística, el reconocimiento de patrones, la inteligencia artificial, bases de datos y sistemas de apoyo a la toma de decisiones (91).

3.4.1. FASES DEL DESCUBRIMIENTO DE CONOCIMIENTO EN BASES DE DATOS.

Al tratarse de un proceso, el DCBD incluye no sólo la obtención de los modelos o patrones (el objetivo de la minería de datos), sino también la preparación de los datos, la evaluación de los resultados y la interpretación de los mismos. En la figura 7 se representa de modo resumido el proceso de DCBD.

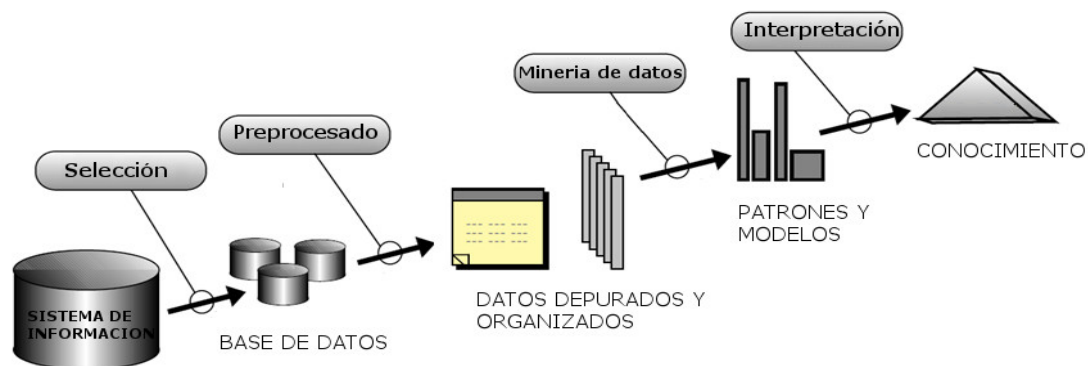


Figura 7. Resumen de las fases que componen el proceso de descubrimiento de conocimiento en bases de datos.

El proceso de DCBD es iterativo e interactivo, involucrando una cadena de pasos en los que el usuario o el experto en el dominio del problema deben tomar determinadas decisiones. Al ser iterativo, la salida de alguna de las fases puede hacer volver a pasos anteriores, siendo a menudo necesarias varias iteraciones para extraer conocimiento de alta calidad.

De manera general, las etapas a desarrollar se pueden concretar en la siguiente secuencia (88):

1. Comprensión del dominio del problema y del conocimiento previo relevante, identificando cual es la meta del proceso de DCBD.
2. Creación o selección de una base de datos sobre la que se llevará a cabo el descubrimiento. En algunos casos, selección de un subgrupo de variables o datos sobre los que se aplicará el proceso.
3. Limpieza y preprocesado de los datos, llevando a cabo operaciones básicas tales como eliminación de ruido sobre los datos, detección de datos de anómalos o atípicos, etc.
4. Reducción de los datos, encontrando características útiles para representar a los datos dependiendo de la tarea a realizar y del objetivo buscado. Esta etapa incluye la reducción de la dimensionalidad o métodos de transformación para reducir el número efectivo de variables consideradas.
5. Elección del algoritmo de minería de datos a utilizar, escogiendo los métodos más adecuados para la búsqueda de patrones en los datos.
6. Fase de minería de datos propiamente dicha, buscando modelos o patrones de interés con alguna forma de representación particular: reglas o árboles de clasificación, árboles de regresión, etc.
7. Interpretación de los patrones obtenidos, posiblemente retornando a alguna etapa anterior para una nueva iteración del proceso. En este paso también se incluye la visualización de los modelos extraídos.
8. Consolidación del nuevo conocimiento, documentándolo e incorporándolo a un sistema nuevo para acciones futuras.

La minería de datos, por tanto, no es más que un paso esencial en el proceso más amplio de DCBD, siendo con frecuencia malinterpretado, asimilándolo al proceso completo de extracción de conocimiento.

Orallo et al. (83) refunde parte de los pasos descritos por Fayyad en el proceso de DCBD, describiendo únicamente cinco fases:

1. Fase de integración y recopilación, constituyendo un almacén de datos a partir de los datos iniciales.
2. Fase de selección limpieza y transformación, eliminando los datos irrelevantes o innecesarios, basándonos en nuestro conocimiento del dominio del problema.
3. Fase de minería de datos, produciendo un conocimiento novedoso que pueda utilizar el usuario.
4. Fase de evaluación e interpretación.
5. Toma de decisiones.

3.5. NOMENCLATURA BASICA UTILIZADA EN LA MINERIA DE DATOS.

A fin de evitar confusión y ambigüedad en las descripciones de la minería de datos, realizamos una serie de definiciones de los términos utilizados más frecuentemente.

- Aprendizaje automático: faceta de la inteligencia artificial especializada en el desarrollo de algoritmos sistemáticos que tienen la capacidad de aprender autónomamente a partir de los datos, sin intervención ni asistencia humana para realizar modificaciones del algoritmo o de su programación.
- Algoritmo: conjunto de reglas que sigue una máquina o computadora para conseguir un resultado determinado. Puede ser concebido como una receta que define las entradas y las instrucciones o pasos requeridos para obtener el resultado a partir de las entradas.
- Modelo: es una estructura obtenida como trasposición o especialización de un algoritmo, El algoritmo indica el “cómo” llevar a cabo una tarea. El modelo particulariza el algoritmo a un conjunto de entradas determinado, concretando una arquitectura específica para esos datos. En el caso de las tareas de clasificación, el modelo puede ser denominado clasificador o predictor.
- Base de datos: hace referencia a una tabla que contiene datos o información estructurada.
- Instancias y atributos: la base de datos está constituida por un conjunto de instancias, también llamadas tuplas, observaciones, ejemplos u objetos, cada una de las cuales tiene, a su vez, una serie de atributos, también llamados parámetros o características, que definen la diferencia entre las observaciones.

En nuestro dominio de estudio, los ejemplos o instancias corresponderían a cada uno de los pacientes incluidos en el estudio. Los atributos o parámetros de las instancias serían los grosores de la CFNR en cada cuadrante y sector horario. En la representación convencional en forma de tabla, las instancias corresponderían a las filas y los atributos a las columnas de la tabla.

- Etiqueta: incumbe a un tipo de atributo especial de las instancias, que define la clase a la que pertenece. Es lo que el algoritmo intenta predecir a través de su aprendizaje. En nuestro estudio, la etiqueta de los ejemplos es la presencia o no de la enfermedad de glaucoma.
- Predicción: es lo que el modelo infiere que será la etiqueta del objeto presentado al sistema, basándose en el valor de sus atributos.

3.6. CLASIFICACION DE LAS TECNICAS DE MINERIA DE DATOS.

Dentro del epígrafe de minería de datos, se engloban una gran cantidad de técnicas dirigidas a llevar a cabo distintas tareas de aprendizaje. Las clasificaciones posibles de estos métodos pueden ser diversas según el dominio de estudio así como el objetivo buscado.

De las posibles subdivisiones, hemos optado por dividir los métodos de minería de datos según dos conceptos:

- Existencia o no de una etiqueta que clasifica los ejemplos de la base de datos entre ejemplos positivos y negativos, enfermos y sanos, etc..
- Grado de complejidad del modelo resultante del análisis de la base de datos.

3.6.1. TECNICAS SUPERVISADAS VS TECNICAS NO SUPERVISADAS.

La diferenciación entre técnicas supervisadas y no supervisadas viene dada por la presencia en el banco de datos de un campo o atributo especial, denominado etiqueta, que determina la propiedad de los ejemplos que queremos predecir a través del modelo. En los métodos supervisados, cada ejemplo pertenecería a una clase distinta en función del valor de su etiqueta. En el área médica, esta etiqueta acostumbra a ser el diagnóstico,

pronóstico u otra característica del paciente que pretendemos predecir con la mayor precisión posible. En el caso de clasificación binaria, se trataría de la presencia o ausencia de la enfermedad.

En el aprendizaje supervisado, el objetivo de los métodos de minería de datos es obtener una descripción de las relaciones entre los atributos o características de los ejemplos que sea capaz de definir la clase a la que pertenece. Se crea un modelo que predice la clase a la que pertenece un ejemplo en función del resto de atributos distintos a la etiqueta. La tarea más frecuente de estas técnicas supervisadas es la clasificación de los ejemplos.

Como características principales de estas estrategias supervisadas podemos destacar (92):

- La tarea que intenta emular es clasificar correctamente los ejemplos nuevos que no han sido presentados al sistema. Su propósito, por tanto, va más allá de clasificar correctamente los objetos conocidos, debiendo predecir adecuadamente la clasificación correcta de los ejemplos no observados previamente.
- La fiabilidad y corrección de los modelos obtenidos dependen directamente de la veracidad de los ejemplos de entrada al sistema utilizados para su aprendizaje.

El aprendizaje no supervisado es una técnica de análisis de datos en la que los ejemplos no están etiquetados con la clase a la que pertenecen. En este caso, la resolución del problema se lleva a cabo agrupando los ejemplos en las particiones más probables según algún criterio o característica preestablecida que nos permite establecer cuanto de diferentes son dos ejemplos entre sí. El criterio empleado con más frecuencia es la distancia euclídea entre los ejemplos.

Los métodos de agrupamiento o conglomerados se vinculan al aprendizaje no supervisado. Su objetivo es encontrar las particiones más probables y con el mayor sentido lógico dentro de todas las posibles. El proceso de agrupamiento no es más que la organización de una colección de elementos en un conjunto de grupos homogéneos. Los algoritmos no supervisados nos permiten llevar a cabo tareas más complejas

comparado a las técnicas supervisadas, aunque sus resultados también pueden ser más impredecibles (93).

La ventaja fundamental que ofrece el aprendizaje no supervisado es su capacidad para encontrar todos los patrones presentes en los datos, al no estar limitado por la presencia de una etiqueta de clase.

De manera gráfica, en la figura 8 se representan una serie de casos no etiquetados dentro de un espacio bidimensional. Tomando como criterio la distancia euclídea entre los casos, representados cada uno por un punto del plano, estos podrían agruparse de manera intuitiva en tres conjuntos, de manera que la distancia entre los elementos dentro de cada grupo sea mínima y la distancia entre los elementos pertenecientes a grupos distintos sea máxima.

Cada partición viene representada por un elemento central denominado centroide. Las regiones se definen minimizando la suma de las distancias cuadráticas entre cada elemento de una partición y el centroide de su correspondiente partición.

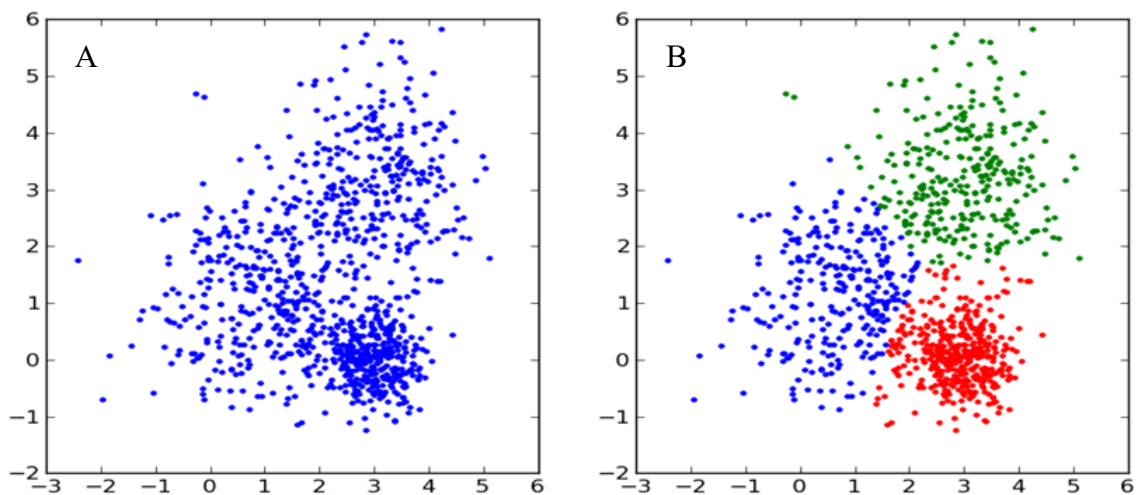


Figura 8: Ejemplo de algoritmo no supervisado para realizar la partición de los datos en agrupamientos homogéneos. A) Conjunto de datos representados por puntos distribuidos en un espacio bidimensional B) Datos agrupados con el algoritmo k-medias en 3 particiones, siguiendo el criterio de minimizar la distancia euclídea

Su objetivo es hallar una estructura o patrón de agrupamiento en una colección de datos no categorizados en clases. Los datos son procesados para encontrar una serie de conjuntos o grupos homogéneos en los que se subdividen los datos. Para que esto sea posible es condición sine qua non que exista algún tipo de asociación que vincule los ejemplos entre sí. Una distribución aleatoria de los ejemplos no encontraría ningún tipo de partición. La presencia de agrupaciones indica que los datos contienen suficiente información como para poder ser divididos en una serie de grupos naturales.

Las diferencias entre aprendizaje supervisado y no supervisado se resumen en la tabla 4.

Tabla 4

Diferencias entre aprendizaje supervisado y no supervisado

	Supervisado	No supervisado
Proceso	Existen variables de entrada y etiqueta de salida.	Solo existen variables de entrada.
Datos	Datos etiquetados	Datos no etiquetados
Algoritmos fundamentales	Clasificación	Agrupación
Complejidad computacional	Baja	Alta
Precisión de resultados	Alta	Baja
Número de clases	Conocido a priori	Desconocido a priori

3.6.2. MODELOS SEGÚN SU GRADO DE COMPLEJIDAD.

La cantidad y variedad de algoritmos posibles en el campo de la minería de datos ha ido creciendo a un ritmo similar al incremento exponencial que ha experimentado la capacidad de almacenamiento de datos. La presencia de repositorios de datos de cada vez mayor tamaño ha originado una demanda de algoritmos de análisis capaces de manejar un enorme volumen de información con métodos eficientes y en tiempos computacionales asumibles por los equipos informáticos actuales.

Por otra parte, la variedad de datos existente, su estructura interna presente o ausente, así como el propósito buscado delimitan cuales son las técnicas que pueden emplearse para la resolución del problema.

Los modelos generados por la aplicación de los métodos de minería de datos a una base de datos caminan paralelos a los algoritmos que los crean. Son por ello tan variados como lo son los algoritmos a los que van aparejados.

Referido la tarea de clasificación, los tipos de modelos posibles pueden categorizarse en función de múltiples parámetros. La figura 9 muestra de manera resumida algunas de las posibilidades que existen de modelos de minería de datos y sus paradigmas.

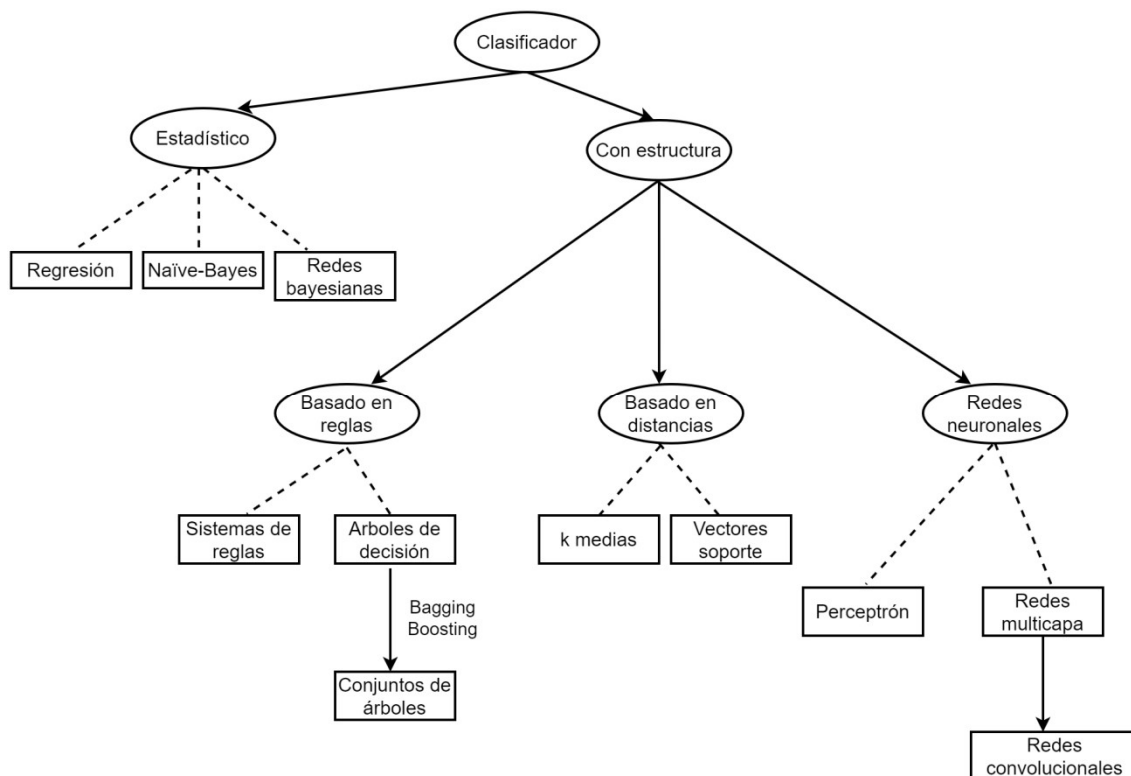


Figura 9. Clasificación de las técnicas de minería de datos según su metodología.

La complejidad del modelo, sin embargo, es una característica especialmente significativa e importante por la estrecha relación que mantiene con otras dos características fundamentales: la precisión del modelo y su interpretabilidad.

La precisión en un modelo de clasificación puede ser calculada en función de diversas métricas. El parámetro más empleado como equivalente a precisión es la proporción de casos nuevos correctamente clasificados, esto es, que la predicción del modelo coincide con la etiqueta real de la observación.

La figura 10 evidencia la relación existente entre el grado de complejidad y la precisión de un modelo, apreciándose la correspondencia directa que hay entre ambas variables. Los modelos más complejos alcanzan mayor porcentaje de observaciones correctamente clasificadas.

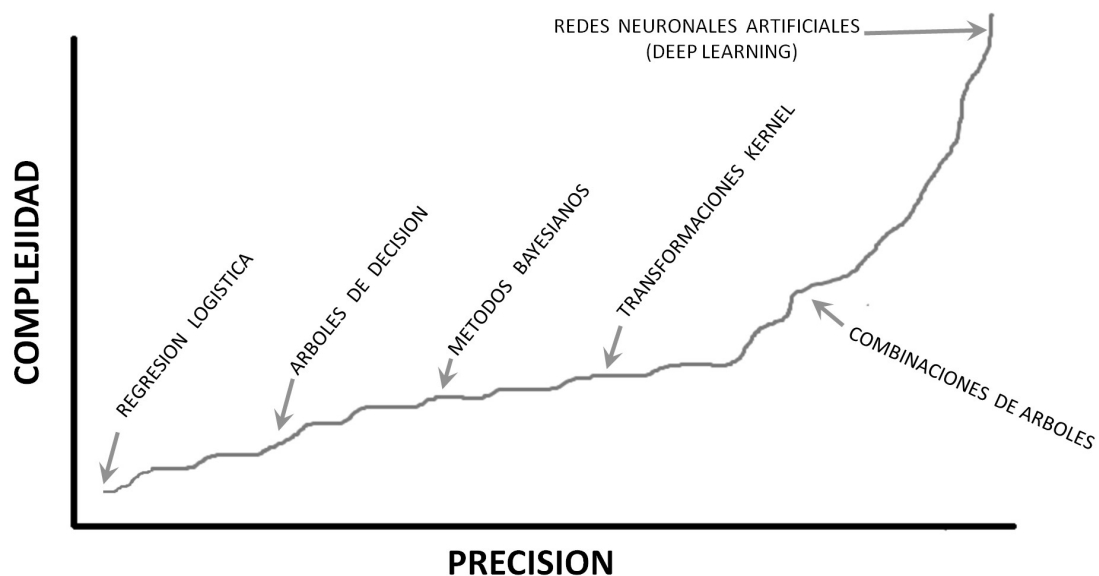


Figura 10. Correlación entre complejidad de un modelo y su precisión.

Algunos de los modelos de menor complejidad están enlazados con técnicas estadísticas convencionales, tales como el análisis de regresión logística. Otros, como las redes Bayesianas, están interpoladas entre la estadística clásica y la minería de datos en sentido estricto.

El mayor grado de complejidad lo alcanzan las redes neuronales artificiales, consiguiendo también el mayor nivel de precisión o clasificación correcta. La contrapartida, sin embargo, es su muy escasa o nula interpretabilidad.

3.6.2.1. ARBOLES DE DECISION O DE CLASIFICACION.

Los árboles de decisión o de clasificación son una de las posibles aproximaciones al aprendizaje supervisado. El algoritmo que realiza su construcción es uno de los métodos de aprendizaje más sencillos de entender y usar.

Un árbol de decisión es una representación gráfica de un procedimiento para clasificar los objetos. Intuitivamente consiste en una secuencia de condiciones organizadas jerárquicamente para llegar a una predicción final. De manera gráfica, viene simbolizado por una estructura de árbol donde cada nodo representa una condición o test sobre alguna característica o atributo de los ejemplos presentados al sistema y cada rama que parte de ese nodo corresponde a un posible valor para ese atributo. Los nodos finales u hojas del árbol están asociados a la predicción de la clase a la que pertenece el ejemplo (93).

En la figura 11 se muestra un ejemplo de árbol de decisión.

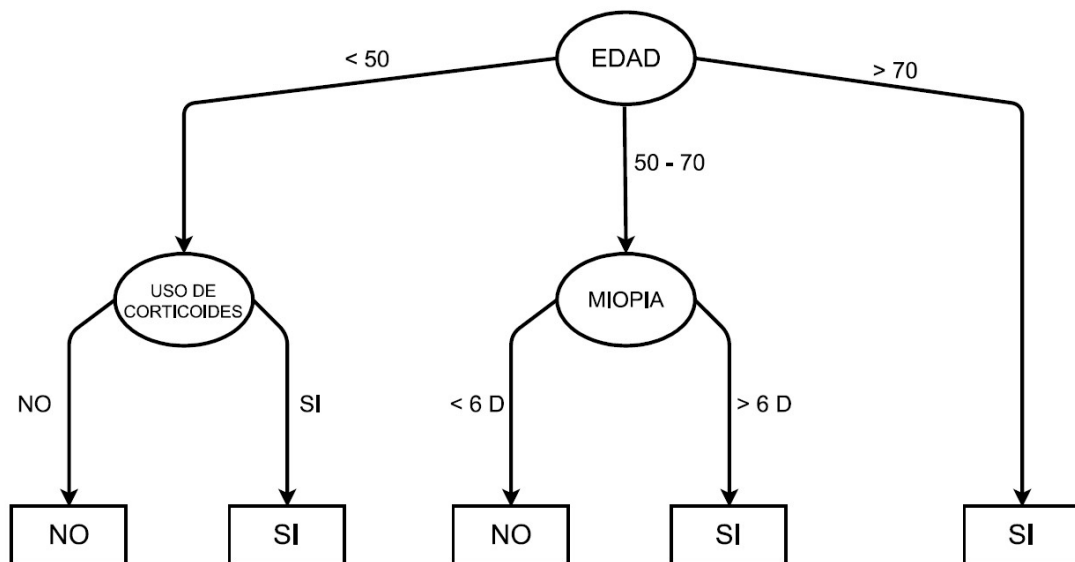


Figura 11. Ejemplo de estructura de árbol de decisión binario para predecir la presencia de catarata. Los nodos internos incluyen parte de las observaciones, y a su vez son divididos en otros nodos hijos siguiendo alguna condición de partición. Las hojas del árbol representan la clasificación final o clase.

El criterio aplicado para llevar a cabo la partición en los nodos internos del árbol debe cumplir dos condiciones:

- Permitir realizar una buena elección de la partición.
- Ser abordable en un tiempo computacional aceptable.

Existen numerosos criterios de partición, tales como el criterio de error esperado, el criterio Gini (94), los criterios de entropía o “Gain” (95), etc. La mayoría estos criterios están basados en las frecuencias relativas de las clases en cada uno de los hijos de la partición respecto a las frecuencias relativas de las clases en el nodo padre. El concepto en el que se fundamenta es la denominada impureza o entropía dentro de cada nodo. De manera intuitiva, podemos asimilar la impureza de un nodo a la homogeneidad o heterogeneidad de las observaciones que contiene. La partición óptima de un nodo interno sería aquella en la que cada nodo hijo contuviera elementos u observaciones de una única clase, es decir, totalmente homogéneos o puros.

Los árboles de decisión son modelos predictivos con una alta comprensibilidad. Su representación gráfica es fácilmente inteligible y legible para el usuario. No obstante, presenta dos grandes debilidades:

- Tiene una alta variabilidad en función de los datos de entrada al algoritmo.
Los árboles de decisión son muy sensibles al conjunto de entrenamiento.
- Es frecuente el sobreajuste a los datos.

El sobreajuste a los datos es un problema común, en mayor o menor grado, a todas las técnicas de minería de datos. El modelo se crea al ser entrenado sobre una base de datos, ajustando su estructura y parámetros para conseguir la predicción más exacta sobre esta base de datos. Pero el propósito de minería de datos no es crear modelos precisos sobre los datos presentados al sistema, sino que busca que estos sean generalizables sobre observaciones nuevas que no han sido presentados al sistema previamente. Predecir sobre datos ya conocidos no tiene valor añadido, el modelo debe ser capaz de predecir la clase correctamente en nuevas observaciones.

Este sobreajuste es más frecuente y de mayor magnitud conforme aumenta la complejidad del modelo. De ahí la importancia de su control mediante diversas técnicas, a fin de que el modelo tenga el mayor grado de generalización posible.

Las técnicas que se emplean para controlar el sobreajuste del modelo incluyen:

- Preprocesamiento de los datos de entrada al sistema, evitando los ejemplos fuera de un rango establecido y manejando los datos perdidos o vacíos.
- Ajuste de las variables de entradas al sistema, eliminando aquellas de dudosa influencia en el resultado del sistema.
- Poda de las ramas del árbol de decisión. Se define un límite inferior en la mejora del grado de impureza de los nodos hijos respecto al nodo padre, de modo que si la mejora es inferior al límite establecido no se lleva a cabo la partición del nodo.
- Validación y estimación de la precisión del modelo mediante un conjunto de ejemplos o datos que no habían sido presentados previamente al sistema (conjunto de datos de test).

3.6.2.2. COMBINACIONES DE ARBOLES DE DECISION.

Los árboles de decisión individualmente considerados son clasificadores débiles, entendiendo como tales a clasificadores con un alto error en la predicción de la clase de las instancias y con una alta variabilidad en función del conjunto de datos de entrenamiento empleado.

Una alternativa para solucionar esta debilidad es fusionar varios árboles en un único clasificador. La idea subyacente es combinar las predicciones de un conjunto de árboles que han sido entrenados sobre una misma base de datos de entrenamiento, aplicando algún tipo de modificación a fin de que exista cierto grado de variabilidad entre los árboles creados..

Para clasificar una nueva instancia, cada uno de los árboles emitiría una predicción, combinándose éstas para generar una salida unificada en función de qué predicción es

mayoritaria (sistema de votación simple) o qué predicción sale ganadora de una votación ponderada.

Los métodos de combinación de modelos son diversos en función de los clasificadores que utiliza y el método utilizado para asociarlos. En estas posibles combinaciones, destacamos dos metodologías:

- Técnica de ensacado o Bagging
- Técnica de potenciación o Boosting.

La técnica de Bagging propuesta por Breimann et al (96) genera subconjuntos de entrenamiento seleccionando aleatoriamente y con reemplazamiento en una muestra de x ejemplos del conjunto original de datos de entrenamiento formado por x ejemplos. La alta sensibilidad que presentan los árboles de decisión al conjunto de entrenamiento hace que los árboles generados con cada subconjunto difieran significativamente. La predicción de nuevos ejemplos se efectúa por votación mayoritaria (83). La figura 12 describe la combinación de árboles usando Bagging. Esta técnica es empleada por el algoritmo Random Forest.

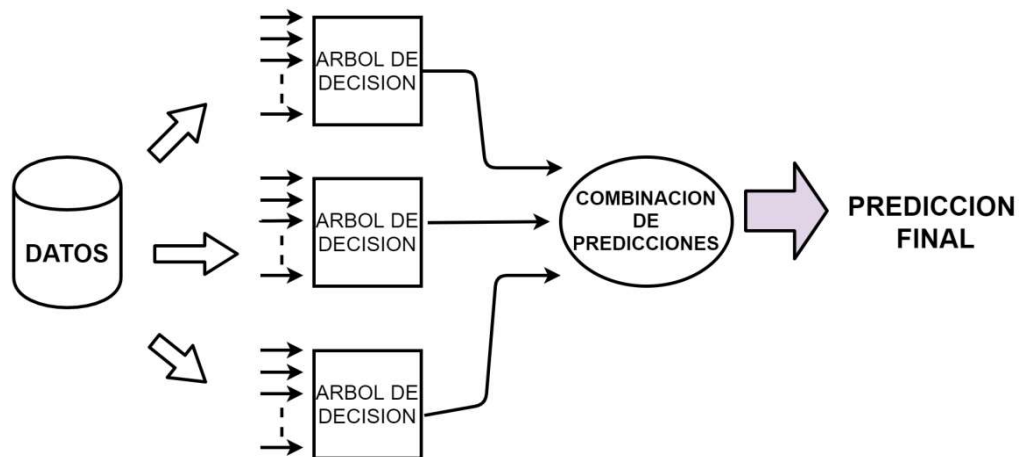


Figura 12: Representación del funcionamiento de la técnica de Bagging para combinar árboles de decisión.

El algoritmo de Boosting o de potenciación de gradiente descrito por Friedman et al (97) es más complicado de implementar que el anterior. La secuencia del algoritmo es la siguiente (92):

1. Se crea un clasificador inicial que es entrenado con los datos de entrada.
2. A continuación se forma un nuevo conjunto de datos de entrenamiento en el que los ejemplos mal clasificados en el primer paso tienen más peso que aquellos que fueron clasificados correctamente.
3. Creación secuencial de sucesivos clasificadores que se especializan en los ejemplos mal clasificados previamente.
4. Una vez construidos cientos de estos clasificadores iterando los pasos 1 a 3, la clase de un ejemplo se calcula como el voto ponderado de las clases que generan como salida todos los clasificadores.

Se ha confirmado que la combinación de clasificadores suele proporcionar predicciones más ajustadas que la de los clasificadores individuales (98). Breiman ha demostrado que el error de clasificación cometido por la combinación de árboles converge hacia un mínimo conforme se incrementa el número de árboles (99).

Dentro de los modelos de clasificación, la combinación de árboles de decisión mediante boosting ha mostrado dar excelentes resultados en multitud de situaciones y estudios (100).

La metodología de potenciación del gradiente extremo o “eXtreme Gradient Boosting” (XGB) descrita por Chen et al. (101) es un algoritmo basado en la combinación de modelos siguiendo una estrategia de Boosting. XGB está optimizado para ser altamente eficiente, flexible y escalable. Estas características, especialmente su alto rendimiento para manejar billones de datos en un tiempo reducido y con recursos computacionales limitados, ha hecho que actualmente sea el algoritmo de Boosting más ampliamente utilizado.

La combinación de clasificadores débiles ha permitido obtener modelos de clasificación con una tasa de clasificación correcta mucho más alta y con menos variabilidad según los datos de entrada. Pero, como contrapartida, ha dado lugar a que la interpretación del

modelo sea mucho más complicada o incluso imposible. El aumento de la complejidad lleva consigo una pérdida en la transparencia e interpretabilidad del modelo, como se observa en la Figura 13. El algoritmo XGB, no obstante, puede obtener una jerarquía de importancia de los atributos o características de las entradas al sistema basándose en la frecuencia con la que una característica es utilizada para realizar la partición de las observaciones en los nodos intermedios de los árboles creados o de acuerdo a la mejora de la entropía observada al hacer la partición en función de ese atributo. Eso hace que el modelo XGB sea óptimo en cuanto que presenta una alta precisión con un cierto grado de interpretabilidad.

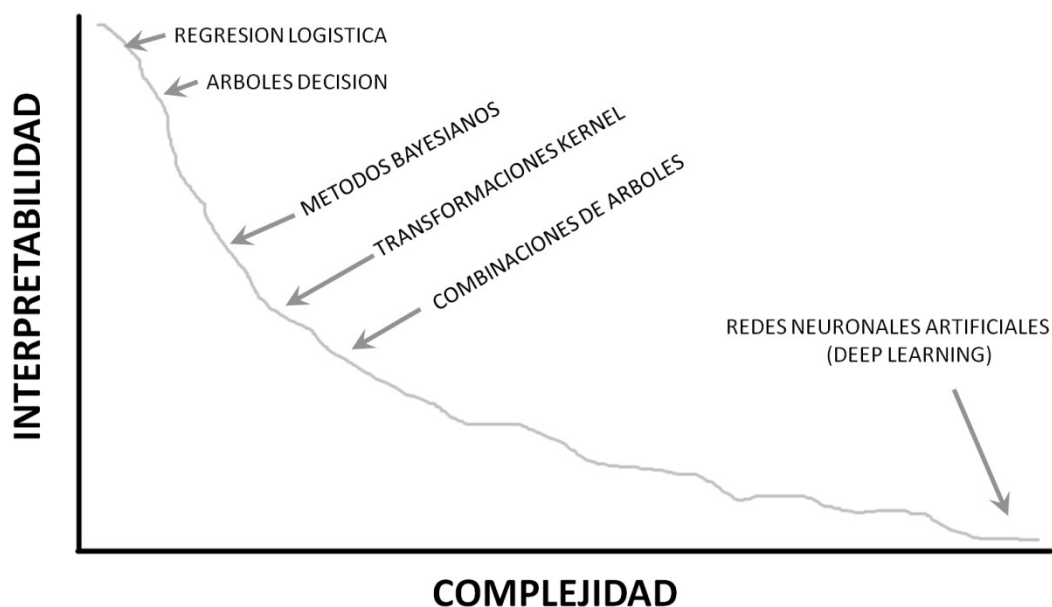


Figura 13. Relación entre complejidad e interpretabilidad del modelo.

3.6.2.3. REDES NEURONALES ARTIFICIALES Y MODELOS DE APRENDIZAJE PROFUNDO (“DEEP LEARNING”)

Las redes neuronales artificiales (RNA) constituyen un método de aprendizaje cuya finalidad inicial era la de emular el comportamiento de las neuronas, como procesadores biológicos de información. El funcionamiento de una neurona biológica puede ser modelado matemáticamente de la manera que se muestra en la figura 14. En el esquema

se observa una serie de señales de entrada a una neurona principal, equivalentes a las sinapsis de las terminaciones axónicas con las dendritas. Estas entradas x_i tienen un efecto sumatorio sobre la neurona según sus pesos sinápticos o w_i . De acuerdo a la suma de las entradas y a una función de activación predeterminada, la neurona emitirá o no una descarga o salida. La respuesta del sistema está asociada a la activación o no activación de la neurona principal.

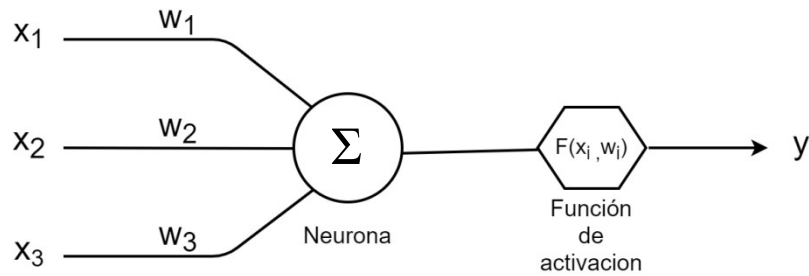


Figura 14. Diagrama de una neurona artificial. x_i : entrada al sistema; w_i : pesos de las entradas; y : salida del sistema.

Una única neurona artificial es una entidad de procesamiento de información demasiado sencilla como para ser de utilidad en procedimientos de clasificación, La potencia de la RNA proviene de la posibilidad de unir multiples neuronas artificiales sencillas para actuar conjuntamente en paralelo, constituyendo una red neuronal artificial. El modelo constaría de una serie de capas de neuronas conectadas entre ellas. La entrada al sistema accedería a la red desde un extremo y se propaga a través de la red, modificándose en cada capa que atraviesa, hasta que la activación alcanza la capa de salida (figura 15).

Al igual que el resto de modelos de aprendizaje automático, las redes neuronales artificiales son entrenadas con un conjunto de ejemplos o datos de entrada. El aprendizaje de la red se lleva a cabo mediante un ajuste de los pesos sinápticos w_i de todas las capas de la red neuronal para obtener el menor error de clasificación posible. La técnica empleada se describe como retropropagación del error, actualizando el error entre una capa a otra y modificando los valores w_i hasta que se consigue optimizar el sistema.

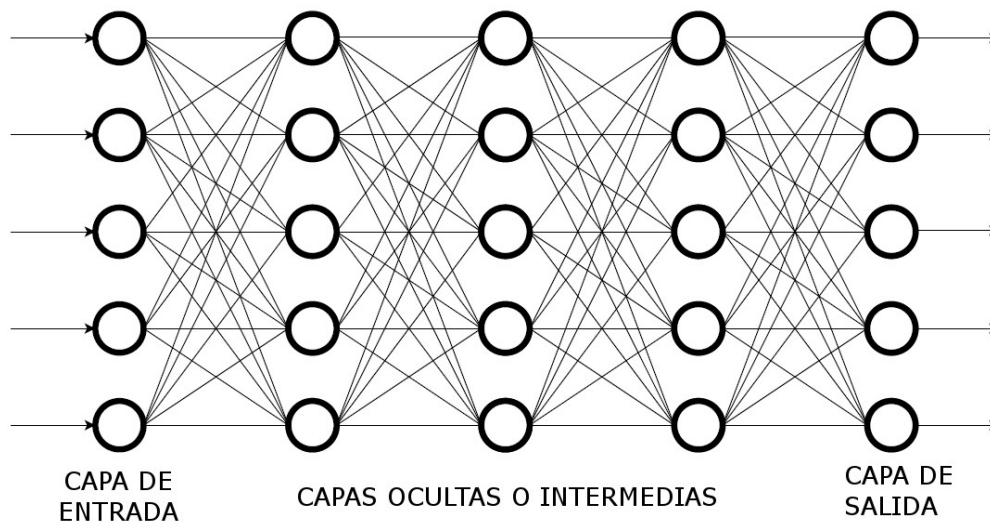


Figura 15. Esquema de una red neuronal artificial con cinco capas de neuronas.

Es posible crear una red neuronal artificial con cualquier número de capas intermedias y con cualquier cantidad de neuronas dentro de cada capa. Conforme aumentamos el número de capas, se incrementa la capacidad del modelo para clasificar correctamente los objetos presentados al sistema, llegando a tener una precisión más alta que cualquier otro modelo de clasificación. Pero igualmente se generan una serie de problemas tales como:

- El coste computacional se eleva enormemente, pudiendo ser inabordable para una computadora convencional.
- El sobreajuste a los datos se incrementa, perdiendo la capacidad de generalización.
- La complejidad de modelo hace imposible conocer cuales son los procesos exactos que tienen lugar entre la entrada y la salida al sistema. El modelo es totalmente opaco en su funcionamiento interno y absolutamente ininteligible.

Un paso más allá de las redes neuronales artificiales, aunque basado en ellas, surge el llamado aprendizaje profundo o “deep learning”. El aprendizaje profundo es un caso particular de las redes neuronales artificiales con la característica de tener múltiples capas intermedias de neuronas. La idea principal del aprendizaje profundo es que las

capas intermedias realizan un procesamiento de los datos de modo jerárquico, obteniendo representaciones de los datos y extrayendo características de un nivel cada vez más alto hasta llegar a la salida del sistema. En el aprendizaje profundo, cada capa aprende una serie de patrones de mayor complejidad según la profundidad de esta.

Las redes convolucionales se engloban en el epígrafe del aprendizaje profundo, añadiendo una serie de capas convolucionales para realizar labores de preprocesamiento de los datos de entrada (figura 16). Su utilización más extendida es el análisis y clasificación de imágenes.

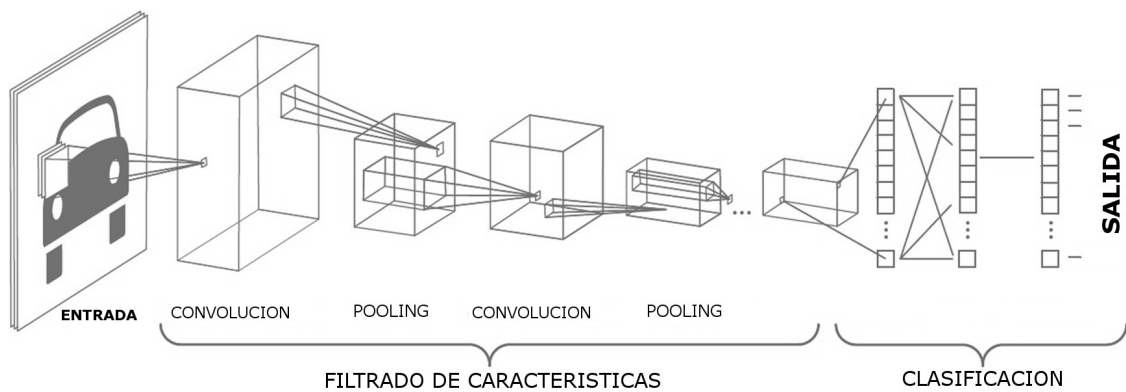


Figura 16. Esquema de una red convolucional.

Tanto las redes neuronales artificiales como su variante de aprendizaje profundo se adscriben en su comportamiento a las denominadas “cajas negras” o “black boxes”. Este término hace referencia a su incapacidad para ser interpretados de manera inteligible para el ser humano. La extrema complejidad de la red hace imposible conocer cual son los patrones o pautas intermedias que llevan al sistema a emitir su respuesta. Como solución intermedia al problema de esta opacidad del sistema, han surgido una serie de métodos que permiten extraer cierta información de lo que acontece en las capas centrales de la red neuronal, obteniendo algunos datos intermedios del procesamiento de la información denominados mapas de activación.

3.7. INTERPRETABILIDAD DE LOS MODELOS DE CLASIFICACION.

La interpretabilidad de un modelo ha sido definida por Miller como el grado en el que el humano puede entender la causa de un resultado del modelo (102). Una mayor interpretabilidad de un modelo se correlaciona con una mayor facilidad para comprender por qué el modelo emite un resultado determinado.

Un modo de hacer que un sistema de aprendizaje automático sea interpretable es utilizar alguno de los modelos que son intrínsecamente explicables. Dentro de estos se encuentran:

- Modelos de regresión lineal.
- Regresión logística
- Árboles de decisión.

Otra opción consiste en utilizar herramientas de interpretación independientes del modelo. En este caso, se construye un modelo no interpretable, por ej. una combinación de árboles o una red neuronal, aplicando después metodologías tales como los gráficos de dependencia o la estimación de la importancia de los atributos en un intento de dilucidar, aunque sea de forma parcial e incompleta, el funcionamiento interno del modelo. Estas herramientas de interpretación están basadas en la medida de los cambios que ocurren en la salida del sistema al modificar su entrada.

A modo de ejemplo, en las redes convolucionales para la clasificación de imágenes, es conocido que el algoritmo asimila la detección de los bordes de los objetos de la imagen y realiza su filtrado en las capas más bajas de la red. Esto es una comprensión de cómo funciona el algoritmo, pero no una interpretación del funcionamiento de un modelo entrenado con una base de datos concreta. Y, sobre todo, es incapaz de dilucidar cómo se ha llegado a una predicción específica de una imagen.

Pero, si un sistema de aprendizaje automático funciona adecuadamente dando predicciones acertadas, ¿por qué no confiar ciegamente en el modelo y obviar el razonamiento interno para llevar a cabo la predicción?. La respuesta puede argumentarse en una serie de razonamientos (103):

- Una simple métrica del funcionamiento del modelo, tal como la precisión o grado de clasificación correcta, es una descripción incompleta y deficiente para los problemas del mundo real
- Algunos modelos quizás no necesiten una explicación de su funcionamiento al ser utilizados en ambientes de bajo riesgo, de manera que un error en su predicción no tiene consecuencias severas (por ej. sistemas de recomendación de películas). Sin embargo, en el ámbito de la medicina, los errores no son permisibles, en tanto que una decisión inadecuada de un sistema experto puede conllevar incluso la muerte de un paciente. La interpretación del sistema nos proporciona una información adicional que puede ser contrastada con el conocimiento médico previo obtenido por métodos de investigación distintos al aprendizaje automático.
- Curiosidad humana por la comprensión del mundo. Cada persona tiene un esquema mental de su entorno, que se actualiza cuando acontece algo inesperado, buscando una explicación de ese evento imprevisto. Si alguien se siente enfermo, buscará el origen de su malestar intentando buscar una comprensión de su origen
- Vinculado a la curiosidad humana, el hombre tiene el deseo permanente de encontrar un significado al funcionamiento del mundo que le rodea. Cualquier persona desea reconciliar contradicciones o inconsistencias entre los elementos de nuestras estructuras mentales.

Para intentar tener cierto conocimiento del funcionamiento de estos modelos de “caja negra” se han implementado unos instrumentos metodológicos que proporcionan un cierto grado de explicación de los procesos internos del algoritmo. La utilización de estas herramientas de interpretación externas, también llamadas agnósticas, aplicadas a un modelo opaco tiene una serie de beneficios, destacando los siguientes (104):

- Flexibilidad en el uso de los modelos de aprendizaje. El desarrollador puede seleccionar el modelo que crea más adecuado para el dominio de conocimiento al que lo aplica, aislándolo de la posible interpretación del mismo.

- Flexibilidad en la estructura de la interpretación así como en su representación. El sistema de explicación del modelo no tiene que usar el mismo formato de las características que utiliza el modelo.

Entre estos instrumentos agnósticos de interpretación del modelo se encuentran la determinación de la importancia relativa de los atributos y el uso de los valores Shapley como estimación del impacto de cada atributo en el resultado del sistema.

El cálculo de la importancia relativa de los atributos está fundamentado en dos principios de la estructura de los árboles de decisión:

- Los atributos que son testados para generar la partición de los nodos tienen mayor importancia relativa a mayor cercanía al nodo raíz. El atributo más importante según este criterio sería aquel que divide la base de datos original, ubicado en el nodo raíz del árbol.
- Un atributo es más importante cuanto más puro son los nodos hijos que genera, concibiendo como grado de pureza al concepto asociado a la homogeneidad de las clases de los elementos incluidos en nodo.

En las combinaciones de árboles de decisión, tal como es el caso del clasificador XGB, la importancia relativa de cada parámetro se establece según estos criterios, haciendo una medida ponderada de todos los árboles que se fusionan para constituir el clasificador final.

Los valores Shapley o SHAP (Explicaciones Aditivas Shapley o “SHapley Additive ExPlanations”) son otra propuesta de herramienta agnóstica para interpretar los modelos opacos introducida en el año 2018 por Lundberg et al (105). Su origen está en la aplicación de la teoría de juegos cooperativos en el análisis de los modelos de clasificación. El objetivo de los valores Shapley es explicar la predicción para una instancia a través de la contribución de sus atributos en la predicción. El valor Shapley asociado con cada parámetro se correlaciona directamente con el impacto en el resultado del predictor.

Para entender de manera muy sencilla cómo se calculan estos valores podemos imaginar a un grupo de personas, que podemos describir como A, B y C, que colaboran de modo conjunto para crear una empresa con una rentabilidad o beneficio. Cada persona contribuye y se involucra en mayor o menor grado en el negocio, por lo que el reparto justo del beneficio entre los socios de la empresa debería ser directamente relacionado al esfuerzo o trabajo que han dedicado a la empresa. Para saber la contribución de cada socio, lo que haremos es ir intercalando a los participantes en las posibles permutaciones según su “orden de colaboración”, digamos ABC, ACB, BCA, etc, e ir calculando la contribución marginal de cada participante en cada vuelta. El promedio de esta contribución marginal en todas las permutaciones posibles es el valor Shapley, indicando cuánto contribuyó cada persona a conseguir el resultado.

Esta forma de entender los juegos cooperativos puede ser trasladada a los modelos de clasificación, donde cada “socio” sería un atributo de la observación y cada valor Shapley indicaría el peso o impacto que ese atributo tiene en el resultado del modelo. El valor SHAP calcula la importancia de un atributo comparando lo que el modelo predice con y sin ese atributo.

Los valores Shapley son utilizados para explicar un modelo ya creado. La estructura de su funcionamiento se reproduce en la figura 17. Los datos son la entrada para la construcción del modelo, que practica la predicción sobre las observaciones nuevas. A partir de este modelo y de los datos originales se calculan los valores Shapley, haciendo una interpretación del modelo mediante su análisis.

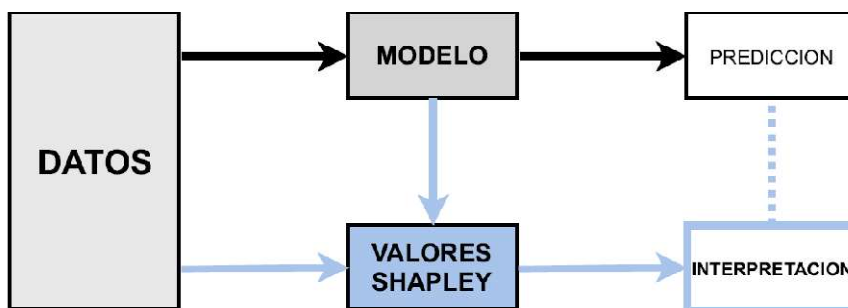


Figura 17. Esquema de la relación entre datos, modelo y valores Shapley. Fuente: Adaptado de SHAP github (106).

En la figura 18 se muestra el funcionamiento de los valores SHAP en un modelo en el que los parámetros de la CFNR son los atributos de las observaciones y el resultado del modelo es el diagnóstico de glaucoma. Cada valor Shapley “empuja” el resultado con una amplitud y una dirección. Valores Shapley positivos desplazan la predicción en el sentido la predicción de glaucoma. Por el contrario, valores negativos favorecen que la salida del modelo sea de ausencia de glaucoma.

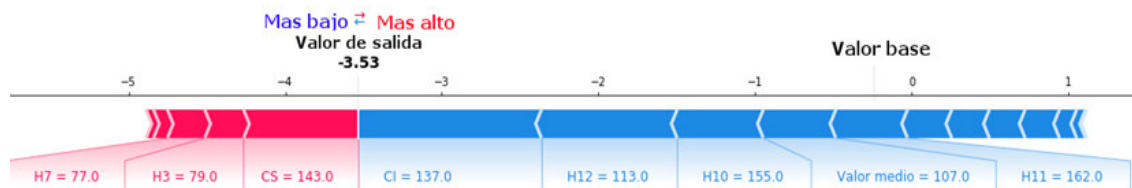


Figura 18. Representación de la contribución de cada parámetro de la CFNR en la predicción del modelo en función de su valor Shapley.

Una característica señalada de los valores Shapley es su consistencia, en contraposición a la inconsistencia en el cálculo de la importancia relativa de los atributos en la estructura de los árboles de decisión. Esta consistencia establece que si el impacto de un atributo en un modelo aumenta o se mantiene igual, su valor Shapley nunca puede reducirse (107,108).

3.8. LA MINERÍA DE DATOS EN OFTALMOLOGIA.

La aplicación de la tecnología de aprendizaje automático y de minería de datos en el estudio de la patología ocular ha tenido su mayor progreso en la últimos años, aunque su metodología era bien conocida durante décadas.

Prácticamente todas las subespecialidades de la oftalmología han recurrido a la minería de datos para la resolución de los distintos problemas planteados.

En el año 2006, Marsolo (109) recurre a los árboles de decisión para diferenciar entre personas con corneas normales, sujetos con miopía post-LASIK y pacientes con queratocono, a partir de los valores de los coeficientes de los polinomios de pseudo-Zernike usados como representación matemática de la morfología de la córnea. La

precisión que alcanza el modelo de clasificación en su estudio oscila entre un 85 % utilizando un árbol estándar, y más del 90% al combinarlo con otras técnicas de meta-aprendizaje. En el análisis comparativo con otras técnicas de clasificación automática, los árboles de decisión presentaron un valor mayor de fiabilidad, velocidad, escalabilidad, interpretabilidad y estabilidad que otras técnicas (Tabla 5)

Tabla 5

Comparación de métodos de clasificación según distintos criterios.

Criterio	Arbol de decisión	Red Neuronal	Naïve-Bayes
Precisión	+	±	±
Velocidad	+	--	+
Escalabilidad	+	--	+
Interpretabilidad	++	--	-
Estabilidad	++	-	-

Schmidt desarrolla un modelo pronóstico basado en un árbol de clasificación para predecir la visión residual tras un traumatismo perforante del globo ocular (110). Toma como datos para crear el modelo los valores de 14 predictores. El modelo final presentó una precisión predictiva excelente, constituyendo una herramienta de gran utilidad para la toma de decisiones relacionadas con el manejo de los traumatismos abiertos del globo ocular.

La adecuación en la indicación de cirugía de catarata ha sido valorada por Quintana et al. mediante la creación de un árbol de decisión (111). Toma para ello los datos sociodemográficos y clínicos, incluyendo la agudeza visual y el índice de función visual determinado con el test VF-14, en una población con un seguimiento prospectivo antes y después de la cirugía de catarata. Los resultados finales se clasificaron entre indicación apropiada o inapropiada, comparándolos con un estándar de referencia de

ganancia de agudeza visual y en el test VF-14. La conclusión del estudio es que el árbol de decisión generado puede ser de ayuda en la identificación del grado de adecuación de la cirugía de catarata.

Egejuru et al. (112) hace uso de diversos algoritmos de aprendizaje automático para identificar los factores de riesgo para predecir la presencia de catarata. Utiliza árboles de decisión, método de Naïve Bayes y redes neuronales artificiales, encontrando que tabaco, alta miopía, uso de lentes de contacto y consumo frecuente de alcohol eran los factores de riesgo más relevantes que auguraban la presencia de catarata.

La investigación de los predictores de un resultado visual desfavorable en la queratoconjuntivitis vernal ha sido llevado a cabo por Sacchetti (113) empleando árboles de regresión y clasificación, proponiendo un nuevo sistema de gradación según los hallazgos clínicos. El estudio encuentra que un número elevado de recaídas y un estadio mayor de la enfermedad en su estado basal fueron los principales predictores de un peor resultado visual final.

El diagnóstico de la retinopatía diabética (RD) es una de las aplicaciones más avanzadas en el campo del aprendizaje automático. Ting et al. han estudiado el rendimiento de las redes neuronales artificiales para la detección de la RD, obteniendo valores de AROC de 0.958 en el sistema que implementaron. Otros grupos de trabajo (114,115) han trabajaron con imágenes de fondo de ojo para el diagnóstico de la RD mediante un sistema de aprendizaje profundo, discriminando entre RD de base, RD que debe ser derivada a especialista y RD proliferativa. Los valores de AROC calculados son superiores a 0.95. Por ello los autores llegan a proponer que estos sistemas de diagnóstico automático basado en aprendizaje profundo podrían introducirse con seguridad en los programas de detección de RD reemplazando al médico, especialmente en países en vías de desarrollo o subdesarrollados.

Otra patología retiniana con amplia aplicación de técnicas de aprendizaje profundo para su detección a partir de imágenes de fondo de ojo ha sido la degeneración macular asociada a la edad (DMAE). Burlina et al (116) desarrolla un sistema de gradación automática para detectar la DMAE. Schlegl et al (117) usa redes convolucionales para detectar y cuantificar la presencia de fluido en área macular. Peng et al (118) analiza la

severidad y el riesgo de progresión de la DMAE avanzada a partir de fotos color de la mácula.

Uno de los campos en los que la aplicación del aprendizaje automático ha sido más extensa ha sido en el diagnóstico del glaucoma. La explosión de valores numéricos y parámetros obtenidos con el estudio del nervio óptico y la CFNR empleando las técnicas de imagen disponibles en la actualidad ha posibilitado una amplia aplicación de métodos de minería de datos en el estudio del glaucoma.

Se considera que los métodos de aprendizaje automático tiene un alto potencial para la detección precoz, diagnóstico y clasificación del glaucoma. Estos métodos también nos proveen de la posibilidad de identificar nuevos factores de riesgo o patrones que pueden ser de utilizad en diagnóstico, progresión o tratamiento de esta patología.

Los primeros estudios de aplicación de aprendizaje automático en glaucoma empezaron a surgir a partir de principios de la década 2000 a 2010. En general se trata de análisis de comparación del rendimiento de distintos modelos de aprendizaje automático en el diagnóstico del glaucoma, tanto supervisados (regresión lineal, análisis lineal discriminante, árboles de decisión, sistemas de vectores soporte, Random-Forest, etc) como no supervisados (análisis de agrupación, análisis de componentes principales, etc). Tanto el daño estructural, con la afectación del nervio óptico y la CFNR, como el daño funcional en el campo visual son susceptibles de ser analizados con estos algoritmos.

En la tabla 6 se presenta un resumen de los estudios de aplicación de la inteligencia artificial para el diagnóstico de glaucoma empleando como datos de entrada al sistema los parámetros o las imágenes obtenidas con OCT. Los estudios son muy dispares, aplicando cada uno de ellos un diseño diferente, un planteamiento distinto, así como métodos de aprendizaje automático, datos de entrada, técnica de validación y métricas de rendimiento del modelo diversas en cada caso. En la tabla 6 se recopilan el diseño y los detalles más importantes de estos trabajos, pero debemos tener presente la complicación que supone la heterogeneidad de estos estudios a la hora de comparar sus resultados.

Tabla 6

Resumen de estudios de aplicación de aprendizaje automático en el diagnóstico del glaucoma usando datos o imágenes de la OCT

Autores	Año	Método de validación	Tipo de algoritmo	Datos de entrada	AROC	Sensibilidad	Especificidad	Comentarios
Huang et al. (119)	2005	Validación cruzada 10 pliegues	ALD, RNA	Parámetros de la OCT	0.821–0.991	72%–100%	80%	–
Burgansky-Eliash et al. (120)	2005	Validación cruzada 6 pliegues	ALD, AD, MVS	Parámetros de la OCT	0.981	92.50%	95%	–
Naithai et al. (121)	2007	Validación cruzada	AD, RF, RNA	Parámetros de la OCT	0.93 -0.98			Mejor resultado con RF
Bizios et al.(122)	2010	Validación cruzada 6 pliegues	MVS,RNA	Parámetros de la OCT	0,98			El tipo de datos de entrada tiene mayor repercusión que el tipo de clasificador.
An et al. (123)	2018	Validación cruzada 10 pliegues	AP	Parámetros de la OCT	0.94	–	–	Combina RNA con RF

Autores	Año	Método de validación	Tipo de algoritmo	Datos de entrada	AROC	Sensibilidad	Especificidad	Comentarios
Kim et al. (124)	2017	Grupo de entrenamiento / Grupo test	RF, MVS,KM	Parámetros de la OCT y CV	0.979	98.3 %	97.5%	Mejor rendimiento con RF
Barella et al (125).	2013	Validación cruzada 10 pliegues	AD, MVS, RF,RNA,	Parámetros de la OCT	0.877	64.9%	80%	-
Christopher et al (126)	2018	Grupo de entrenamiento / Grupo test	ANS	Parámetros de la OCT	0.95	-	-	-
Shigueoka et al (127)	2018	No descrito	AD,MVS,	Parámetros de la OCT y CV	0.80-0.93	-	.-	
Sugimoto et al (128)	2013	No descrito	AD, RF	Parámetros de la OCT	0.75-0.90			Análisis progresión de glaucoma. Mejor resultado con RF
Ting et al (129)	2017	Grupo de entrenamiento / Grupo test	AP	Fotografía de fondo de ojo	0.942	96.40%	87.2%	-
Medeiros et al.(130)	2019	Grupo de entrenamiento / Grupo test	AP	Fotografía de fondo de ojo	0.944	90%	80%	

Autores	Año	Método de validación	Tipo de algoritmo	Datos de entrada	AROC	Sensibilidad	Especificidad	Comentarios
Maetschke et al.(131)	2019	Grupo de entrenamiento / Grupo test	MVS, RF, NB, GB,AP	Imágenes de OCT	0.94	–	–	Interpretación con mapas de activación
Zhang et al. (132)	2019	Grupo de entrenamiento / Grupo test	AP	Imágenes de OCT	0.90	–	–	–
Peiyu et al. (133)	2019	Validacion cruzada 5 pliegues	MVS, KM,RC	Imágenes de OCT	0.90-0.92	81 – 86%	88 -93%	
Hemelings et al. (134)	2020	No descrito	AP	Fotografía de fondo de ojo	0.995	98%	91%	

ANS: Análisis no supervisado. ALD: Análisis lineal discriminante: AP: Aprendizaje profundo (“Deep learning”); MVS: Máquina de vectores soporte; KM: K-medias; RF: Random-Forest; RNA: Redes neuronales artificiales.; AD: Arbol de decisión; NB: Naïve-Bayes; GB: Potenciación de gradiente (Gradient Boosting)

Hothorn et al. (135) desarrolla el problema de la clasificación entre los ojos con glaucoma o normales mediante la estrategia de un clasificador basado en 62 variables obtenidas de los datos clínicos y las imágenes de la cabeza del nervio óptico adquiridas con el láser confocal de barrido. Utiliza y compara tres modelos: análisis discriminante, árboles de clasificación y técnica de bagging aplicada a árboles de clasificación. El estudio concluye que la técnica de “bagging” aplicada a árboles de clasificación mejora el rendimiento de los árboles sencillos, siendo a su vez este método superior al análisis discriminante.

En el año 2007, Naithani (121) conduce un trabajo en el que compara el rendimiento para la detección del daño glaucomatoso de un conjunto de clasificadores automáticos basados en la OCT con los incluidos en el paquete estadístico del HRT II. Los valores de la OCT que utiliza como datos de entrada a los clasificadores son los parámetros de la CFNR y de la cabeza del nervio óptico. Los clasificadores analizados son la función lineal discriminante, las redes neurales artificiales y los árboles de clasificación y regresión. Las funciones incluidas en el HRT II con las que practica la comparación son el análisis de regresión de Moorfield y las funciones dicriminantes lineales desarrolladas por Mikelberg et al (136) y Bathija et al (137). Al confrontar los valores del área bajo la curva ROC de cada clasificador observa que todos los clasificadores basados en la OCT muestran un rendimiento superior que las fórmulas del HRT para diferenciar a los pacientes con glaucoma inicial o moderado de sujetos normales. Destaca, además, que, si bien la función discriminante fue la que mostró un mejor comportamiento en la labor de clasificación, el análisis mediante la estructura de árbol aportó una estructura fácilmente comprensible y con el menor número de variables, ya que únicamente incluía el valor de la E/D vertical y el grosor medio de la CFNR peripapilar.

Ya en el año 2013, Barella et al (125) estudia el rendimiento de diez algoritmos de aprendizaje automático para discriminar pacientes con glaucoma precoz utilizando entre otros los parámetros de grosor de la CFNR. Sus resultados indican que el mejor clasificador es Random-Forest, una combinación de árboles de decisión mediante la técnica de Bagging. El rendimiento de clasificación que reporta es de un AROC de 0.805.

Bizios et al.(122) comparan los modelos de vectores de soporte y redes neuronales en el diagnóstico del glaucoma. Los valores de AROC que alcanzan son de 0.982 para la máquina de vectores de soporte y 0.982 para las redes neuronales artificiales.

En el estudio de Huang et al (119), el algoritmo de aprendizaje basado en gradiente descendiente alcanzó una capacidad diagnóstica con AROC de 0.74. El mejor clasificador presentado en el trabajo de Burgansky-Eliash et al. (2005) consiguió un AROC de 0.987.

Vidotti et al. (138) analiza el AROC de múltiples modelos de clasificación basados únicamente en los grosores de la CFNR en el glaucoma precoz y moderado, hallando valores de AROC que oscilan entre 0.818 para el modelo basado en Bagging y 0.785 en el modelo basado en Boosting.

Como vemos, en la primera mitad de la década de 2010 a 2020, la mayor parte de los estudios tomaban como entrada al sistema de clasificación los parámetros que proporcionaba la OCT en su estudio del nervio óptico y la capa de fibras nerviosas peripapilar. Al avanzar a la década posterior a 2010, con el advenimiento de una mayor capacidad computacional y de bancos de datos de mayor tamaño, así como el desarrollo teórico de las redes convolucionales, los estudios publicados pasaron a ser en su mayoría centrados en el análisis de imágenes o fotografías directas del nervio óptico con técnicas de aprendizaje profundo, utilizando como entrada al clasificador los píxeles de las imágenes en color del área peripapilar.

Los trabajos que utilizan métodos de aprendizaje profundo con imágenes de fondo de ojo consiguen, en general, mejores resultados de clasificación que aquellos basados en los parámetros de la OCT. Esto es debido, en gran parte, a que los bancos de datos utilizados para su entrenamiento son de dimensiones mucho mayores, y con una tendencia a ser cada vez amplios conforme se incrementa la capacidad de almacenamiento y de los sistemas de manipulación de las bases de datos. La faceta negativa de estos métodos es que, si bien realizan una clasificación casi perfecta, desconocemos completamente cómo se construye la predicción al carecer de la característica de la interpretabilidad.

JUSTIFICACION

La prevención de la pérdida de la función visual como consecuencia del glaucoma es una de las principales metas de la oftalmología. Diversos estudios poblacionales estipulan que aproximadamente un 4 % de la población general padecerá la enfermedad a los 70 años de edad, aumentando esta prevalencia conforme se incrementa la longevidad.

En su práctica médica, el oftalmólogo se enfrenta a diario a la dificultad que supone el diagnóstico del glaucoma. Para ello dispone en la actualidad de una gran cantidad de valores numéricos y probabilísticos ofrecidos a partir del examen del paciente con las nuevas tecnologías de imagen. El profesional confronta esta información con los datos obtenidos en la exploración clínica para clasificar al paciente como sujeto sano o paciente con glaucoma. La integración de todos los datos por un oftalmólogo experto, que aplique su experiencia adquirida, es posiblemente el mejor método para el diagnóstico de la enfermedad en la actualidad.

No obstante, la detección de glaucoma en su fase inicial sigue constituyendo un reto para el oftalmólogo, dada la superposición de datos clínicos y estructurales que existe entre pacientes sanos y pacientes con glaucoma en estadio precoz. En el comienzo de la enfermedad ninguna de las exploraciones disponibles resulta definitiva para definir la enfermedad. La campimetría y la exploración de la papila y la capa de fibras nerviosas constituyen los pilares diagnósticos básicos, pero ninguno de ellos tiene la suficiente sensibilidad y/o especificidad para establecer la presencia de glaucoma en todos los casos.

La excepcional evolución tecnológica de la última década ha ido acompañada del desarrollo de técnicas de imagen dirigidas a la exploración de las diversas capas de la retina y de la cabeza del nervio óptico. La oftalmología se ha visto inundada con una enorme cantidad de valores numéricos que estiman los tamaños, grosores y volúmenes de las estructuras oculares afectadas en diversas patologías oculares, entre ellas en el glaucoma. La aparición de la tomografía de coherencia óptica (OCT) ha permitido determinar del grosor de la capa de fibras nerviosas de la retina, así como el cálculo de los datos topográficos de la cabeza del nervio óptico. Este cúmulo de información, no obstante, ha generado un nuevo problema: su interpretación.

Disponemos de un enorme volumen de información generada al aplicar las últimas técnicas de imagen en la exploración del paciente. Cuando examinamos estos datos, el análisis

practicado suele estar guiado por los resultados que esperamos encontrar. Habitualmente utilizamos algún paquete estadístico para localizar correlaciones entre variables, establecer medidas o varianzas e intentar modelar de esta forma la información recopilada.

Esta explosión de datos, sin embargo, no supone un aumento directo de nuestro conocimiento. Por otra parte, en esa montaña de cifras existe información que no puede ser encontrada con los procedimientos habituales de trabajo. La minería de datos nos ayuda a dar un paso más en este sentido, sacando a la luz relaciones, patrones o tendencias significativas escondidas entre los datos, que no son evidentes en una valoración más superficial, y que podrían ser de gran utilidad en el proceso de diagnóstico o toma de decisiones terapéuticas en nuestros pacientes.

El objetivo principal del presente estudio es la creación de un sistema de clasificación, basado en los valores numéricos obtenidos mediante el estudio con OCT de la capa de fibras nerviosas peripapilar, que sirva de apoyo para discernir entre pacientes con glaucoma en estadio inicial y pacientes sanos. Para ello, se recurrirá a la creación de un clasificador de alta complejidad, con excelentes resultados en sus predicciones pero con escasa interpretabilidad. A través del análisis de este clasificador inicial, podemos sacar una serie de conclusiones que serán aplicadas para obtener un clasificador con alta interpretabilidad como es el árbol de decisión.

Deseamos destacar que, dentro de este objetivo, la finalidad principal del presente estudio es la prevención del deterioro en la calidad de vida del paciente con glaucoma. La mayor eficacia del tratamiento en las etapas tempranas del glaucoma justifica todo el esfuerzo que pueda hacerse orientado a su diagnóstico precoz, en aras de mantener la calidad visual de nuestros pacientes.

OBJETIVOS

El presente estudio propone la creación de un modelo de clasificación para establecer el diagnóstico del glaucoma en fase precoz de la enfermedad apoyado únicamente en los datos de grosor de la CFNR peripapilar, haciendo uso de las herramientas que proporciona la minería de datos. Idealmente, este modelo debería tener una alta interpretabilidad directa, lo cual redundará en una aplicabilidad clínica inmediata.

Los objetivos concretos que se plantean pueden resumirse en los siguientes puntos:

1. Creación de una base de datos de pacientes con glaucoma en etapa temprana de la enfermedad y sujetos sanos, de tamaño suficientemente grande como para sustentar la aplicación de herramientas de minería de datos.
2. Construcción de un modelo de clasificación de alta complejidad y alta precisión para predecir la presencia de glaucoma precoz a partir de los parámetros de grosor de la CFNR peripapilar..
3. Explicar el modelo de alta complejidad mediante herramientas agnósticas de interpretación.
4. Construcción de un clasificador con alta interpretabilidad directa, basado en árboles de decisión, para la predicción de la enfermedad del glaucoma precoz sustentado únicamente en los parámetros de grosor de la CFNR. Este modelo debe cumplir las condiciones de tener un sobreajuste controlado, y ser fácilmente aplicable en la práctica clínica.

MATERIAL Y METODOS

1. CARACTERISTICAS GENERALES DEL ESTUDIO.

El trabajo desarrollado corresponde a un estudio de tipo observacional transversal de pacientes atendidos en el centro sanitario Hospital Universitario de Fuenlabrada durante el periodo de tiempo comprendido entre enero 2017 y diciembre 2019.

Todos los sujetos incluidos en este estudio procedían del Area Sanitaria 9 de la Comunidad de Madrid, siendo atendidos en las consultas de Oftalmología del Hospital Universitario de Fuenlabrada como primera visita o como visita sucesiva. Los pacientes con glaucoma fueron seleccionados de manera consecutiva entre aquellos atendidos en la consulta de la subespecialidad de glaucoma durante el periodo de reclutamiento y que presentaban un estadio inicial de la enfermedad. Los sujetos sanos, que constituyen el grupo control, se seleccionaron aleatoriamente entre aquellos pacientes que acudieron a la consulta de oftalmología por un motivo no relacionado con la enfermedad del glaucoma y que presentaron una exploración oftalmológica normal.

El estudio se ha realizado de acuerdo con los principios de la Declaración de Helsinki y ha sido aprobado por el Comité de Etica e Investigación del Hospital Universitario de Fuenlabrada (Anexo 1. Código Registro: APR 18/16). A todos los pacientes que participaron en el estudio se les entregó una hoja informativa y firmaron un consentimiento informado que explicaba la naturaleza y objetivo de la investigación (Anexo 2).

El número total de sujetos reclutados para el estudio fue de 328 en total, de los que 151 corresponden al grupo de pacientes con glaucoma en estadio precoz y los 177 restantes al grupo control de sujetos sanos.

2. PROTOCOLO DE EXPLORACION OFTALMOLOGICA.

A todos los pacientes susceptibles de ser incluidos en el estudio se les realizó una historia y exploración oftalmológica completa, que comprendía:

- Encuesta y resumen de antecedentes personales de enfermedades sistémicas, enfermedades oculares y cirugías oculares previas.
- Mejor agudeza visual corregida en escala decimal
- Refracción ocular en visión lejana y cercana.

- Biomicroscopía del segmento anterior.
- Estudio del ángulo mediante gonioscopia con lente de tres espejos de Goldman,
- Medida de presión intraocular mediante tonometría de aplanación con tonómetro de Goldman. Se realizaron 3 medidas consecutivas, tomando como valor válido la media de las 3 medidas realizadas.
- Estudio funduscópico con dilatación pupilar mediante colirio de tropicamida al 1%, realizando una valoración meticulosa del aspecto de la papila óptica mediante biomicroscopía con lente no contacto de +90 dioptrías.

Posteriormente, a los pacientes que cumplían los criterios de inclusión para algunos de los grupos se les realizaron las siguientes exploraciones:

- Perimetría automatizada con el algoritmo 24-2 del analizador de campo visual Humphrey II 740 (Humphrey Field Analyser, Zeiss/Humphrey Systems, Dublin, California, EEUU), utilizando un estímulo de tamaño III y estrategia de umbral SITA estándar (Swedish Interactive Threshold Algorithm).

Se realizaron al menos 3 campimetrías consecutivas en el espacio de tiempo máximo de 2 meses, a fin de determinar la reproducibilidad de los resultados del estudio del campo visual.

El estudio de campo visual se llevó a cabo con la corrección para visión cercana.

- Estudio de la cabeza del nervio óptico y capa de fibras peripapilares mediante Tomografía de Coherencia Óptica de dominio espectral (Topcon 3D OCT 2000 MarkII, FastMap v. 3.40, Topcon Corp., Japan). Los grosores de la capa de fibras peripapilares fueron obtenidos en la circunferencia estándar de 3,4 mm de diámetro centrada en el centro de la cabeza del nervio óptico.

El estudio mediante SD-OCT fue practicado tras dilatación pupilar con colirio de tropicamida al 1%.

Aquellos pacientes en los que los resultados de la OCT mostraban errores de segmentación o índice de intensidad de la señal (SSI) inferior a 45 fueron descartados.

Todos los exámenes de OCT practicados fueron revisados por 2 oftalmólogos con amplia experiencia en la enfermedad de glaucoma, y fueron descartados ante cualquier sospecha de estar descentrados o no ser completamente fiables.

Todas las exploraciones de campo visual y OCT fueron llevadas a cabo dentro de un lapso de tiempo máximo de 60 días, a fin de evitar posibles cambios del estadio del glaucoma, obteniendo de esto modo una imagen del estado de la enfermedad lo mas transversal posible.

3. SELECCION DE PACIENTES.

3.1. CRITERIOS DE INCLUSION.

3.1.1. GRUPO DE PACIENTES CON GLAUCOMA EN ESTADIO INICIAL.

Para este grupo de pacientes los criterios de inclusión fueron los siguientes:

- Edad mayor a 40 años.
- Mejor agudeza visual con corrección igual o superior a 0.5.
- Defecto refractivo esférico inferior a - 5 D y + 5D
- Defecto refractivo cilíndrico inferior a 3 D
- Diagnóstico definitivo de glaucoma crónico de ángulo abierto, uni o bilateral, establecido mediante criterio clínico.
- Presión intraocular igual o superior a 21 mmHg sin tratamiento hipotensor ocular.
- Campimetría fiable y reproducible en al menos 3 campos visuales consecutivos, con un daño perimétrico leve. Siguiendo los criterios de clasificación de Hodapp-Parish-Anderson (139), se consideró defecto perimétrico leve a los siguientes casos:
 - Defecto campimétrico con morfología arcuata de localización en área de Bjerrum y con desviación media mejor a -6 dB.
 - Resultado fuera de límites normales del test de hemicampo de glaucoma.
 - Valor patológico del índice CPSD/PSD con significación estadística $p < 0.05$.

El campo visual fue considerado fiable cuando el valor de falsos negativos fue inferior al 15%, el valor de falsos positivos fue inferior al 10% y el porcentaje de pérdidas de fijación inferior al 10%.

- Ausencia de cirugía ocular, excepto cirugía de facoemulsificación con implante de lente intraocular en saco capsular sin complicaciones intraoperatorias ni postoperatorias.
- Ausencia de patología ocular que pudiera inducir un defecto en el campo visual.
- Ausencia de cualquier enfermedad neurológica, especialmente referido a aquellas que pudieran presentar alguna alteración en el campo visual.

3.1.2. GRUPO CONTROL DE SUJETOS SANOS.

Los criterios de inclusión para los sujetos sanos que constituyen el grupo control fueron:

- Edad mayor a 40 años.
- Ausencia de antecedentes de enfermedad ocular previa.
- No presentar antecedente familiar de glaucoma en familiares de primer grado.
- Mejor agudeza visual con corrección igual o superior a 0.5.
- Defecto refractivo esférico inferior a ± 5 D.
- Defecto refractivo cilíndrico inferior a 3 D
- Ausencia de enfermedad ocular, con exploración oftalmológica normal, salvo presencia de facoesclerosis o catarata incipiente.
- Presión intraocular inferior a 21 mmHg.
- Campo visual normal, fiable y reproducible en al menos 3 campos visuales consecutivos.
- Ausencia de cirugía ocular, salvo facoemulsificación con implante de lente intraocular en saco capsular sin complicaciones intraoperatorias ni postoperatorias.
- Ausencia de enfermedad sistémica que pudiera presentar manifestaciones oculares de cualquier tipo.

3.2. CRITERIOS DE EXCLUSION.

3.2.1. GRUPO DE PACIENTES CON GLAUCOMA EN ESTADIO INICIAL.

Los criterios de exclusión empleados en el grupo de pacientes con glaucoma fueron los siguientes:

- Edad menor a 40 años.
- Agudeza visual con corrección inferior a 0.5.
- Defecto refractivo esférico superior a ± 5 dioptrías.
- Defecto refractivo cilíndrico superior a 3 dioptrías.
- Glaucomas congénitos o glaucomas secundarios de ángulo abierto.
- Anomalías del ángulo camerular.
- Angulo camerular estrecho o cerrado.
- Baja fiabilidad en la realización de los campos visuales, con índice de falsos negativos superior al 15%, falsos positivos superior a 10% o porcentaje de pérdidas de fijación superior a 15%
- Enfermedad ocular concomitante a la presencia de glaucoma.
- Enfermedad sistémica o neurológica que pudiera provocar defectos del campo visual.
- Cirugía ocular distinta de facoemulsificación con implante de lente intraocular no complicada.
- Anomalías congénitas o adquiridas de la cabeza del nervio óptico, excepto neuropatía glaucomatosa.

3.2.2. GRUPO CONTROL DE SUJETOS SANOS.

Los criterios de exclusión aplicados al grupo de sujetos normales del grupo de sujetos sanos fueron los siguientes:

- Edad menor a 40 años.
- Agudeza visual con corrección inferior a 0.5.
- Defecto refractivo esférico superior a ± 5 dioptrías.
- Defecto refractivo cilíndrico superior a 3 dioptrías.
- Anomalías adquiridas o congénitas del ángulo camerular.

- Baja fiabilidad en la realización de los campos visuales, con índice de falsos negativos superior al 15%, falsos positivos superior a 10% o porcentaje de pérdidas de fijación superior a 15%
- Presencia de enfermedad ocular, salvo facoesclerosis o catarata incipiente.
- Presencia de enfermedad sistémica o neurológica que pudiera provocar defectos del campo visual.
- Cirugía ocular distinta de facoemulsificación con implante de lente intraocular no complicada.
- Anomalías congénitas o adquiridas de la cabeza del nervio óptico, incluida la disversión papilar.

4. VARIABLES INCLUIDAS EN EL ESTUDIO.. PARAMETROS SELECCIONADOS PARA LA CREACIÓN DEL MODELO.

4.1. VARIABLES EPIDEMIOLOGICAS.

Se han incluido las siguientes variables epidemiológicas:

- Edad.

Se tomó como válida la edad en el momento inmediatamente anterior a la conclusión de todas las exploraciones de campo visual y OCT en el paciente.

- Sexo del paciente.
- Ojo afectado.

En aquellos casos en los que ambos ojos de un mismo sujeto cumplieran los criterios de inclusión, uno de los ojos fue seleccionado aleatoriamente para ser incluido en el estudio.

4.2. VARIABLES OBTENIDAS A PARTIR DE PARAMETROS DE LA OCT

Los parámetros de la OCT seleccionados como datos de entrada al sistema han sido:

- Valor promedio del grosor de la CFNR en los 360° peripapilares.

- Valor del grosor de la CFNR peripapilar en los cuadrantes superior, nasal, inferior y temporal.
- Valor del grosor de la CFNR peripapilar en los 12 sectores horarios en los que el dispositivo de la OCT subdivide de modo automatizado la CFNR alrededor del nervio óptico.

Todos los grosores de la CFNR fueron calculados a una distancia fija de 1,7 mm del centro de la cabeza del nervio óptico por la OCT y medidos en micrómetros (μm).

- Tamaño en milímetros cuadrados de la cabeza del nervio óptico, determinado automáticamente por la OCT.
- Índice señal/ruido de la OCT, como estimación de la calidad de la imagen obtenida.

4.2.1. IDENTIFICACION DE CUADRANTES Y SECTORES HORARIOS.

Dado que fueron incluidos tanto ojos derechos como izquierdos, se ha tomado como estándar para establecer la localización de los sectores horarios la correspondiente a la del ojo derecho, siguiendo un sentido horario y con las 12 horas en el polo superior del nervio óptico.

Los sectores horarios han sido identificados mediante la secuencia de H0 a H11, siguiendo un sentido horario, de modo que los sectores horarios H1 a H5 se ubicarán en la región nasal del nervio óptico, y los sectores horarios H7 a H11 se localizarán en el área temporal del nervio óptico (figura 19).

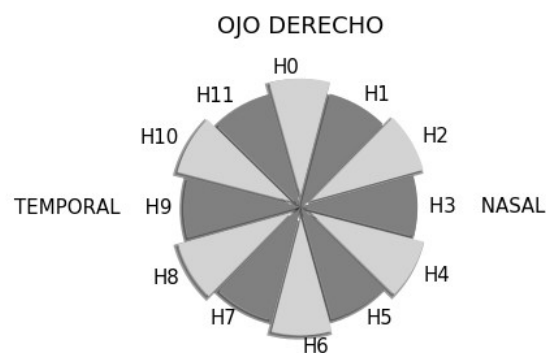


Figura 19. Ubicación de los sectores horarios en el ojo derecho.

4.2.2. CORRESPONDENCIA ENTRE OJO DERECHO y OJO IZQUIERDO.

Para la identificación de los sectores horarios en los ojos izquierdos, se ha llevado a cabo una trasposición respecto al ojo derecho, de modo que los sectores horarios de los ojos izquierdos han sido identificados de H0 a H11 siguiendo un sentido antihorario, con el sector horario H0 a las 12 horas en el polo superior del nervio óptico. De esto modo, se consigue la coincidencia correcta entre los sectores horarios nasales y temporales, independientemente de que se trate de un ojo derecho o izquierdo, tal y como se muestra en la figura 20.

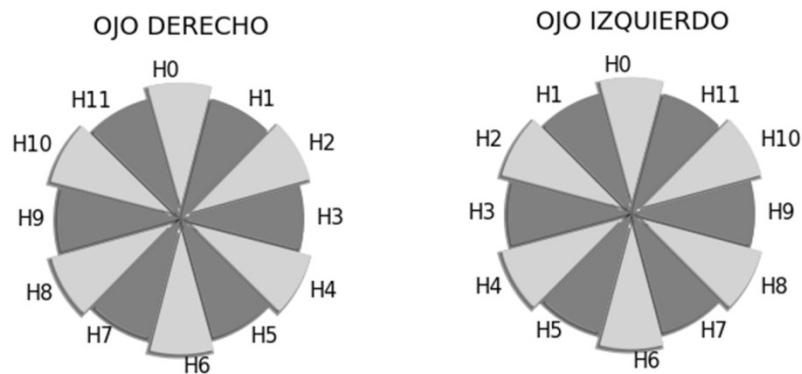


Figura 20. Correspondencia de los sectores horarios de la CFNR entre ojo derecho e izquierdo.

4.2.3. CONTROL DE LA CALIDAD DE IMAGEN DE OCT.

La calidad de la imagen ha sido controlada a través del índice de señal/ruido proporcionado automáticamente por la OCT.

Una baja calidad de imagen, estimada por un índice de intensidad de la señal inferior a 45 en la exploración de la CFNR mediante OCT, es un criterio de exclusión del paciente del estudio.

4.3. VARIABLES OBTENIDAS A PARTIR DE PARAMETROS DEL CAMPO VISUAL.

- Desviación media (DM).

El valor de DM elegido para el estudio ha sido el del último campo visual practicado al paciente antes de la realización del estudio del nervio óptico mediante OCT.

- Desviación estándar del patrón o PSD.

El valor de PSD escogido para la base de datos corresponde al del último campo visual realizado antes de llevar a cabo el análisis del nervio óptico mediante OCT.

- Hemicampo visual donde se localiza el defecto campimétrico.

La localización del defecto campimétrico en el grupo de pacientes con glaucoma se ha categorizado en las clases de:

- Superior.
- Inferior.
- Ambos hemicampos visuales.

5. SOFTWARE PARA ANALISIS DE DATOS.

Para el análisis estadístico y la aplicación de métodos de minería de datos se han utilizado los siguientes paquetes software:

- IBM@ SPSS versión 21 para Windows.

Con este software se han realizado los análisis preliminares, aplicando técnicas estadísticas clásicas para el estudio descriptivo inicial de las variables. También se han aplicado otros métodos convencionales, tales como:

- o Test de Kolmogorov-Smirnov, para evaluar el ajuste a la distribución normal de las variables cuantitativas.

- Prueba de la t de Student y U de Mann Whitney para muestras independientes, para comparar el valor medio de las variables cuantitativas.
 - Prueba de Chi cuadrado, para la comparación de la distribución en las variables cualitativas entre los grupos.
- R free software versión 3.5.2, ejecutado desde la plataforma RStudio versión 1.1.463 para sistema operativo Windows.

El entorno de análisis estadístico R ha sido el utilizado en la fase de preprocesamiento de los datos, previo a la creación del modelo, haciendo uso de técnicas de agrupación para la detección y eliminación de los valores anómalos o aberrantes de la base de datos.

- Librerías de software libre del lenguaje de programación Python necesarias para desarrollar los métodos de minería de datos y para la visualización de los resultados, incluyendo los paquetes:
- Pandas, NumPy y SciPy, para manipulación y operaciones científicas básicas de los datos.
 - SciKit-Learn, con operaciones matemáticas de minería de datos.
 - XGBoost, para la creación del modelo de clasificación de potenciación de gradiente.
 - Shap, con los métodos fundamentales para el cálculo de los valores Shapley.

6. ANALISIS ESTADISTICO GENERAL DE LA BASE DE DATOS.

6.1. ANALISIS ESTADISTICO DE LAS VARIABLES EPIDEMIOLOGICAS Y DEL CAMPO VISUAL.

Los test t-student y U de Mann Whitney han sido empleados para examinar la igualdad de medias en las variables cuantitativas, haciendo uso del primero en caso de ajuste a la distribución normal y del segundo en caso contrario.

La distribución de las frecuencias en las variables cualitativas ha sido comparada entre los grupos control y de pacientes con glaucoma mediante el test chi-cuadrado, aplicando el test exacto de Fisher si la frecuencia esperada en más de una celda de la tabla de contingencia era inferior a 5.

6.2. ANALISIS ESTADISTICO DEL GROSOR DE LA CFNR ENTRE LOS GRUPOS.

Los grosores de la CFNR de los grupos de pacientes con glaucoma y pacientes sanos han sido comparados mediante contraste de hipótesis de igualdad de medias con el test de t-student, y U de Mann Whitney, según su ajuste a la distribución normal.

Se han construido las curvas ROC para cada parámetro de la CFNR, calculando el área bajo la curva AROC como medida de su capacidad para discriminar entre glaucoma precoz y controles sanos

Para cada parámetro de la CFNR, se han determinado el grosor para el que la especificidad alcanza el valor mínimo de 0.7, y su sensibilidad asociada.

6.3. COMPARACION DEL TAMAÑO DEL DISCO OPTICO ENTRE GRUPOS.

El tamaño del disco óptico de los grupos del estudio ha sido comparado mediante el contraste de hipótesis de igualdad de medias aplicando el test de t-student para muestras independientes.

7. METODOLOGIA DE MINERIA DE DATOS EMPLEADA PARA EL ANALISIS DE LA BASE DE DATOS.

Los procedimientos de análisis de minería de datos practicados en el presente estudio se han representado en el diagrama de flujo de la figura 21 y figura 22. De manera resumida, los pasos seguidos han sido:

1. Creación de una base de datos etiquetada tras la selección de los pacientes en función de los criterios de inclusión y exclusión.

2. Construcción de una base de datos modificada, al eliminar la etiqueta de “Diagnóstico de glaucoma”, de manera que no pueda identificarse si el caso corresponde al grupo control o al grupo de pacientes con glaucoma.
3. Análisis mediante el algoritmo no-supervisado k-medias, determinando el número óptimo de agrupaciones y detectando los casos anómalos o extremos, que serán excluidos de la base de datos.
4. Recuperación de las etiquetas de “Diagnostico de glaucoma” para conformar una base de datos depurada y etiquetada.
5. Subdivisión aleatoria de la base de datos en un conjunto de datos de entrenamiento y un conjunto de datos de test o validación.
6. Construcción del modelo de clasificación de potenciación del gradiente basado en la estructura de árboles (modelo eXtreme Gradiente Boosting o XGB) y validación con los datos de test.
7. Establecimiento de una jerarquía de los parámetros de la CFNR según su importancia relativa en el modelo XGB.
8. Análisis de modelo XGB mediante el cálculo de los valores Shapley.
9. Determinación del impacto de cada variable en el resultado del clasificador XGB según los valores Shapley.
10. Análisis del tipo de correlación existente entre los valores de grosor para cada parámetro de la CFNR y su impacto en la predicción final del modelo, estimado por su valor Shapley.
11. Categorización de los parámetros de la CFNR, pasando de ser variables cuantitativas a variables cualitativas ordinales,
12. Construcción de un clasificador con alta interpretabilidad, de tipo árbol de decisión, con los parámetros modificados a forma cualitativa ordinal y que han mostrado mayor impacto en el modelo.
13. Validación del árbol de clasificación final con los datos de test.

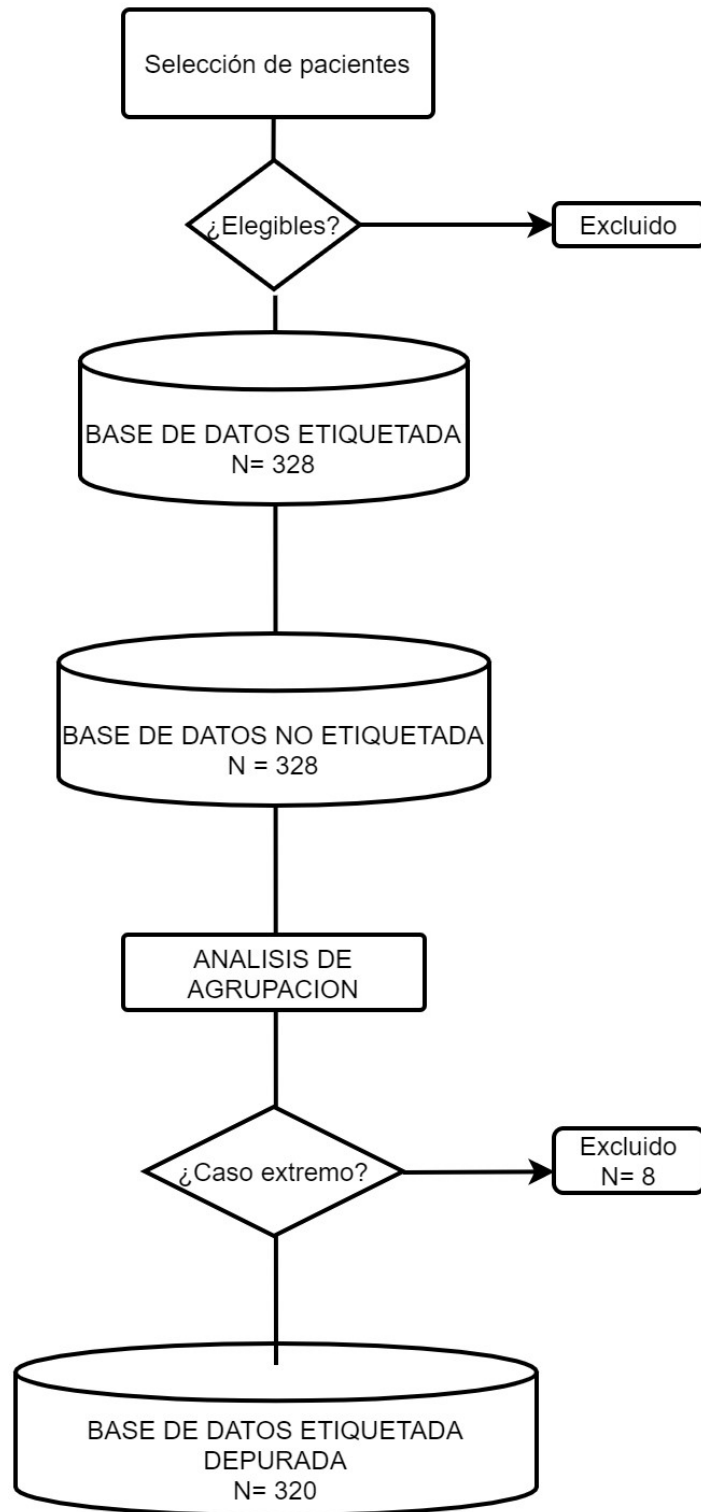


Figura 21. Flujo de trabajo hasta la producción de la base de datos etiquetada depurada.

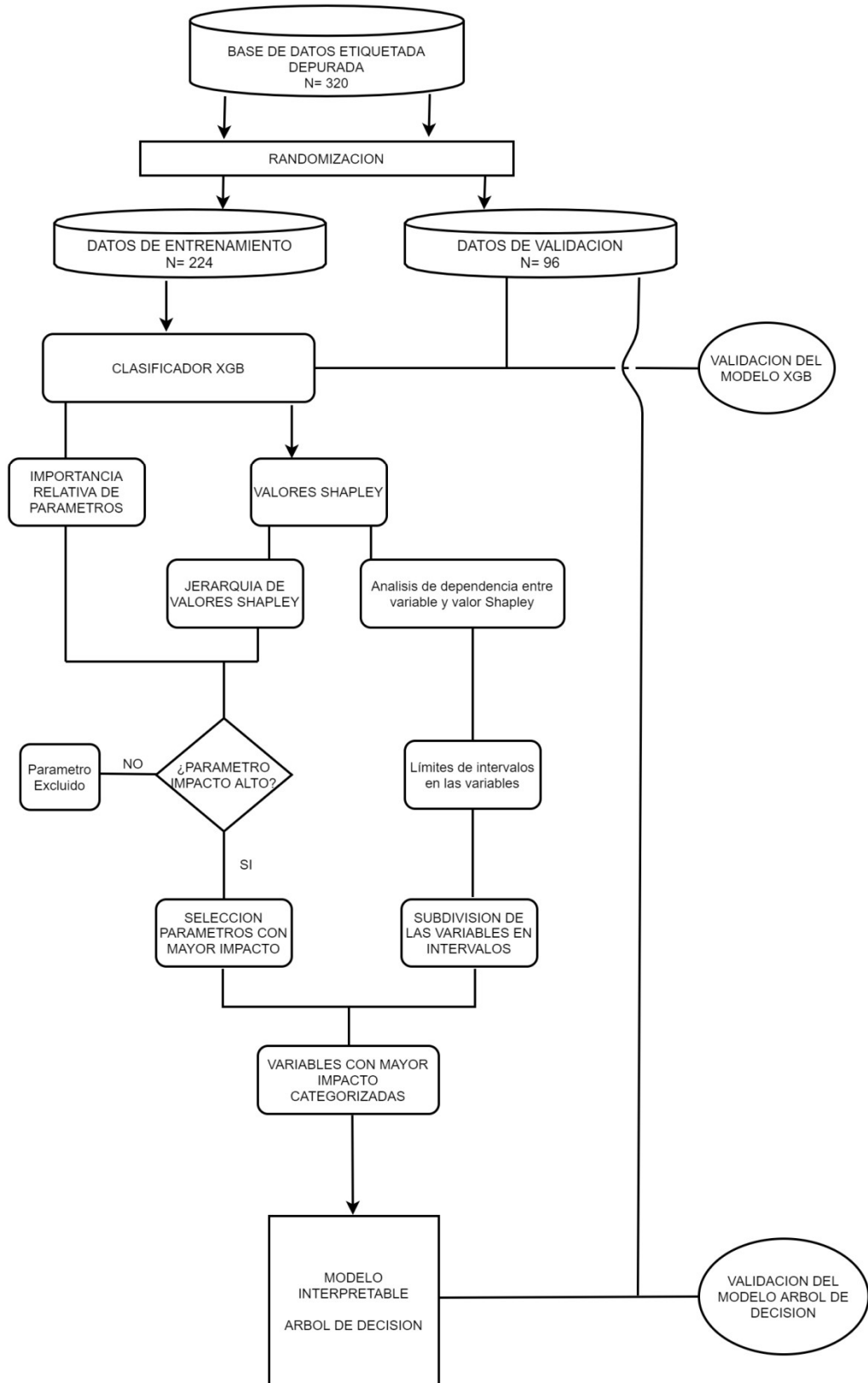


Figura 22. Flujo de análisis de datos desde la base de datos etiquetada tras eliminar casos extremos.

7.1. DETERMINACION DEL NÚMERO ÓPTIMO DE AGRUPACIONES.

A partir de la base de datos inicial, se ha generado una nueva base de datos modificada en la cual se ha eliminado el campo de “diagnóstico”, obteniendo de este modo una base de datos no etiquetada, constituida únicamente con los grosores de la CFNR. Sobre este banco de datos secundario, se ha llevado a cabo un análisis de agrupación no supervisado usando el software estadístico R y la metodología k-medias, con los objetivos de:

- Determinar el número óptimo de agrupaciones teniendo en cuenta únicamente los grosores de la CFNR.
- Determinar los valores extremos presentes en la base de datos, que serán excluidos de la misma, evitando que interfieran en los resultados de los análisis posteriores

El número óptimo de agrupaciones de los datos se ha deducido a través del cálculo del porcentaje de la varianza total que es explicado por el establecimiento de agrupaciones de los casos. La varianza hace referencia a la dispersión de las distancias existentes entre cada elemento del grupo a su punto central o centroide. El número óptimo de agrupaciones “n” ha sido aquel para el que se alcanza el máximo en el porcentaje de la varianza total del sistema que es explicada por la presencia de esas “n” particiones en los datos.

7.2. DETECCION Y ELIMINACION DE LOS CASOS EXTREMOS.

Tomando como valores de entrada los grosores de la CFNR, se ha realizado un análisis de agrupación, obteniendo dos conjuntos disjuntos de puntos – cada punto representando a un sujeto del estudio - con la característica de que los puntos dentro de un mismo conjunto tienen la mayor similitud entre ellos, a la vez que muestran la mayor diferencia con los componentes del otro conjunto. El valor central de cada conjunto se representa mediante el denominado centroide, correspondiente al punto más cercano al valor medio de los elementos del conjunto. Aquellos puntos localizados a una distancia mayor a la suma de la distancia media de los puntos al centroide de cada conjunto más dos veces su desviación estándar fueron considerados como valores extremos o anómalos y excluidos de la base de datos para los análisis posteriores.

Todos los análisis practicados a partir de este momento han sido llevados a cabo utilizando la base de datos original reducida al excluir los valores extremos o anómalos detectados.

7.3. CREACION DEL MODELO DE CLASIFICACION MEDIANTE EL ALGORITMO DE POTENCIACION DEL GRADIENTE BASADO EN ARBOLES DE CLASIFICACION.

Una vez eliminados aquellos casos que han sido considerados como extremos en el análisis de agrupación, se ha llevado a cabo la construcción de un modelo de clasificación basado en la metodología de potenciación del gradiente extremo (XGB) basado en árboles de decision. Este modelo ha sido elegido por su alta eficacia predictora.

La base de datos original ha sido dividida en dos subconjuntos:

- Conjunto de datos de entrenamiento, constituido por el 70% de los datos depurados. Corresponden a los datos usados para construir el clasificador.
- Conjunto de datos de validación o test, formado por el 30% de los datos originales. Son los datos utilizados para probar el modelo y confirmar la calidad y precisión del mismo.

La partición se ha realizado de modo aleatorio utilizando la generación de números aleatorios proporcionado por el paquete `random` del lenguaje Python.

La proporción de casos de pacientes con glaucoma y sujetos sanos en el conjunto de datos de entrenamiento y en el conjunto de datos de validación ha sido comparada mediante el test chi-cuadrado.

La obtención del clasificador más preciso se ha logrado eligiendo la configuración adecuada del algoritmo XGB para nuestra base de datos a través del ajuste de los hiperparámetros o variables internas del modelo. Los hiperparámetros que han sido ajustados han sido:

```
learning_rate
n_estimators
max_depth
```

```
min_child_weight  
gamma  
subsample  
colsample_bytree  
reg_alpha  
objective
```

7.4. ESTUDIO DE LA IMPORTANCIA DE LAS VARIABLES DEL MODELO EN EL CLASIFICADOR XGB.

Utilizando el grado de impureza o entropía dentro de los nodos como referencia, se han estimado cuales son los parámetros de la CFNR que muestran un mayor peso en el resultado del clasificador XGB, estableciendo una relación jerárquica en función de su importancia relativa en la predicción, es decir, a la hora de discriminar si un paciente es sano o presenta glaucoma en estadio inicial.

7.5. CALCULO E INTERPRETACION DE LOS VALORES SHAPLEY DEL MODELO DE CLASIFICACION XGB.

Una vez creado el modelo XGB para la predicción de la clase de las observaciones, se han calculado los valores Shapley de los parámetros de la CFNR para cada uno de los ejemplos de entrada al modelo – cada ejemplo se corresponde con un paciente.

Los valores Shapley han sido computados utilizando el paquete software `Shap` integrado dentro del lenguaje de programación Python.

7.5.1. ANALISIS DEL IMPACTO DE LAS VARIABLES DE LA CFNR EN LA PREDICCIÓN DEL MODELO XGB.

Teniendo en cuenta la relación existente entre el valor Shapley de cada parámetro de la CFNR y su repercusión en la predicción del modelo, se ha establecido una ordenación de los parámetros de la CFNR en función de la media de sus valores Shapley considerados en valor absoluto.

De este modo, se ha creado una nueva jerarquía de los parámetros de la CFNR en razón a su peso en el resultado del clasificador.

7.5.2. ANALISIS DE LA RELACION ENTRE GROSOR DE LA CFNR E IMPACTO EN EL RESULTADO DEL CLASIFICADOR: ANALISIS DE DEPENDENCIA.

La conexión entre los grosores de la CFNR medidos en micrómetros y el grado de impacto en la predicción del clasificador para ese grosor ha sido estudiada mediante la creación de gráficos bidimensionales en los que se correlaciona cada grosor con su valor Shapley. Estos gráficos son nombrados como análisis de dependencia.

Esta representación nos permite profundizar en la relación exacta que mantiene el grosor de la CFNR para cada parámetro y su valor Shapley.

Haciendo uso del método no-supervisado k-medias, en cada gráfico de análisis de dependencia se ha practicado la partición de los casos según su valor Shapley, deduciendo el grosor de la CFNR que establece la frontera divisoria entre las agrupaciones identificadas por el algoritmo.

7.6. SELECCIÓN Y CATEGORIZACION DE LOS PARÁMETROS DE MAYOR IMPACTO EN EL CLASIFICADOR.

Los índices de importancia relativa obtenida desde el modelo XGB y los valores Shapley han sido empleados como criterio para escoger los parámetros de la CFNR con mayor peso en el resultado del clasificador.

Los parámetros con mayor influencia en el clasificador XGB han sido seleccionados como variables de entrada al sistema para los modelos creados en fases posteriores del presente estudio, desestimando el resto de variables como entradas al modelo.

Los valores de grosor de la CFNR que constituyen los límites entre las agrupaciones registradas por el algoritmo k-medias en el análisis de dependencia han sido utilizados para la transformación de las variables de cuantitativas a variables cualitativas ordinales. De este modo, se han definido una serie de intervalos de grosor para cada parámetro de la

CFNR, convirtiendo la variable de grosor de ser cuantitativa y continua a variable discreta cualitativa.

7.7. CREACION DE UN CLASIFICADOR DE ALTA INTERPRETABILIDAD BASADO EN ARBOLES DE DECISION.

Una vez practicada la discretización de los parámetros de la CFNR con mayor impacto en la predicción del modelo XGB, se ha construido un nuevo clasificador de alta interpretabilidad como paso final del presente estudio.

El clasificador escogido corresponde a una estructura de árbol de clasificación, a fin de poder ser interpretado y contrastado con el conocimiento científico previo de la enfermedad del glaucoma.

RESULTADOS

1. ANALISIS DESCRIPTIVO DE LA BASE DE DATOS

1.1. VARIABLES POBLACIONALES

1.1.1. EDAD.

En la tabla 7 se muestra las características descriptivas de la variable edad para cada uno de los grupos, indicando el valor medio, la desviación estándar (DE) y el rango para cada grupo. Mediante la prueba de Kolmogorov-Smirnov (K-S) se comprobó el ajuste a una distribución normal de las muestras ($p > 0.05$). La comparación del valor medio de la variable edad mediante el test t-student muestra diferencia significativa entre los grupos, con mayor edad media en los pacientes con glaucoma ($p < 0.01$).

Tabla 7
Estadísticos de la variable edad (años)

	Media	DE	Rango
Glaucoma	62.55	8.81	43 – 81
Control	58.97	8.27	42 – 84

1.1.2. SEXO.

El reparto de los pacientes del estudio según la variable sexo viene detallado en la tabla 8. No se observaron diferencias significativas al comparar las proporciones de esta variable entre los grupos de glaucoma y sujetos sanos mediante el test de chi-cuadrado ($p > 0.05$).

Tabla 8
Estadísticos de la variable sexo

	Mujer		Hombre	
	Número	Proporción(%)	Número	Proporción(%)
Glaucoma	72	47.7	79	52.3
Control	99	55.9	78	44.1

1.2. VARIABLES OCULARES

1.2.1. PROPORCION DE OJO DERECHO / OJO IZQUIERDO.

Las proporciones de ojos derechos e izquierdos estudiados en cada grupo han sido comparadas mediante el test de chi-cuadrado, no obteniendo diferencias significativas entre los grupos de glaucoma y sujetos sanos ($p > 0.05$).

Tabla 9
Estadísticos de la variable ojo derecho / ojo izquierdo

	Número		Proporción (%)	
	OD	OI	OD	OI
Glaucoma	71	80	47	53
Control	100	77	56.5	43.5

1.2.2. VALORES DE GROSOR DE LA CAPA DE FIBRAS NERVIOSAS PERIPAPILARES.

La distribución de las medidas del grosor de la CFNR peripapilar muestra un ajuste a la distribución gaussiana para los parámetros de grosor promedio, grosor en cuadrantes superior, inferior, nasal y temporal, y grosor en los sectores horarios H1, H3, H4, H5, H6, H7, H9, H10 y H11 (Test K-S, $p > 0.05$). La distribución del grosor en los sectores horarios H2 y H8 no se ajusta a la distribución normal (Test K-S, $p < 0.05$).

Al comparar la media de cada parámetro del la CFNR – valor promedio, cuadrantes y sectores horarios - mediante los test t-Student y U de Mann Whitney para datos independientes se ha observado una diferencia significativa entre los valores de los grupos de glaucoma y control para todos los parámetros ($p < 0.01$).

La capacidad para el diagnóstico de glaucoma precoz del grosor de la CFNR ha sido evaluada mediante el cálculo del área bajo la curva ROC para cada parámetro. La representación de las curvas ROC y los valores del AUC obtenidos se representan en la figura 23 y la tabla 10.

Tabla 10

Valores de AUC para cada parámetro de la CFNR y intervalo de confianza 95%

	AUC	IC 95%
Valor promedio	0,858	0,817 – 0.898
CS	0,785	0,735 – 0.834
CN	0,682	0,624 – 0.740
CI	0,858	0,817 - 0.898
CT	0,712	0,655 – 0.769
H0	0,706	0,649 – 0.763
H1	0,705	0,648 – 0.761
H2	0,576	0,514 – 0.637
H3	0,640	0,580 – 0.699
H4	0,664	0,606 – 0.722
H5	0,730	0,676 – 0.783
H6	0,807	0,760 – 0.855
H7	0,693	0,637 – 0.750
H8	0,616	0,555 – 0.676
H9	0,670	0,612 – 0.729
H10	0,679	0,621 – 0.736
H11	0,724	0,669 – 0.778

AROC: Area bajo la curva ROC; IC: Intervalo de confianza

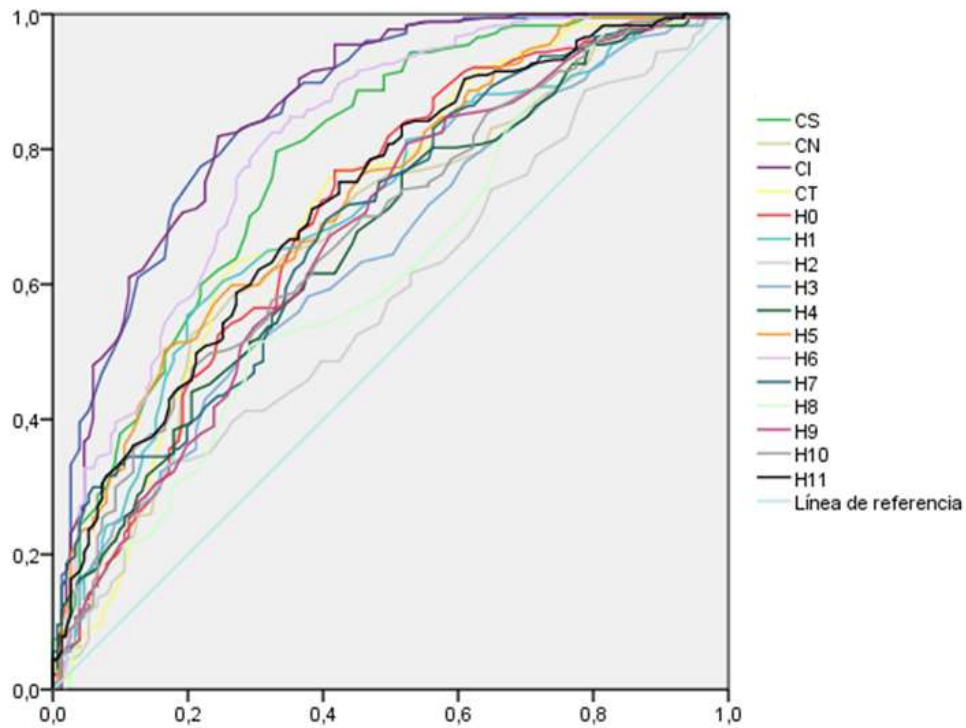


Figura 23. Curva ROC para cada parámetro individual de la CFNR

Los valores de corte para cada parámetro de la CFNR para los que muestra una especificidad superior a 0.7 y su nivel sensibilidad correspondiente se indican en la tabla 11.

Considerando el AROC como índice de la capacidad diagnóstica, el orden de los parámetros de la CFNR, organizados de mayor a menor capacidad diagnóstica, es el representado en la tabla 12. Los parámetros con mayor capacidad discriminativa son el grosor promedio de la CFNR peripapilar, el grosor en el cuadrante inferior, el grosor del sector horario H6 y el grosor del cuadrante superior de la CFNR.

Tabla 11

Valores de corte del grosor de CFNR con especificidad superior a 0.7 y nivel de sensibilidad.

	Grosor (μm)	Sensibilidad
Valor medio	86	0.83
CS	108	0.70
CT	63	0.63
CI	106	0.83
CN	71	0.58
H0	104	0.56
H1	89	0.62
H2	69	0.41
H3	65	0.19
H4	88	0.51
H5	117	0.59
H6	116	0.78
H7	75	0.48
H8	55	0.50
H9	66	0.53
H10	104	0.53
H11	115	0.62

1.2.3. TAMAÑO DEL DISCO OPTICO.

El tamaño medido en milímetros cuadrados de la cabeza del nervio óptico ha sido comparado entre los grupos de glaucoma y sujetos sanos. Los estadísticos obtenidos se recogen en la tabla 13. En ambos grupos la variable mostró un ajuste a la distribución normal (test K-S, $p > 0.05$).

Tabla 12

Parámetros de CFNR ordenados según AROC en orden descendente

	AROC
Valor promedio	0.858
CI	0.857
H6	0.807
CS	0.785
H5	0.730
H11	0.724
CT	0.712
H0	0.706
H1	0.705
CN	0.682
H7	0.693
H10	0.679
H9	0.670
H4	0.664
H3	0.640
H8	0.616
H2	0.576

Se encontraron diferencias significativas en los valores medios de tamaño discal entre los grupos de glaucoma y pacientes sanos (test t-student, $p < 0.01$), con un área discal de tamaño mayor en el grupo de pacientes con glaucoma.

Tabla 13

Estadísticos descriptivos de la variable tamaño del nervio óptico (mm²)

	Media	DE	Rango
Glaucoma	2.04	0.44	1.02 - 3.46
Control	1.90	0.33	1.18 – 2.96

1.2.4. VARIABLES DEL CAMPO VISUAL.

1.2.4.1. DESVIACION MEDIA (DM).

La media, desviación típica y rango de la DM de las perimetrías llevadas a cabo en cada grupo se presenta en la tabla 14.

Mediante el test K-S se confirmó el ajuste a la normalidad de la variable en ambos grupos. El test t-student encontró diferencia significativa en el valor medio de DM entre los grupos de glaucoma y sujetos sanos ($p < 0.01$), con una DM inferior en el grupo con glaucoma.

Tabla 14

Estadísticos descriptivos de la variable Desviación Media del campo visual.

	Media	DE	Rango
Glaucoma	-3.40	1.46	-5.96 – 0.15
Control	-0.95	1.45	-4.82 – 2.14

1.2.4.2. DESVIACION ESTANDAR DEL PATRON (“PATTERN STANDARD DEVIATION” – PSD).

La media y desviación típica de la variable Desviación Estándar del Patrón en cada grupo están recopiladas en la tabla 15.

La variable mostró un ajuste a la distribución normal en ambos grupos (test K-S $p>0.05$). El test t-student para comparación de medias de muestras independientes confirmó una diferencia estadísticamente significativa de PSD entre los grupos, con un valor mayor en el grupo de pacientes con glaucoma ($p<0.01$).

Tabla 15

Estadísticos descriptivos de la variable Desviación Estándar del Patrón

	Media	DE	Rango
Glaucoma	4.10	0.76	1.22 – 9.28
Normal	1.96	1.85	1.04 – 5.73

1.2.4.3. INDICE DEL CAMPO VISUAL (“VISUAL FIELD INDEX” o VFI)

La tabla 16 muestra los valores medios, DE y rango de la variable de índice del campo visual en los grupos de glaucoma y pacientes normales. En ambos grupos se apreció una distribución acorde a la curva gaussiana (Test K-S, $p>0.05$). El test t-student reveló diferencias estadísticamente significativas entre los valores medios entre el grupo control y el grupo con glaucoma ($p<0.01$).

Tabla 16

Estadísticos descriptivos de la variable Índice de Campo Visual

	Media	DE	Rango
Glaucoma	91.15	4.06	82 – 97
Control	98.54	1.47	93 – 100

1.2.4.4. AFECTACION DE HEMICAMPO VISUAL SUPERIOR O INFERIOR.

La distribución de la localización en hemicampo superior o inferior de los defectos campimétricos detectados en los pacientes con glaucoma está representada en la tabla

17. No se encontraron pacientes con defectos simultáneos en ambos hemisferios visuales superior e inferior. El test chi-cuadrado reveló que no existían diferencias significativas en la frecuencia de afectación de hemisferio superior o inferior ($p > 0.05$).

Tabla 17

Frecuencias de la variable Hemisferio afectado en el grupo de glaucoma precoz

	Número	Proporción (%)
Hemisferio superior	79	52.3
Hemisferio inferior	72	47.6

2. NÚMERO ÓPTIMO DE AGRUPACIONES. EXCLUSIÓN DE VALORES EXTREMOS DE LA BASE DE DATOS.

La figura 24 representa la relación existente entre el porcentaje de la varianza total que es explicado por la presencia de agrupaciones frente al número de agrupaciones creado mediante el algoritmo no-supervisado “k-medias”.

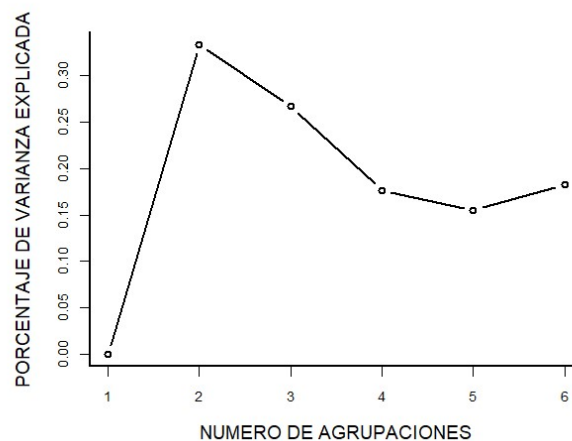


Figura 24. Relación entre porcentaje de varianza total explicada y número de particiones.

El número de conjuntos óptimo, con la mayor varianza dentro de los grupos respecto a la varianza total, es de 2 agrupaciones.

La figura 25 representa esquemáticamente la distribución de los casos en cada uno de los 2 grupos construidos por el algoritmo k-medias, utilizando el análisis de componentes principales a fin de poder hacer una representación bidimensional de los casos.

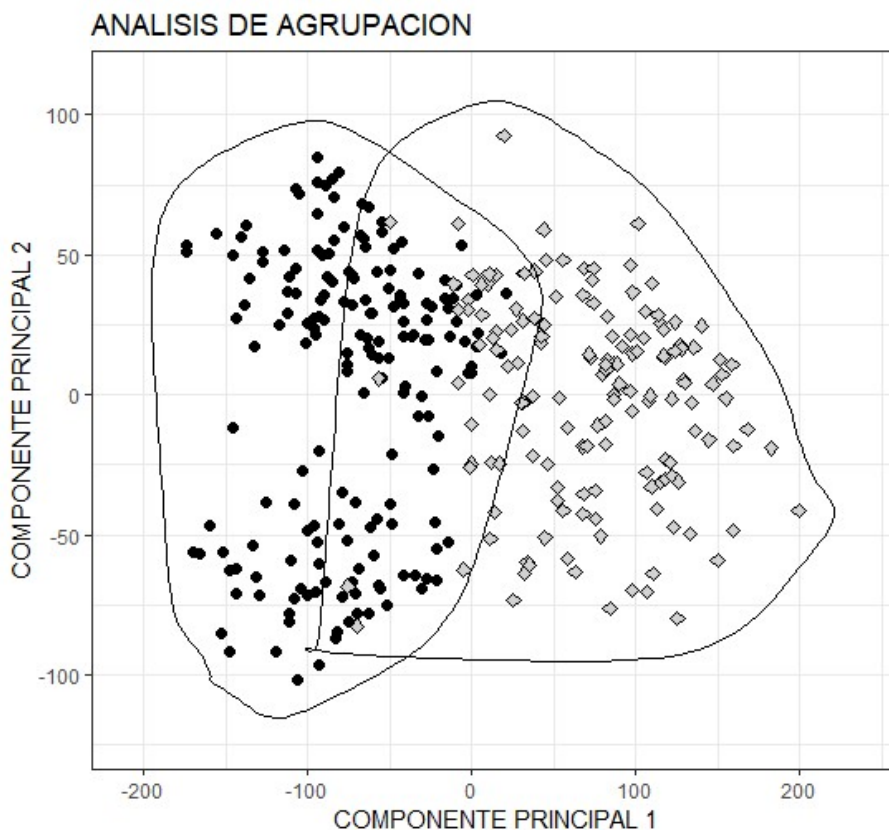


Figura 25: Representación gráfica bidimensional de los agrupamientos generados por el algoritmo k-medias según los dos primeros componentes principales.

Una vez creados los dos conjuntos de casos mediante el algoritmo no-supervisado k-medias, y computadas las distancias entre los componentes de cada conjunto a su centroide o valor central, se han identificado los valores extremos para cada grupo, excluyéndolos de la base de datos original. Los valores de la distancia media y la desviación estándar de las distancias al centroide dentro de cada grupo creado por el algoritmo vienen representadas en la tabla 18.

El número de casos extremos detectados ha sido de 5 para el grupo 1, con diagnóstico mayoritario de glaucoma en los casos incluidos en este grupo, y 3 casos para el grupo 2, con diagnóstico mayoritario de sujetos controles.

Tabla 18

Valor medio y desviación estándar de las distancias al centroide

	Media	DS	Límite anormalidad
Grupo 1 (CONTROLES)	86.18	30.00	146.19
Grupo 2 (GLAUCOMA)	76.90	23.04	122.98

En las figuras 26 y 27 se representa las distancias al centroide (ordenadas por valor decreciente) para los dos conjuntos que se han obtenido en el análisis de agrupaciones, indicándose la distancia establecida como límite de normalidad mediante la línea horizontal dibujada.

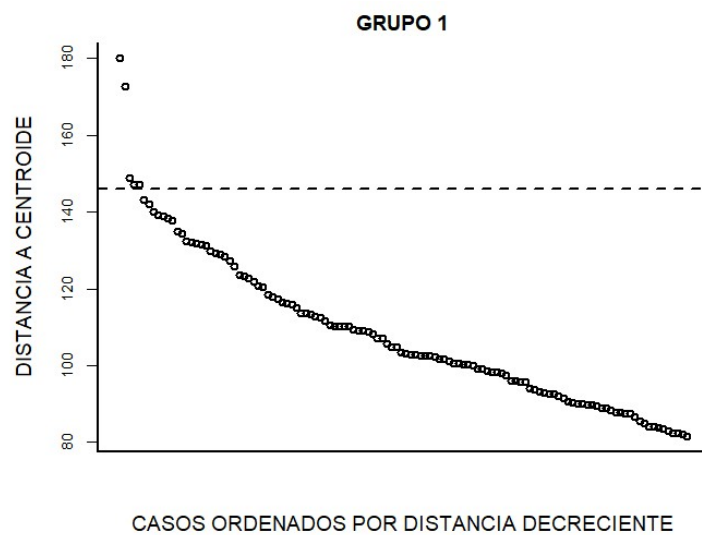


Figura 26: Representación de los casos en la partición [1] según su distancia a centroide (ordenador por distancia decreciente). La línea horizontal marca el límite para definir los casos extremos.

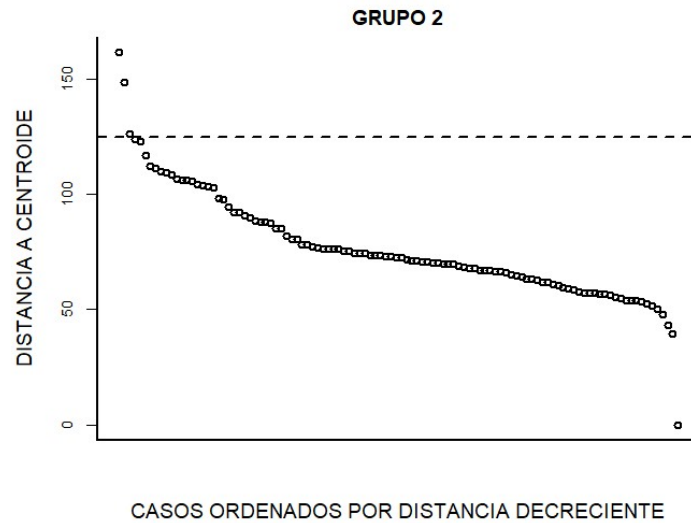


Figura 27: Representación de los casos en la partición [2] según su distancia a centroide (ordenador por distancia decreciente). La línea horizontal marca el límite para definir los casos extremos.

3. MODELO DE CLASIFICACION DE POTENCIACION DEL GRADIENTE BASADO EN ARBOLES DE DECISION (ALGORITMO EXTREME GRADIENT BOOSTING – XGB).

Una vez eliminados de los datos originales los casos extremos detectados en el paso anterior, la base de datos ha sido segmentada aleatoriamente en dos subconjuntos, correspondientes a:

- Datos de entrenamiento, usados para la construcción del modelo.
- Datos de validación o datos test, utilizados para la validación del modelo.

El número de casos incluidos en cada subconjunto viene especificado en la tabla 19, con la proporción de 70% para datos de entrenamiento y 30% para datos de test.

No se han encontrado diferencias significativas en la proporción de pacientes con glaucoma en los datos de entrenamiento y los datos de validación del modelo (test chi-cuadrado $p > 0.05$).

Tabla 19

Número de casos en los datos de entrenamiento y datos de validación y proporción de glaucoma vs sanos.

	Número	Proporción (%)	
		Glaucoma	Sanos
Datos de entrenamiento	224	44,2	55,8
Datos de validación	96	45,8	55,2

El modelo XGB de clasificación, creado utilizando los datos de entrenamiento, ha sido configurado ajustando los hiperparámetros internos del algoritmo para conseguir el máximo rendimiento posible, tomando como medida de rendimiento el error de clasificación de las entradas al modelo.

Los hiperparámetros para los que se ha obtenido el clasificador con el mejor rendimiento han sido:

```

learning_rate =0.1
n_estimators = 500
max_depth = 4
min_child_weight=7
gamma=0.1
subsample=0.8
colsample_bytree=0.8
reg_alpha=0.01
objective="binary:logistic"

```

El hiperparámetro `max_depth` establece la máxima profundidad de árbol, limitando el sobreajuste del modelo. En el parámetro `objective = "binary:logistic"` dispone la construcción de un modelo de clasificación binaria.

Las métricas del rendimiento del modelo XGB, evaluado sobre el conjunto de datos test, están recopiladas en la tabla 20. La exactitud o precisión del modelo se ha valorado como el porcentaje de casos correctamente clasificados.

Tabla 20

Medidas de precisión, AROC, sensibilidad y especificidad del modelo XGB.

	Precisión	AROC	Especificidad	Sensibilidad
Modelo XGB	83.33 %	0.828	0.88	0.77

3.1. IMPORTANCIA RELATIVA DE LAS VARIABLES EN EL MODELO DE CLASIFICACION XGB.

El impacto de cada parámetro de la CFNR en el resultado del clasificador ha sido estimado mediante el valor de la importancia relativa de cada variable de entrada al modelo. El valor de importancia relativa es proporcionado por un método encapsulado en el paquete de software XGBoost, utilizando como criterio de comparación la mejora de la entropía o impureza de los nodos del árbol. La figura 28 representa gráficamente el resultado obtenido, ordenando los parámetros de entrada al modelo según una importancia relativa decreciente.

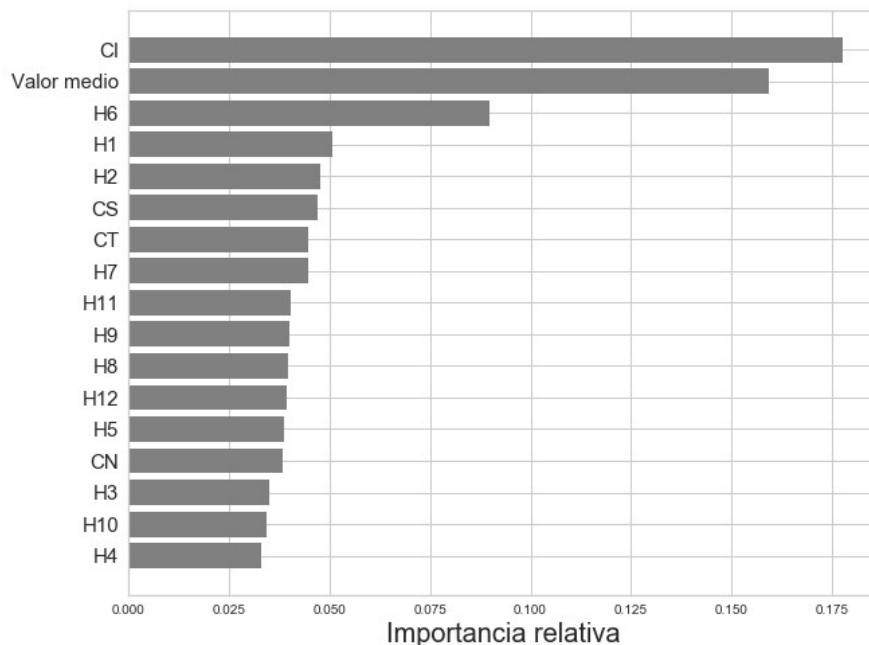


Figura 28: Gráfico de la importancia relativa de los parámetros de la CFNR ordenados en orden decreciente. CI: Cuadrante inferior; CS: Cuadrante superior; CT: Cuadrante temporal; CN: Cuadrante nasal; Hi: Sector horario “i”.

Los valores numéricos de la importancia relativa calculada para cada parámetro de la CFNR, ordenador por orden decreciente, se recogen en la tabla 21.

Tabla 21

Valores de importancia relativa de las variables de entrada al modelo XGB

	Importancia relativa
Valor promedio	0.17923
CS	0.03456
CT	0.01299
CI	0.21188
CN	0.04152
H0	0.07203
H1	0.04018
H2	0.001
H3	0.03621
H4	0.00890
H5	0.07337
H6	0.05824
H7	0.03512
H8	0.07866
H9	0.02863
H10	0.06823
H11	0.02016

De acuerdo a esta jerarquía, el grosor de la CFNR en el cuadrante inferior es el parámetro con la mayor importancia relativa en el clasificador, seguido del grosor promedio de la CFNR peripapilar, el grosor en los sectores horarios H6, H1, H2 y grosor en el cuadrante superior.

4. EVALUACION DE LA IMPORTANCIA DE LAS VARIABLES DEL MODELO MEDIANTE CALCULO DE LOS VALORES SHAPLEY.

El valor absoluto promedio de los valores Shapley computados para cada parámetro de la CFNR se encuentran recogidos en la tabla 29. Los parámetros con mayor impacto en la predicción, por orden decreciente, son los grosores en el cuadrante inferior, los sectores horarios H1 y H6, el cuadrante superior, el sector horario H2 y el valor promedio del grosor de la capa de fibras nerviosas peripapilares.

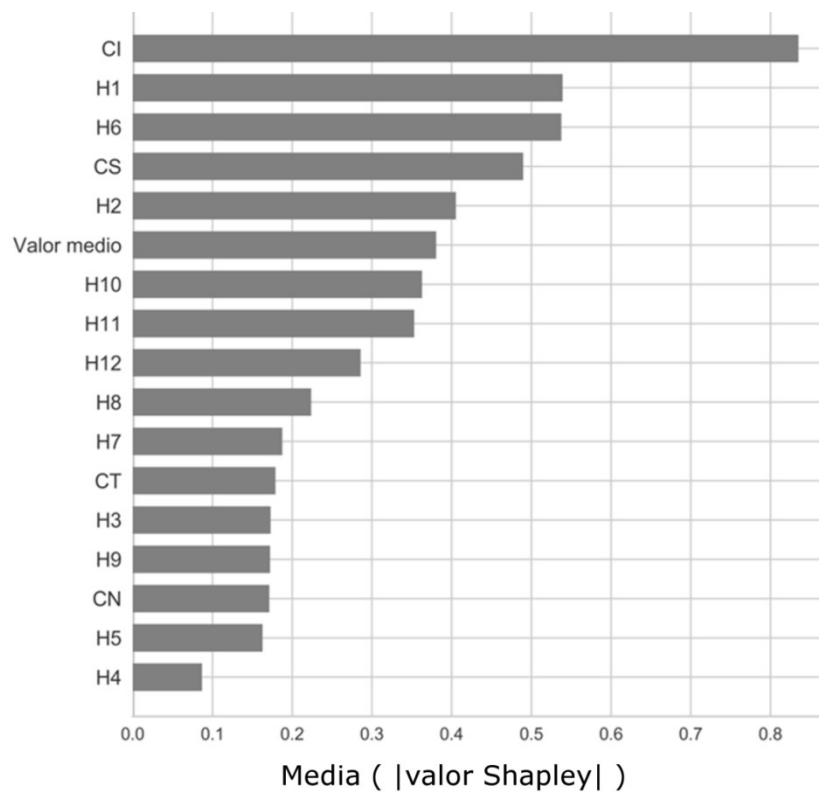


Figura 29. Ordenación de los parámetros de la CFNR según la media del valor Shapley sin signo (valor absoluto). CI: Cuadrante inferior; CS: Cuadrante superior; CT: Cuadrante temporal; CN: Cuadrante nasal; Hi: Sector horario “i”.

En nuestro modelo XGB, un valor Shapley positivo se asocia a una mayor probabilidad de que el caso sea clasificado como glaucoma precoz. Por el contrario, un valor Shapley negativo se corresponde con una mayor tendencia a que la predicción del modelo sea la de ser un sujeto sano.

La figura 30 muestra la distribución de los valores Shapley para cada parámetro de la CFNR en relación con su grosor en valor numérico (representado en codificación de color – rojo para grosores mayores y azul para grosor menor).

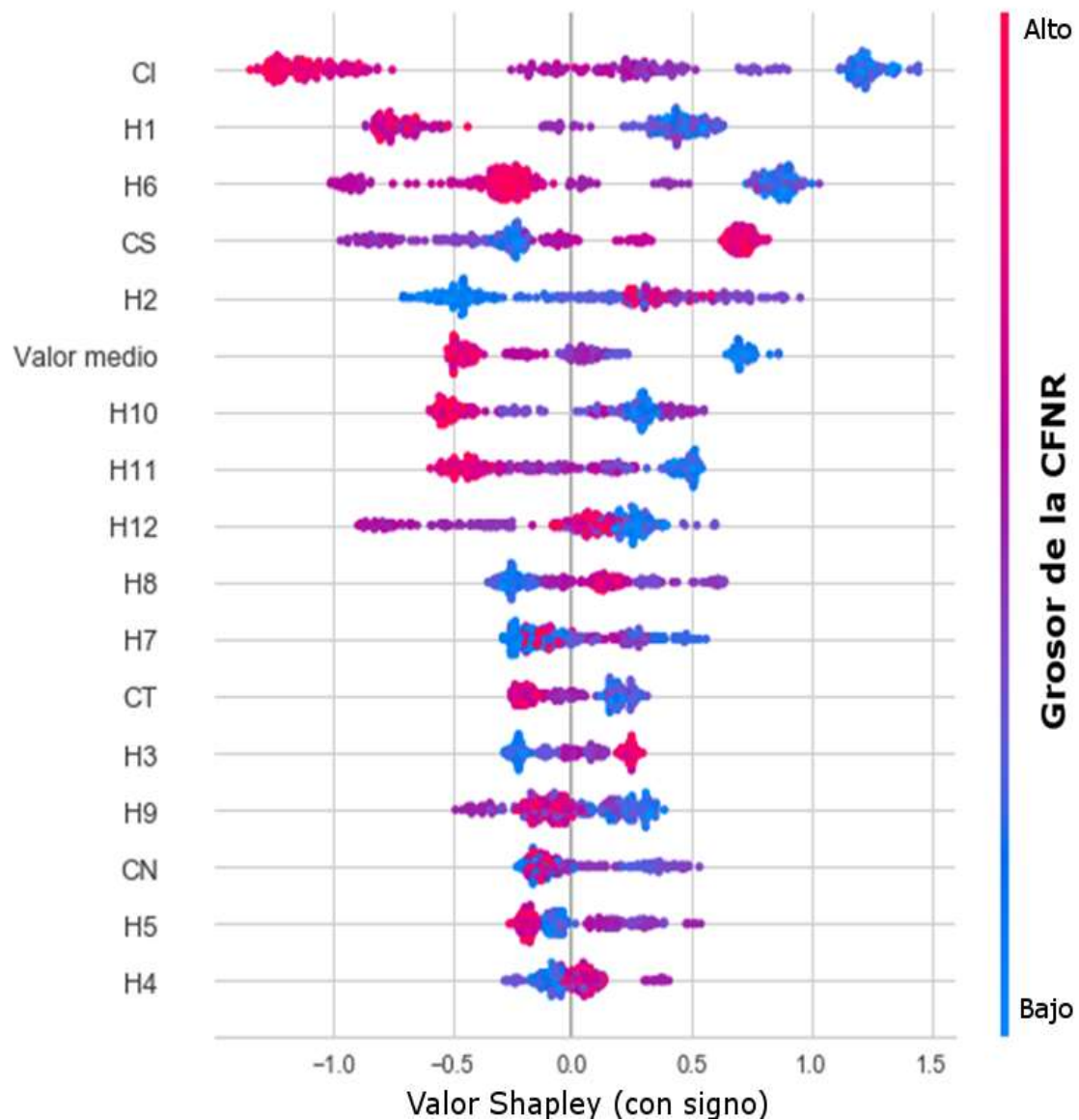


Figura 30. Representación de los valores Shapley según el valor numérico del grosor de la CFNR para cada parámetro. El color rojo y azul indica valores altos o bajos en el grosor de la CFNR respectivamente. La localización horizontal muestra si el efecto de ese valor de grosor causó una predicción mayor o menor en el clasificador y en qué sentido de predicción (glaucoma vs control). CI: Cuadrante inferior; CS: Cuadrante superior; CT: Cuadrante temporal; CN: Cuadrante nasal; Hi: Sector horario “i”.

De nuevo, el cuadrante inferior es el parámetro con mayor impacto en el resultado del modelo, tanto en valores Shapley positivos como negativos, de manera que un grosor mayor de la CFNR en el cuadrante inferior se corresponde con un valor Shapley negativo, y por tanto, con mayor probabilidad de que el caso sea clasificado como sujeto sano.

Esta correspondencia entre mayor valor del grosor de la CFNR y mayor tendencia a la clasificación como sujeto sano se observa también en los sectores horarios H1 y H6, así como el cuadrante superior y valor promedio de la CFNR.

El grosor en el sector horario H2, sin embargo, exhibe un comportamiento inverso, de modo que valores más bajos de la CFNR en este sector horario se corresponden con una mayor posibilidad de que el sujeto sea clasificado como sujeto sano.

5. ANALISIS DE DEPENDENCIA ENTRE EL VALOR DEL PARAMETRO DE ENTRADA Y EL IMPACTO EN LA PREDICCIÓN DEL CLASIFICADOR.

Los 6 parámetros con mayor importancia en la predicción del clasificador, tanto en el estudio según el grado de entropía en el modelo XGB como según la jerarquía de los valores Shapley obtenidos para esos parámetros, han sido seleccionados para llevar a cabo un análisis de dependencia, estudiando en detalle el tipo de correlación existente entre cada grosor de la CFNR y los valores Shapley.

5.1. Análisis del grosor de la CFNR en el cuadrante inferior.

La representación para el grosor de la CFNR en el cuadrante inferior frente al valor Shapley exhibe una relación lineal entre los valores de 95 μm y 120 μm . Para valores inferiores a 95 μm o valores superiores a 120 μm encontramos un valor Shapley en meseta, indicando que la influencia en la predicción del modelo es estable (figura 31).

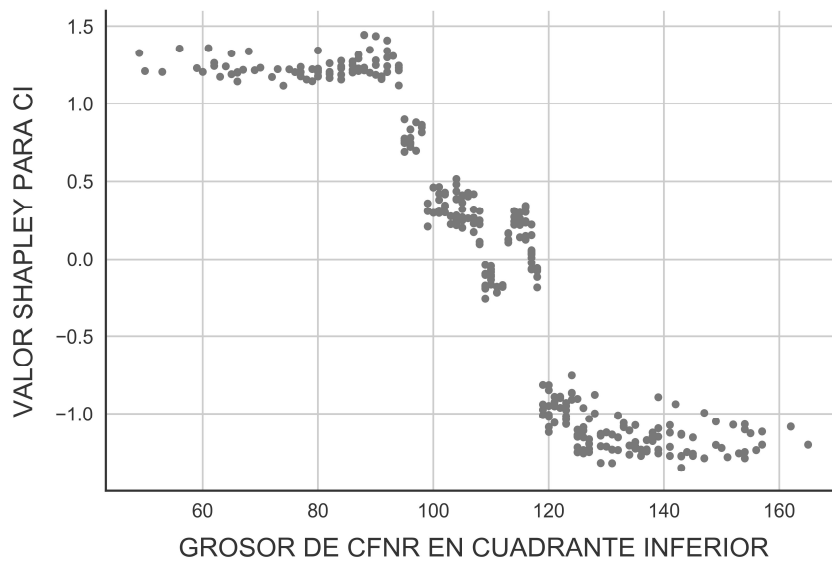


Figura 31: Gráfico de análisis de dependencia del grosor de CFNR en cuadrante inferior.

En el análisis no-supervisado de agrupamientos de los casos en función del valor Shapley, utilizado para delimitar las distintas regiones del gráfico, se ha encontrado que el número óptimo de agrupaciones es de 3. Los valores de grosor que delimitan los grupos creados han sido 95 μm y 120 μm (figura 32).

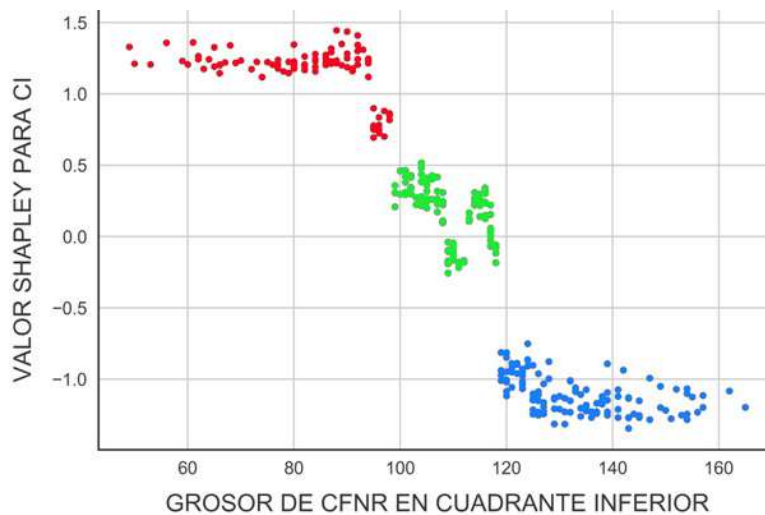


Figura 32. Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el cuadrante inferior.

5.2. Análisis del grosor en el sector horario H1 de la capa de fibras nerviosas.

El gráfico bidimensional para el sector horario H1 muestra dos agrupaciones, ambas en meseta, separadas por valor del grosor de H1 de 96 μm . Valores inferiores a 96 μm se asocian un valor Shapley positivo, y por lo tanto con una mayor probabilidad a predecir el caso como glaucoma. Ocurre lo opuesto para valores de H1 superiores a 96 μm (figura 33).

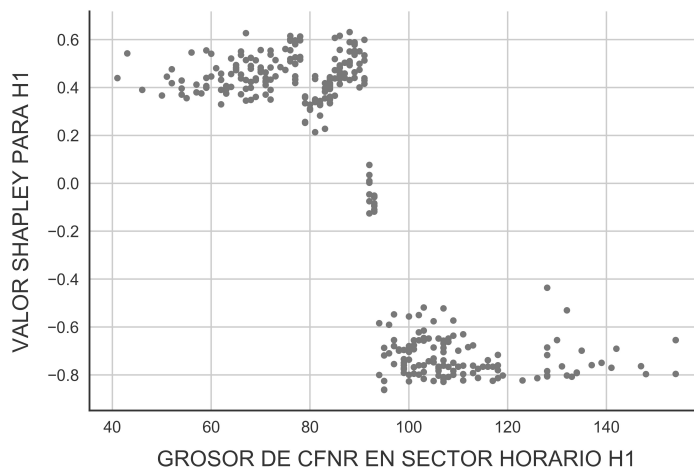


Figura 33: Gráfico de análisis de dependencia del grosor de CFNR en sector horario H1.

En el análisis no-supervisado encontramos que el número óptimo de agrupaciones es de 2, y el grosor que delimita los grupos es 96 μm (figura 34).

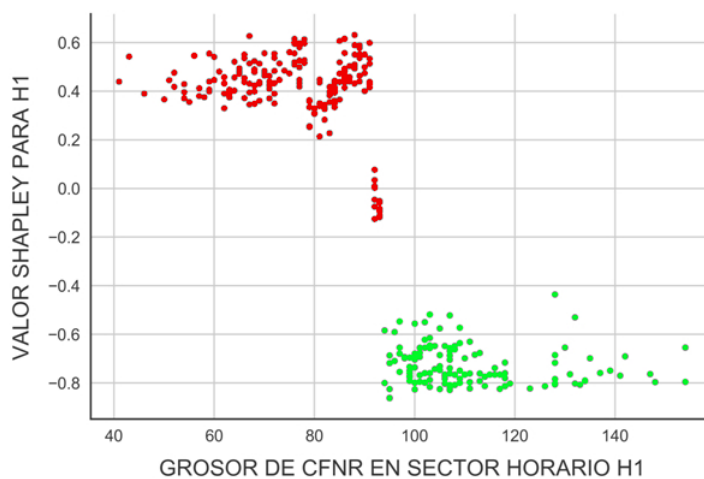


Figura 34. Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el sector horario H1.

5.3. Análisis del grosor en el sector horario H6 de la capa de fibras nerviosas.

El estudio del grosor de CFNR en el sector horario H6 muestra un patrón típico. Valores inferiores a 100 μm se asocian a un valor Shapley positivo, y por lo tanto a mayor tendencia a un resultado de glaucoma en el clasificador. Por otro lado, los valores superiores a 125 μm se correlacionan con un valor Shapley cercano a 0, por lo que su influencia en predicción del modelo es escasa o nula (figura 35). Los grosores entre 100 y 125 μm tienen valores Shapley con una alta variabilidad, alcanzando incluso valores Shapley negativos.

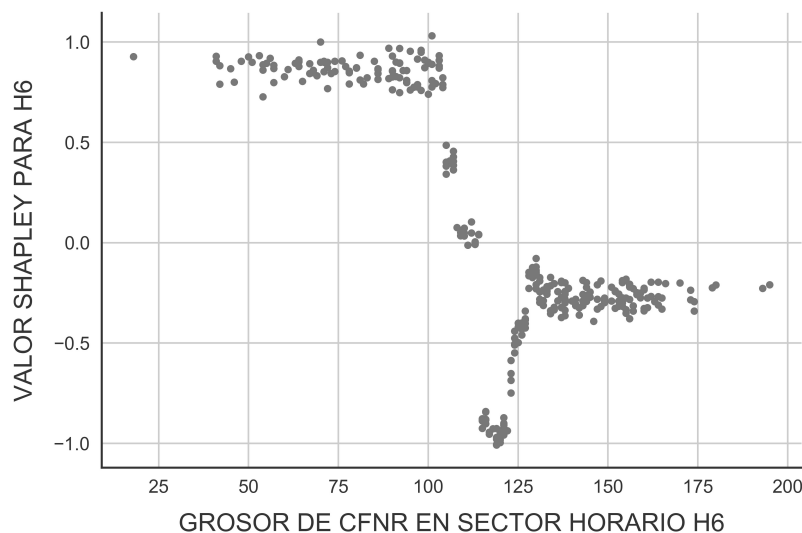


Figura 35: Gráfico de análisis de dependencia del grosor de CFNR en sector horario H6.

Al realizar el análisis de agrupamiento se definen dos grupos con una frontera en el grosor en el sector horario H6 de 100 μm (figura 36).

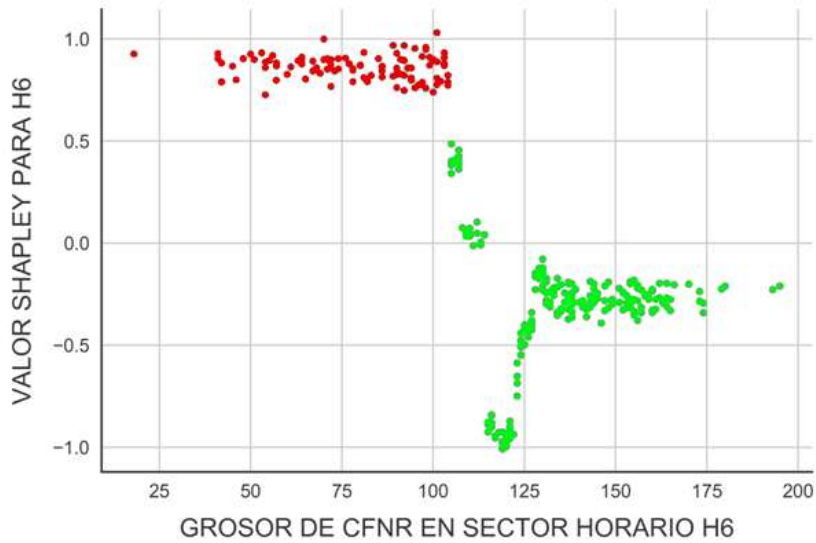


Figura 36. Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el sector horario H6.

5.4. Análisis del grosor en el sector horario H2 de la capa de fibras nerviosas.

En la representación de la capa de fibras en el sector horario H2 frente al valor Shapley se aprecia que un menor grosor de la CFNR en este sector horario se correlaciona con valores Shapley negativos. De este modo, valores más delgados de grosor en H2 muestran una tendencia a que el caso sea clasificado como sujeto sano (figura 39).

Por otra parte, el gráfico muestra que grosores superiores a 80 μm tienen una gran dispersión y están asociados a un valor Shapley cercano a 0, por lo que su influencia en la predicción es escasa.

Por estos motivos, añadido a que la variable de grosor en este sector horario H2 no se ajustaba a una distribución normal, se ha tomado la decisión de eliminar el grosor de la CFNR en el sector horario H2 como variable de entrada al clasificador creado como parte final del presente estudio.

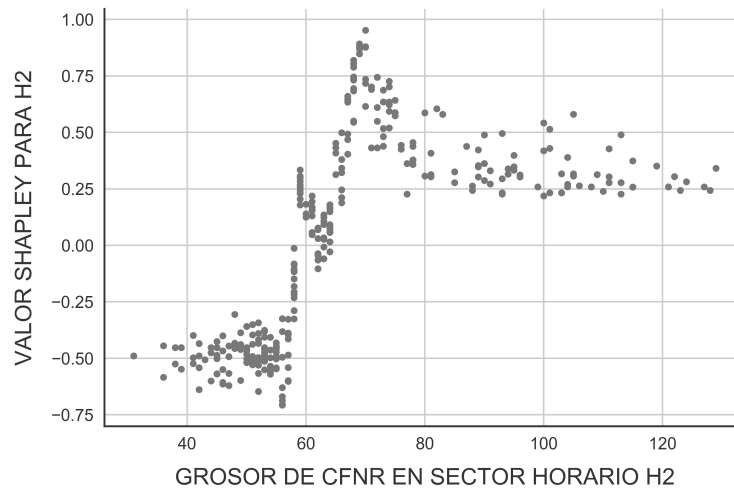


Figura 37: Gráfico de análisis de dependencia del grosor de CFNR en sector horario H2.

5.5. Análisis del grosor en el cuadrante superior de la capa de fibras nerviosas.

El grosor de la CFNR en el cuadrante superior muestra un gráfico atípico al ser representado frente a su valor Shapley, mostrando una morfología en V (figura 38). Un grosor inferior a 100 μm tiene un valor Shapley cercano a cero y, por tanto, con escaso impacto en el resultado del clasificador.

Los grosores entre 100 y 110 μm se asocian a valores Shapley negativos, con mayor tendencia a clasificar la observación como sujeto sano. Los valores más altos de grosor se asocian a un valor Shapley positivo, lo cual indica que la predicción tiende a ser hacia el diagnóstico de glaucoma, al contrario de lo que sería de esperar de acuerdo al conocimiento patogénico existente de la enfermedad glaucomatosa.

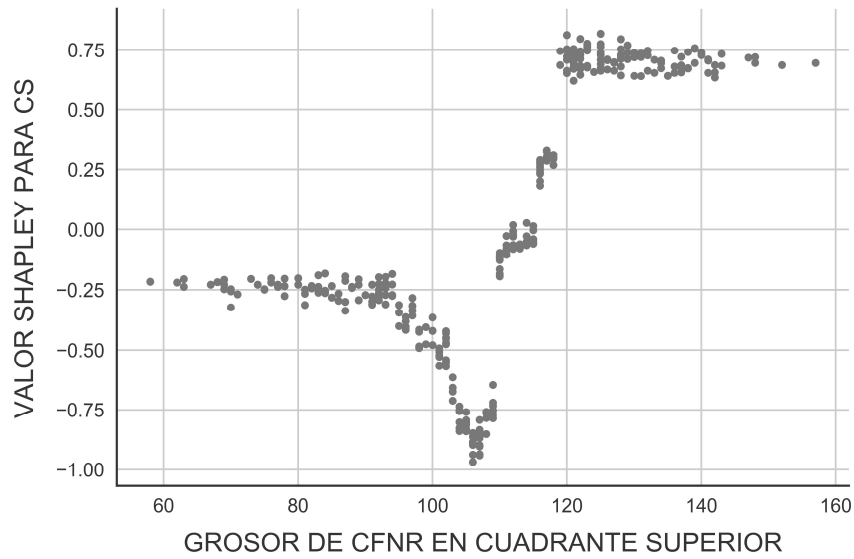


Figura 38: Gráfico de análisis de dependencia del grosor de CFNR en cuadrante superior.

El análisis por algoritmo no-supervisado k-medias evidencia que el número óptimo de agrupaciones es de 2, y el grosor frontera entre los grupos es 120 μm (figura 39).

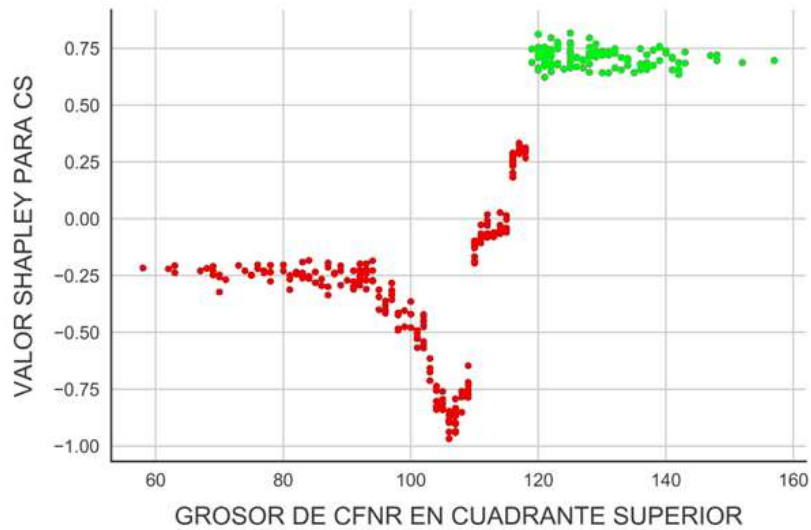


Figura 39. Agrupamientos creados mediante algoritmo no-supervisado de la CFNR en el cuadrante superior.

5.6. Análisis del grosor promedio de la capa de fibras nerviosas.

El estudio del grosor medio de la CFNR gráficamente revela la presencia de dos mesetas y una zona intermedia, delimitadas por los valores de grosor medio de la CFNR de 78 μm y 92 μm .

Un grosor inferior a 78 μm se asocia a valores Shapley negativos, con mayor tendencia a que el caso sea clasificado como glaucoma. Por el contrario, un grosor medio superior a 92 μm se corresponde a valores Shapley positivos, y por tanto, a mayor tendencia a que el caso se clasifique como sujeto sano (figura 40).

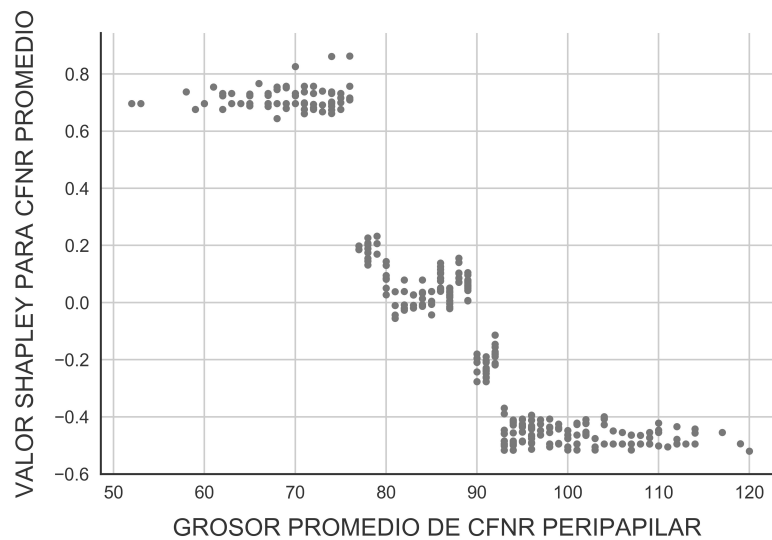


Figura 40: Gráfico de análisis de dependencia del grosor promedio de CFNR peripapilar.

El algoritmo k-medias encuentra 3 particiones de las observaciones, delimitados por los valores de grosor de 78 μm y 92 μm (figura 41).

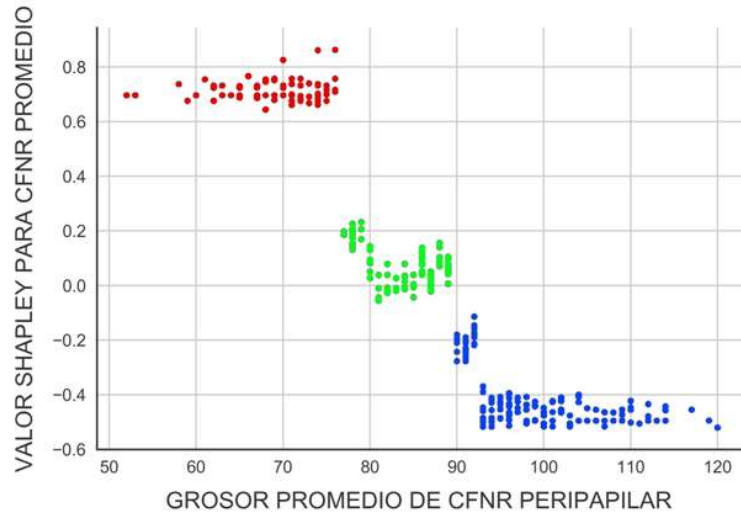


Figura 41. Agrupamientos creados mediante algoritmo no-supervisado para el valor promedio de la CFNR peripapilar.

6. CONSTRUCCION DE CLASIFICADOR DE ALTA INTERPRETABILIDAD BASADO EN ARBOLES DE DECISION.

A través del análisis de la importancia relativa de las variables de entrada a nuestro modelo XGB así como por el estudio del impacto en el clasificador según el valor Shapley, se deduce que las variables con mayor significación en la predicción del modelo son los grosores de la CNFR de los siguientes sectores:

- Cuadrante inferior.
- Sector horario H1.
- Sector horario H6.
- Cuadrante superior.
- Valor promedio de CFNR peripapilar.

Estas variables han sido seleccionadas como variables de entrada para la construcción de un clasificador con alta interpretabilidad de tipo árbol de clasificación.

El grosor en el sector horario H2 ha sido eliminado como parámetro de entrada en el árbol de decisión dado que el análisis de dependencia indica una relación entre grosor y

predicción de glaucoma que no es acorde al conocimiento previo sobre la enfermedad del glaucoma, así como la ausencia de ajuste de esta variable con la distribución normal.

Las variables seleccionadas se han categorizado y transformado de cuantitativas a variables cualitativas ordinales. Para ello se ha utilizado como límites entre las categorías los valores de grosor que constituyen las fronteras entre los distintos agrupamientos encontrados en el análisis de dependencia al representar el grosor de la CFNR para cada sector frente a su valor Shapley.

De este modo, las variables han sido discretizadas con las correspondencias mostradas en la tabla 22.

Tabla 22

Categorías para la conversión de las variables a cualitativas ordinales

	Categoría	Grosor de la CFNR (μm)
CI	1	< 95
	2	95 – 120
	3	>120
H1	1	< 96
	2	>96
H6	1	<100
	2	>100
CS	1	<120
	2	>120
Valor promedio	1	<78
	2	78-92
	3	>92

CI: Cuadrante inferior; H1: Sector horario 1; H6: Sector horario 6;
CS: Cuadrante superior

El árbol de clasificación ha sido construido utilizando los mismos datos de entrenamiento con los que se ha creado el modelo de clasificación XGB, con sus variables categorizadas a formato cualitativo ordinal. Para su validación se ha recurrido a los mismos datos de test utilizados en la validación del modelo XGB, obteniendo las métricas incluidas en la tabla 23.

Tabla 23

Medidas de precisión, AUC y valor de especificidad y sensibilidad del modelo árbol de clasificación

	Precisión	AUC	Especificidad	Sensibilidad
Modelo Arbol de clasificación	83.3 %	0.8216	0.96	0.70

La estructura de árbol obtenida se representa en la figura 42. Los nodos finales de la rama correspondiente a un grosor de CI superior a 120 μm agrupan al 36.5% de las observaciones, y es clasificado como sujeto sano en más de un 80% de los casos.

Los casos con grosor en el cuadrante menor a 95 μm y grosor del sector horario H1 inferior a 96 μm se predicen como glaucoma precoz con una probabilidad superior al 90%. Un grosor promedio inferior a 78 μm aumenta todavía más la probabilidad de que el caso sea clasificado como glaucoma.

Aquellos casos con un grosor del CI entre 95 y 120 μm tienen un alto error de clasificación, independientemente del grosor promedio de la CFNR.

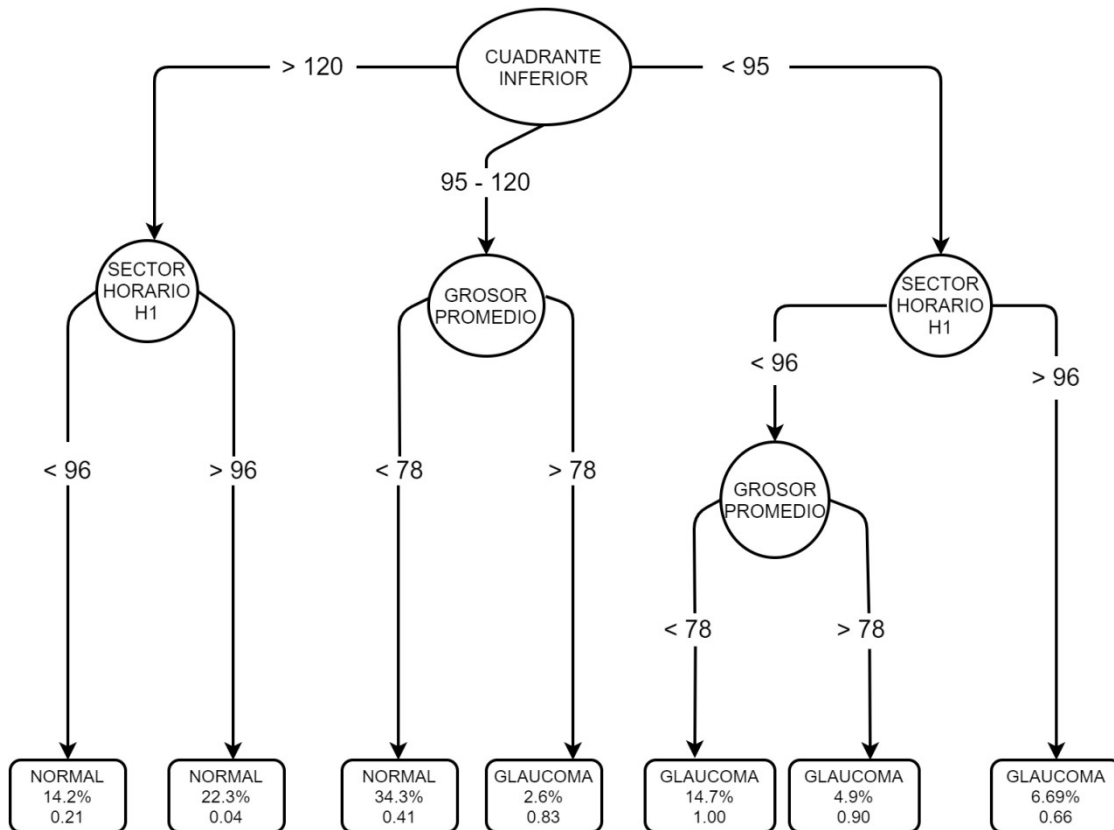


Figura 42. Árbol de clasificación construido con las variables categorizadas. Las líneas de unión entre los nodos incluyen el grosor de la CFNR en el que se realiza la partición. Los nodos finales muestran la predicción del modelo, así como el porcentaje de observaciones en el nodo y la probabilidad calculada de glaucoma precoz.

DISCUSSION

1. CONSIDERACIONES GENERALES Y METODOLOGICAS DEL ESTUDIO

El glaucoma es una de las causas principales de pérdida visual a nivel mundial. El diagnóstico en las primeras fases de la enfermedad nos otorga la oportunidad de realizar un tratamiento precoz. Hay evidencia científica de que en las etapas tempranas del glaucoma el tratamiento es más efectivo, siendo más sencillo alcanzar un control adecuado que pudiera evitar la pérdida visual asociada al avance de la enfermedad. Los ensayos clínicos prospectivos randomizados Early Manifest Glaucoma Trial y Ocular Hypertension Treatment Study han probado que el tratamiento precoz del glaucoma es beneficioso para disminuir el riesgo de progresión (140,141).

La importancia de la detección precoz ha sido justificada por Mansberger et al (142) al exponer que en el glaucoma confluyen las siguientes características:

- Es un problema de salud pública muy significativo.
- Es una enfermedad tratable
- Puede ser detectado en estadios tempranos.

Según diversos estudios, la prevalencia del glaucoma en la población general oscila entre un 2 y 5 % (11,14,15). La transcendencia socioeconómica y su impacto social son enormes, y con una tendencia a crecer con el paso de los años debido al envejecimiento de la población.

El glaucoma es una enfermedad con manifestaciones clínicas tardías y muchas veces indetectable por el paciente, lo cual dificulta su rastreo en la población general. Más del 50% de los enfermos desconocen que sufren glaucoma (19–22), hecho que se acentúa más aún en los países en vías de desarrollo o con difícil acceso a la asistencia sanitaria (23,24).

En el momento actual se considera que la evaluación clínica por un oftalmólogo experimentado es el procedimiento con la mayor fiabilidad para la detección del glaucoma. Los pilares sobre los que se asienta el diagnóstico son la exploración del nervio óptico, la medida de la presión intraocular y la presencia de escotomas en el campo visual (26). Las alteraciones en la CFNR son uno de los acontecimientos más

tempranos en la enfermedad (50,51). No obstante, el adelgazamiento de la CFNR muchas veces pasa desapercibido en la simple observación clínica del fondo de ojo, incluso para un oftalmólogo con amplia experiencia.

En nuestro estudio, la selección de los pacientes con glaucoma ha sido según los criterios clínicos que definen la enfermedad. Los paciente elegidos han sido aquellos que presentaban un glaucoma precoz, correspondiente al estadio de glaucoma leve establecido por Parrish et al (139). La superposición en la apariencia del nervio óptico entre sujetos normales y pacientes con formas incipientes de glaucoma determina que la detección del glaucoma precoz sea, en no pocas ocasiones, un dilema complicado de resolver. Los límites entre normalidad y patología en las fases iniciales son difusos y poco delimitados.

La exploración del campo visual se engloba dentro de los test psicofísicos, y como tal, tiene un alto grado de subjetividad y variación en las respuestas obtenidas al repetir la prueba en un mismo sujeto. Las anomalías en el campo visual con frecuencia son difíciles de evaluar debido a la alta variabilidad entre campos visuales seriados en pacientes con baja reproducibilidad de la prueba. Este obstáculo es especialmente manifiesto en sujetos de edad avanzada o con algún tipo de deficiencia cognitiva (143). La incidencia del glaucoma aumenta con la edad (15), siendo éste uno de los grupos que muestran mayor dificultad para la realización correcta de la prueba del campo visual.

En el comienzo de la enfermedad, los defectos perimétricos son poco significativos, restringiéndose a unos pocos puntos conexos con reducción en el umbral luminoso. La detección fiable de estos defectos sutiles se torna de una gran dificultad ante el menor grado de variabilidad intercampo, no siendo infrecuente que no podamos asegurar la presencia de dichos escotomas ni su tamaño ni profundidad.

Por otra parte, el campo visual, si bien ha sido parte del patrón de referencia en el diagnóstico del glaucoma, es una exploración que consume una gran cantidad de recursos y de tiempo para su realización, incluso con las nuevas estrategias perimétricas rápidas de SITA o SITA Fast. Si bien este no debe de ser un factor decisivo, es un aspecto a tener en cuenta a la hora de organizar los recursos dentro de la organización sanitaria.

Por ello, la implementación de un procedimiento de diagnóstico precoz del glaucoma que esté asentado en variables distintas a la presencia de los defectos del campo visual tendría un gran valor intrínseco en el manejo de estos pacientes. En el presente trabajo, el campo visual ha sido evaluado por un comité de oftalmólogos, dictaminando si el paciente sufría o no glaucoma según los hallazgos del campo visual así como del resto de la exploración clínica. Pero una vez establecido el diagnóstico, los parámetros del campo visual no han sido incluidos como datos de entrada para la construcción del modelo de clasificación. De este modo, el clasificador obtenido puede realizar la predicción sobre cada observación únicamente utilizando los parámetros de grosor de la CFNR, pudiendo ser aplicado en el amplio grupo de sujetos en los que el campo visual no es valorable por su baja fiabilidad o no puede realizarse por algún motivo. Dado que la etiqueta de diagnóstico de cada paciente de nuestra base de datos ha sido conferida teniendo en cuenta el estudio perimétrico, ha sido importante la exclusión del estudio de todos aquellos pacientes que pudieran presentar anomalías del campo visual de causa distinta al glaucoma, ya sea de origen ocular o neurológico.

Para que podamos obtener un modelo a partir de un banco de datos, tienen que cumplirse tres condiciones:

- Que existan datos en número suficiente
- Que los datos tengan algún patrón entre ellos.
- Que el patrón que relaciona los datos entre sí sea estable.

Una de las condiciones sine qua non para poder crear un modelo de calidad que proporcione predicciones acertadas es que debe haber datos que sean correctos y válidos con los que entrenar el sistema. En la tarea de clasificación es inexcusable que los datos estén correctamente etiquetados. De datos incorrectamente etiquetados obtendremos un modelo cuyas predicciones serán también incorrectas. Esto justifica el uso del campo visual como criterio de inclusión en el presente estudio, intentando que la etiqueta diagnóstica de cada sujeto fuera lo más precisa posible.

La CFNR es una de las estructuras que se afectan en los primeros estadios del glaucoma. Se ha demostrado la correlación existente entre el grosor calculado en cortes histológicos de la retina y el grosor medido mediante OCT in vivo. La reducción en el

número de fibras axonales de las células ganglionares es un suceso que antecede hasta 5 años a la aparición de escotomas en el campo visual (50). Por otra parte, hay una fuerte relación entre la reducción del grosor de la CFNR y la disminución de la sensibilidad lumínica en el campo visual (49). Es innegable, por tanto, que el estudio de la CFNR nos podría permitir anticiparnos a la aparición de anomalías detectables en el campo visual.

La OCT es un instrumento de exploración ocular accesible y actualmente presente en la práctica totalidad de los sistemas sanitarios a nivel mundial. Todas las OCT vienen equipadas con un software para la segmentación de las capas de la retina, permitiendo determinar el grosor de la CFNR en el área peripapilar. Existen otros paquetes software que pueden ser instalados como complementos en la OCT para obtener información adicional de las imágenes, pero habitualmente tienen un alto coste económico y no tienen una alta difusión en el sistema sanitario, siendo utilizados en pocas ocasiones y con acceso muy limitado.

En nuestro estudio, el clasificador ha sido creado con los parámetros de grosor de la CFNR proporcionados por el software convencional de la OCT. El hecho de no hacer uso de software especializado, sino únicamente de los datos de OCT a los que cualquier oftalmólogo general puede acceder, dan mayor universalidad a nuestro modelo, estando disponible para poder ser utilizado por la inmensa mayoría de los oftalmólogos.

La CFNR es una de las estructuras oculares que han adquirido mayor preponderancia en el estudio del glaucoma en las últimas décadas. Mauschitz et al (144) han indagado los factores que muestran una influencia significativa en el grosor de la CFNR peripapilar en la población europea, hallando que los siguientes factores se asocian a un menor espesor de la CFNR:

- Edad más avanzada.
- Presión intraocular más elevada.
- Agudeza visual menor.
- Historia de hipertensión arterial.
- Diagnóstico de demencia, aunque en este caso no encuentra asociación significativa.

Un dato de suma importancia es la reproducibilidad en las mediciones del grosor de la CFNR mediante OCT. Como hemos indicado previamente, un modelo válido sólo puede obtenerse con datos correctos. Por ello es esencial que los parámetros de la OCT usados como datos de entrada al sistema sean válidos y reproducibles. Esta reproducibilidad ha sido avalada en múltiples estudios (60–62). Un valor adecuado del índice señal-ruido o SSI en las imágenes de la OCT es igualmente primordial para asegurar que los datos de grosor de la CFNR son válidos (59).

2. CARACTERISTICAS BASALES DE LA BASE DE DATOS.

En el presente estudio se han recogido los datos correspondientes a 328 sujetos. De ellos, 177 fueron sujetos con exploración oftalmológica normal, tomados como controles, y 151 fueron diagnosticados de glaucoma en fase precoz, aplicando los criterios de inclusión y exclusión establecidos.

La comparación de la variable edad entre los dos grupos de estudio no presentó diferencias significativas. Si bien es conocido que hay una reducción del grosor de la CFNR con el aumento de la edad (145,146), la homogeneidad entre los grupos comparados asegura que esta variable no ha interferido en los resultados obtenidos.

Diversos estudios encuentran diferencias en el grosor de la CFNR según el sexo, con mayor grosor en la mujer respecto al hombre (146,147). En nuestros datos, la variable sexo se ha distribuido por igual en los grupos control y de glaucoma, por lo que descartamos la influencia de esta variable en el resultado. Igualmente, la proporción de ojo derecho /izquierdo no ha mostrado diferencias significativas entre los grupos.

Como era presumible, los parámetros de DM, PSD y VFI difieren en los grupos control y grupo de glaucoma, corroborando la exactitud del diagnóstico de glaucoma precoz y la validez nuestro banco de datos.

El análisis individual de los grosores de la CFNR ha mostrado una distribución ajustada a la curva normal para la mayoría de los parámetros, tal y como corresponde a una variable de origen biológico. No obstante, los sectores horarios H2 y H8 no han

mostrado este ajuste. Este hallazgo podría ser explicado por una mayor variabilidad o menor reproducibilidad de las medidas en estos sectores, ya que los sectores anexos a ellos sí que presentan una distribución gaussiana.

El AROC de cada parámetro de la CFNR puede ser empleado para establecer una gradación de su capacidad diagnóstica, evaluada según su sensibilidad y especificidad. Budenz et al (63) establece en su trabajo que los parámetros de la CFNR con mayor AROC son el grosor del cuadrante inferior seguido del valor promedio, el cuadrante superior y el sector horario H6. Bowd et al (148) comparando sujetos control con pacientes con glaucoma precoz, con una MD media en el campo visual de -4.00 ± 4.2 dB, encuentra que los parámetros con mayor AROC son el cuadrante inferior, el sector horario H6 y el valor promedio de la CFNR. Kanamori et al (149) evalúa a un grupo de sujetos con glaucoma precoz, con MD de -3.1 ± 1.8 dB. Las curvas ROC para el cuadrante inferior y el sector horario H7 fueron las de mayor AROC.

Chen et al (150) efectúan una revisión de las publicaciones acerca del uso de la OCT para detectar cambios asociados con el diagnóstico del glaucoma. Uno de sus apartados recoge específicamente los parámetros con el mismo instrumento que hemos empleado en el presente estudio, 3D OCT-2000 de la casa comercial TOPCON, indicando que los parámetros de la CFNR con mejor potencial para diagnosticar el glaucoma son, por orden descendente:

1. Grosor promedio de la CFNR peripapilar (AROC: 0.93 -0.97)
2. Grosor en el cuadrante inferior (AROC: 0.89 -0.95)
3. Grosor en el cuadrante superior (AROC: 0.82 -0.99)

En nuestro trabajo hemos encontrado que los parámetros de la CFNR con mayor AROC son, en orden decreciente, el valor promedio, el cuadrante inferior, sector horario H6 y cuadrante superior, siendo estos hallazgos consistentes con los resultados de estudios previos. La primacía del cuadrante inferior como parámetro con mayor sensibilidad es un dato repetido en casi todos los trabajos. Ello podría ser explicado por el mayor grosor de la CFNR a este nivel, haciendo que sea más sencillo detectar su adelgazamiento, o por el hecho de que el hemisferio visual superior sea afectado más habitualmente que el inferior. Este último dato, sin embargo, no se sustenta en nuestro

estudio, ya que no hay diferencia entre la localización del defecto perimétrico en hemicampo superior e inferior.

El tamaño del disco óptico ha mostrado diferencias significativas entre los grupos control y de glaucoma, siendo mayor en el grupo de pacientes con glaucoma. Esta observación es congruente con lo que describe Quigley et al (151) al apreciar mayor tamaño discal en un grupo de glaucoma comparado a controles sanos. Esta diferencia, sin embargo, desaparece al hacer la corrección para la edad, sexo y raza. Quigley hace su investigación calculando el área discal mediante oftalmoscopia corregida para el error refractivo, lo cual no es confrontable a la determinación del área discal que establece automáticamente la OCT. Sekhar et al (152) encuentran que los discos ópticos en pacientes con hipertensión ocular son significativamente menores que aquellos de sujetos sanos o pacientes con glaucoma manifiesto. En cualquier caso, la literatura sobre la posible relación entre tamaño discal y glaucoma es escasa y con resultados poco concluyentes. Por ello, razonamos que la diferencia de tamaño discal que hemos encontrado en nuestro estudio es improbable que tenga repercusión para el resto del análisis.

3. FUNDAMENTOS DE NUESTRO MODELO DE CLASIFICACION.

El modelo de clasificación propuesto en esta tesis tiene como finalidad la predicción de la presencia de glaucoma precoz en un sujeto a partir de los datos obtenidos en la exploración de la CFNR mediante OCT. Los conceptos en los que hemos cimentado este planteamiento podemos concretarlos en dos ideas esenciales:

1. La afectación de la CFNR antecede a las manifestaciones funcionales del glaucoma, expresadas como escotomas del campo visual. El diagnóstico precoz, por tanto, debe sustentarse en el estudio de las estructuras afectadas más tempranamente.
2. Las alteraciones detectables en la cabeza del nervio óptico son consecuencia de la reducción en el número de fibras del rodete neuroretiniano (RNR). Múltiples estudios han mostrado un rendimiento diagnóstico mayor en el análisis de la

CFNR que en el estudio de los cambios en el nervio óptico para el diagnóstico del glaucoma (153,154). Por otra parte, al incluir a pacientes con glaucoma precoz, es de esperar que los cambios detectables de reducción del RNR del disco óptico sean poco manifiestos en estas etapas de la enfermedad.

Por ello, no se han incluido como datos de entrada al modelo los parámetros obtenidos mediante el estudio con OCT del disco óptico.

Al omitir los parámetros del campo visual como datos de entrada para la predicción conseguimos una aplicabilidad más extensa del modelo, al poder emplearse en pacientes en los que los campos visuales no sean valorables por su baja fiabilidad o impracticables por alguna deficiencia física o cognitiva.

4. ANALISIS DE AGRUPACIONES EN LOS DATOS DE ENTRADA.

El análisis de los datos mediante el algoritmo no supervisado k-medias ha revelado que el número óptimo de agrupaciones en las que podemos dividir los datos de entrada ha sido de 2 particiones. Este hallazgo nos ratifica dos hechos:

1. Los datos contienen suficiente información almacenada en los parámetros de grosor de la CFNR como para permitir una subdivisión en dos grupos disjuntos.
2. El número de conjuntos óptimo es de dos grupos, que idealmente deberían de representar a los grupos de sujetos sanos y pacientes con glaucoma. Sería de esperar que uno de los grupos contuviera mayoritariamente a sujetos sanos entre sus elementos, y el otro grupo a pacientes con glaucoma precoz.

La significación de estos resultados da validez a los procedimientos posteriores del presente estudio. Los datos que serán utilizados para construir el clasificador han mostrado ser divisibles de manera óptima en exactamente dos grupos. Un resultado superior a dos grupos indicaría la existencia de posibles particiones de los casos en tres o más grupos, lo cual no es consistente con una clasificación binaria como la que buscamos en el presente trabajo.

Por otra parte, y en aras de controlar el problema del sobreajuste que aparece en todos los modelos creados mediante técnicas de minería de datos, se ha utilizado la metodología de agrupamiento no supervisado para detectar la presencia de datos aberrantes o anómalos. Para ello se ha hecho uso de la distancia euclídea entre los elementos de cada partición y su elemento central o centroide. Los casos en los que la distancia fue superior al valor medio de las distancias más dos desviaciones estándar fueron catalogados como casos extremos y eliminados de la base de datos, siguiendo el método propuesto por Duan et al (155). El número de casos extremos ha sido de 8 observaciones, que corresponde a un 2.4% de la base de datos. Esto apoya la adecuación de los datos recogidos para crear el clasificador.

5. ANALISIS DEL MODELO DE CLASIFICACION XGB.

Una vez excluidos los casos extremos, la base de datos ha sido dividida de manera aleatoria en un grupo de entrenamiento y un grupo de test o de validación, con la proporción de 70% y 30 % respectivamente. Este reparto nos permitirá disponer de datos no presentados al sistema durante la fase de entrenamiento, a partir de los que estimaremos el rendimiento del modelo. Buscamos con ello conocer el grado de generalización de nuestro modelo. Esta separación de la base de datos inicial en los grupos de entrenamiento y validación es una técnica habitual en la minería de datos. La proporción que se aconseja para constituir cada grupo no está claramente definida, aunque lo más habitual es un grupo de entrenamiento mayor que el grupo test y con una relación entre 6:4 y 8:2 entre un grupo y otro, en función del tamaño de la base de datos original.

La técnica de validación cruzada es otro método de estimación del rendimiento de un modelo utilizado ampliamente. En este procedimiento se selecciona aleatoriamente un porcentaje de los elementos de la base de datos, habitualmente un 90%, que es utilizado para entrenar el modelo, reservando el 10% restante para calcular las métricas de rendimiento. Esta operación se repite un número determinado de veces, eligiendo en cada ocasión un grupo aleatorio distinto para entrenar y validar el sistema. La media de

las métricas obtenidas en cada vuelta o “pliegue” es tomada como la estimación válida del rendimiento del modelo.

En nuestro estudio, hemos optado por la creación de un conjunto de datos de entrenamiento y otro de validación. Si bien la validación cruzada es un método con validez reconocida para determinar la eficiencia de un modelo, el método empleado en nuestro estudio muestra un potencial mayor para verificar la capacidad de generalización en la población general. Como contrapartida, la base de datos original empleada para el entrenamiento se reduce en el número de casos que contiene, aunque consideramos que al disponer de 320 observaciones originales alcanza un tamaño suficiente como para utilizar esta metodología.

Juzgamos importante verificar que la proporción de pacientes con glaucoma y controles es similar en los grupos aleatorios de entrenamiento y test, ya que una diferencia en su frecuencia relativa podría llevar a resultados irregulares en el cálculo del rendimiento del modelo. El test de chi-cuadrado no encontró diferencias significativas en la proporción de sujetos sanos y glaucoma entre estos grupos.

La creación del modelo XGB basado en árboles de decisión se ha llevado a cabo utilizando el paquete de software XGBoost para lenguaje Python, ajustando sus hiperparámetros manualmente para conseguir la mayor precisión posible con los datos de entrenamiento. La evaluación del modelo creado con los datos de test prueba una alta precisión del modelo, clasificando correctamente al 83,3% de los casos nuevos, con un AROC de 0.828. Estudios previos utilizando los parámetros de la OCT para la creación de clasificadores con técnicas de aprendizaje automático han obtenido unos valores de AROC que oscilan entre 0.821 y 0.981 (119–123,125,128). No hay, sin embargo, un criterio homogéneo en estos trabajos en cuanto a la fase precoz o tardía de glaucoma de los pacientes incluidos. Esto justifica el gran rango de valores de AROC reportado por los distintos autores. En nuestro estudio, consideramos que una predicción acertada en el 83,3% con una especificidad de 0.88 muestra un alto rendimiento de nuestro clasificador XGB.

El modelo XGB como técnica de minería de datos ha sido utilizado sólo en muy contadas ocasiones en estudios previos en medicina, pese a haber demostrado una alta

validez para la resolución de problemas de clasificación. Hasta el momento ha sido aplicado fundamentalmente en estudios genéticos o moleculares (156,157), y muy rara vez en estudios de contenido oftalmológico (158,159). XGBoost ha sido el ganador en repetidas ocasiones en las plataformas de competiciones de minería de datos sobre datos científicos tales como Kaggle (<http://www.kaggle.com/>).

Si bien el modelo XGB que hemos alcanzado tiene un buen rendimiento en la clasificación, el hecho de tratarse de una combinación de clasificadores más simples fusionados mediante Boosting lo convierte en un modelo críptico, no siendo factible conocer de manera directa cual son las operaciones, interacciones o apreciaciones internas que realiza el algoritmo para llegar a la predicción. No obstante, el algoritmo XGBoost proporciona una extensión que permite cierta interpretación del modelo a través del análisis de la importancia relativa de las variables de entrada al sistema. Aplicando esta extensión hemos obtenido una gradación jerárquica de los parámetros de la CFNR según su importancia relativa en la predicción. Los parámetros que muestran un mayor peso en el clasificador son el grosor de la CFNR en el cuadrante inferior, el valor promedio de la CFNR peripapilar y el grosor en el sector horario H6.

Utilizando un árbol de clasificación, Huang et al (160) encuentra que el cuadrante inferior de la CFNR es el parámetro con mayor importancia en el clasificador. En el trabajo de Medeiros et al (71) la mayor potencia discriminadora corresponde al cuadrante inferior. Baskaran et al (161) reporta que el grosor promedio y en el cuadrante inferior discrimina los pacientes con glaucoma mejor que cualquier otro parámetro individual de la CFNR. Nuestros resultados son acordes a la literatura, preponderando al cuadrante inferior y al valor promedio de la CFNR como los parámetros más significativos para la clasificación entre glaucoma y sujetos sanos. Bizios et al (122) compara el método de aprendizaje automático de máquina de vectores soporte con RNA en el diagnóstico del glaucoma. Haciendo uso de los valores de AROC, establece que la secuencia de mayor a menor AROC corresponde a los parámetros de grosor promedio de la CFNR, cuadrante superior, cuadrante inferior y sectores horarios H11 y H6.

6. ANALISIS DEL MODELO XGB MEDIANTE LOS VALORES SHAPLEY.

El cálculo de los valores Shapley nos abre una nueva ventana dentro de la caja negra del modelo, facultándonos para explicar el funcionamiento del algoritmo con mayor detalle.

El valor Shapley de cada parámetro de la CFNR ha sido calculado para cada paciente de los datos de entrenamiento. Su valor para un parámetro (ej. cuadrante inferior) nos indica el impacto o influencia que tiene ese parámetro en la predicción final del modelo.

La media de los valores Shapley en valor absoluto para cada parámetro de la CFNR está asociado al impacto que ese parámetro tiene en la predicción del modelo. Un valor medio más alto indica que el parámetro tiene un impacto mayor. Siguiendo este criterio, los parámetros han sido organizados en orden decreciente según su impacto en el resultado del clasificador, estimado por su valor Shapley. El grosor en el cuadrante inferior es el parámetro con mayor influencia en el clasificador, seguido de los sectores horarios H1 y H6.

Comparando esta ordenación de los parámetros de la CFNR con la obtenida según su importancia relativa en el modelo XGB, observamos que los 6 primeros parámetros más relevantes en ambas escalas coinciden plenamente. Estos parámetros son:

- Cuadrante inferior.
- Sector horario H1.
- Sector horario H6.
- Cuadrante superior.
- Sector horario H2.
- Valor promedio de la CFNR peripapilar.

7. ANALISIS DE DEPENDENCIA DE LOS PARAMETROS RESPECTO AL VALOR SHAPLEY.

La evaluación del impacto de los parámetros de la CFNR en el resultado del clasificador puede estudiarse más a fondo mediante la representación en un gráfico bidimensional de

la correspondencia existente entre el grosor de la CFNR en un sector determinado y su valor Shapley. Estos gráficos son denominados análisis de dependencia.

En el análisis del grosor de la CFNR en el cuadrante inferior respecto a sus valores Shapley se aprecia que el impacto del grosor en el clasificador no es lineal. Grosos inferiores a 95 μm se asocian a un valor Shapley positivo estable. Ocurre igual con los grosos mayores a 120 μm , correlacionados con un valor Shapley negativo prácticamente invariable. Este hallazgo nos lleva a concluir la existencia de intervalos de grosor en los que la influencia en la predicción es fija. Tendrá la misma influencia en el resultado del clasificador un grosor de 50 μm o de 95 μm . Igual ocurriría con un grosor de 120 μm y otro de 140 μm . Haciendo uso de este hallazgo, la variable de grosor en el cuadrante inferior se ha convertido en una variable con tres categorías, según los intervalos de “menor a 95 μm ”, “entre 95 y 120 μm ” y “superior a 120 μm ”.

Del análisis de dependencia en el grosor del sector horario H6 se desprende que en este caso la frontera entre las mesetas del gráfico se sitúan en el grosor de 100 μm . El análisis de agrupación establece dos particiones separadas por el nivel de 100 μm . Al igual que en el cuadrante inferior, la variable se ha convertido de cuantitativa a cualitativa ordinal, con división entre los grupos en el grosor de 100 μm .

En el estudio del impacto del grosor en el sector horario H2 mediante este análisis apreciamos un comportamiento inusual. Los grosos más delgados, especialmente por debajo de 60 μm , se correlacionan con un valor Shapley negativo, y por lo tanto, con una tendencia a que la clase que predice el modelo sea de sujeto control. Por el contrario, un mayor grosor se corresponde con una mayor probabilidad de que la predicción sea de paciente con glaucoma. Este hallazgo es opuesto al conocimiento etiopatogénico y clínico previo sobre la enfermedad del glaucoma, en la que a lo largo de la evolución de la enfermedad ocurre un adelgazamiento progresivo de la CFNR por pérdida de las células ganglionares de la retina (37). El cuadrante superior tiene un gráfico con distribución similar, aunque muestra mucha menor dispersión de las observaciones. No comprendemos enteramente la significación de este descubrimiento. La interacción con algún otro parámetro de la CFNR podría dar lugar a este comportamiento. Al incluir únicamente a pacientes con formas precoces de glaucoma,

el adelgazamiento de la CFNR que acontece con la enfermedad podría ser irregular y más tardío en este sector horario H2. Por otra parte, el hecho de que esta variable no mostrara una distribución gaussiana podría contribuir a este hallazgo. En cualquier caso, consideramos que esta peculiaridad observada en el grosor del sector horario H2, ubicado en el área nasal del disco óptico, debe ser confirmada en estudios posteriores. Dada la incoherencia de esta relación entre el sector horario H2 y la predicción de glaucoma, y hasta tener datos adicionales de otros estudios que confirmen o desmientan esta observación, hemos tomado la decisión de excluir a este parámetro para la construcción del árbol de decisión, paso final de la presente tesis.

En la representación del cuadrante superior llama la atención la morfología en “V” de la gráfica de la relación entre grosor y valor Shapley. Semejante a los que ocurre con el sector horario H2, valores de grosor en el cuadrante menores a 100 μm se correlacionan con valores Shapley negativos, prediciendo por tanto una tendencia menor a que la observación sea clasificada como glaucoma. No obstante, no encontramos otros datos que nos hagan dudar de la validez de los datos de grosor en este cuadrante en nuestro banco de datos. Al contrario que con el grosor en el sector horario H2, en el cuadrante superior el grosor muestra una distribución ajustada a la distribución normal. Por otra parte, se ha demostrado que la reproducibilidad de la medida de la CFNR es mayor en las medidas de grosor en los cuadrantes que en los sectores horarios, siendo a su vez inferior a la reproducibilidad del valor promedio de la CFNR. Por ello, valoramos que es importante no excluir a este cuadrante superior como variable de entrada para las fases posteriores de este estudio.

El análisis de dependencia del grosor promedio de la CFNR peripapilar tiene un comportamiento semejante al del cuadrante inferior, con dos zonas en meseta y una zona central intermedia, definidas por los límites de 78 μm y 92 μm de grosor. La variable se ha convertido a forma ordinal, siguiendo un razonamiento semejante a la discretización del resto de sectores.

Haciendo uso de los resultados del análisis de dependencia, los parámetros de la CFNR han sido transformados de variable continua a variables discretas, utilizando como límite para crear los intervalos las fronteras observadas entre los grupos del análisis de

dependencia. Esta transformación no es baladí, ya que al estar fundamentada en los hallazgos del clasificador XGB, permitirá controlar parte de las debilidades del árbol de decisión creado como última etapa de nuestro trabajo.

Este tipo de análisis aplicado a los parámetros de la CFNR, hasta lo que es conocido por el autor del presente estudio, no ha sido empleado hasta la fecha, por lo que no podemos comparar los resultados con trabajos previos.

8. CLASIFICADOR DE ALTA INTERPRETABILIDAD: ARBOL DE DECISION.

Tomando como entradas los cinco parámetros que han mostrado mayor importancia relativa en el análisis del clasificador, tras excluir el sector horario H2 por su comportamiento anómalo en el análisis de dependencia, se ha llevado a cabo la construcción de un árbol de clasificación, modelo caracterizado por tener una alta interpretabilidad directa.

La selección de las variables de entrada al árbol, eliminando aquellas variables con un comportamiento discrepante al conocimiento previo de la enfermedad el glaucoma, así como la discretización de los parámetros de la CFNR, nos permiten controlar el sobreajuste y la falta de generalización propio de un clasificador débil como es el árbol de decisión.

La profundidad del árbol ha sido controlada para evitar el sobreajuste. Se ha elegido la mínima profundidad que permite una mejoría en el ajuste del árbol a los datos test.

Las variables que forman parte del árbol de decisión son las siguientes:

- Grosor en cuadrante inferior.
- Grosor promedio.
- Grosor en sector horario H1.

La interpretación del árbol de clasificación es directa y sencilla, siendo fácilmente aplicable en la práctica clínica habitual. La predicción del clasificador viene dada por la

clase del nodo hoja al que se llega siguiendo las ramas del árbol, según el resultado de las comparaciones practicadas en cada nodo interno.

Un grosor en el cuadrante inferior superior a 120 μm es un buen predictor de ausencia de glaucoma precoz. Si además, el grosor en el sector horario H1 es superior a 96 μm , la probabilidad de glaucoma precoz es únicamente de un 4%. En caso contrario, con grosores de sector horario H1 inferiores a 96 μm , la probabilidad de glaucoma es de cerca del 20%.

Un grosor en el cuadrante inferior menor a 95 μm , junto con grosores en el sector horario H1 menores a 96 μm , hace el diagnóstico de glaucoma sumamente probable. Si a ello se suma un grosor promedio de la CFNR por debajo de 78 μm , la totalidad de los casos se predicen como glaucoma. Cuando el sector horario H1 tiene grosores mayores a 96 μm , si bien se predice la presencia de glaucoma, la probabilidad de enfermedad alcanza sólo el nivel del 66%.

Los valores de grosor de la CFNR en el cuadrante inferior entre 95 y 120 μm nos llevan a predicciones con mayor índice de error. Al combinarse con un grosor promedio de la CFNR peripapilar mayor a 78 μm , la predicción que realiza es de ausencia de glaucoma, aunque la probabilidad calculada de sufrir glaucoma precoz en esta situación es de un 41%. Si el grosor promedio es menor a 78 μm , el resultado es presencia de glaucoma precoz, con una probabilidad del 83 %. En este último caso, el nodo hoja incluye únicamente un 2,6% de las observaciones, por lo que en este caso el resultado del clasificador consideramos que es poco concluyente.

El árbol de decisión construido como terminación del presente estudio tiene una tasa de clasificación correcta de los casos del grupo de validación de un 83,3%, semejante al obtenido con el clasificador XGB, pese a que la complejidad del árbol es mucho menor. El AROC calculado es 0.821, y la especificidad alcanza el valor de 0.96.

La literatura dispone de muy pocos estudios en los que se ha recurrido a árboles de decisión para la resolución del dilema del diagnóstico en el glaucoma. Huang et al (160) recurre a un árbol de decisión en el diagnóstico del glaucoma entrenándolo con los grosores obtenidos con OCT. Incluye como variables de entrada tanto parámetros de la

cabeza del nervio óptico como de la CFNR, obteniendo una precisión del 73% de observaciones correctamente clasificadas, con una especificidad el 83%. El tamaño de la base de datos que utiliza es más reducido que el de nuestro estudio, incluyendo únicamente 64 pacientes con glaucoma y 71 sujetos controles, por lo que pensamos que nuestros resultados tienen mayor fiabilidad. En un trabajo posterior, el mismo autor recurre a un árbol de decisión para la detección del glaucoma, pero utiliza los parámetros obtenidos con GDx VCC, por lo que no es confrontable a nuestro estudio. Badalá et al (162) construyen un árbol de clasificación introduciendo medidas de OCT y de HRT en el modelo para el diagnóstico del glaucoma, por lo que sus resultados tampoco son asimilables a nuestro modelo.

La debilidad que supone el sobreajuste de nuestro árbol de decisión ha sido manejado a través de:

1. Eliminación de las observaciones extremas o aberrantes.
2. Selección de los parámetros con mayor importancia en la predicción de glaucoma en el modelo de alta complejidad XGB, siguiendo un doble criterio:
 - Nivel de importancia relativa.
 - Impacto en la predicción evaluado mediante el valor Shapley.
3. Estudio de la correlación entre grosor en aquellos sectores de la CFNR con mayor impacto en el clasificador y la predicción de glaucoma, excluyendo aquellos que no muestran un comportamiento consistente con el conocimiento etiopatogénico del glaucoma en el momento actual.
4. Discretización de las variables de grosor de la CFNR, transformándolas a forma cualitativa ordinal.

A partir de la estructura del árbol podemos derivar una lista de reglas de aplicación directa en la práctica oftalmológica diaria, tales como:

- Un grosor en el cuadrante inferior menor a 95 μm sumado a un grosor en el sector horario H1 inferior a 96 μm establecen una probabilidad muy alta de glaucoma (probabilidad de glaucoma mayor a 95%)

- Un grosor en el cuadrante inferior mayor a 120 μm se asocia a una baja probabilidad de glaucoma en estadio temprano (probabilidad de glaucoma: 25%).
- Un grosor en el cuadrante inferior mayor a 120 μm junto a un grosor en el sector horario H1 superior a 96 μm hacen sumamente improbable la predicción de glaucoma (riesgo de glaucoma menor a 5%).

CONCLUSIONES

1. La aplicación de técnicas de minería de datos en nuestro estudio nos ha permitido crear un modelo capaz de predecir, analizando los valores de grosor de la capa de fibras nerviosas retinianas (CFNR) peripapilar, la presencia de glaucoma en estadio precoz con una tasa de acierto del 83%.
2. El análisis de la importancia relativa de los parámetros de la CFNR en el modelo obtenido en nuestro estudio revela que los sectores con mayor impacto en el diagnóstico de glaucoma precoz son el grosor en el cuadrante inferior, los sectores horarios H1, H6 y H2, cuadrante superior y grosor promedio de la CFNR peripapilar.
3. Los resultados de nuestro estudio nos permiten afirmar que el grosor de la CFNR en el cuadrante inferior es el parámetro con más relevancia para predecir el diagnóstico de glaucoma en estadio precoz.
4. El análisis de la correlación existente entre los valores Shapley calculados en nuestro modelo predictivo y el grosor de la CFNR en el cuadrante inferior muestra que un menor grosor en este segmento se asocia a una mayor probabilidad de diagnóstico de glaucoma precoz, siendo consistente con las teorías existentes de la etiopatogenia del glaucoma. Esta correspondencia inversa entre grosor y predicción de glaucoma ha sido observada también para los segmentos horarios H1 y H6 así como en el grosor promedio de la CFNR peripapilar.
5. El análisis de la correlación presente entre los valores Shapley calculados en nuestro modelo y el grosor en el sector horario H2 y en el cuadrante superior evidencia que un menor grosor en estos segmentos de la CFNR se corresponde con una menor probabilidad de diagnóstico de glaucoma precoz. Este hallazgo es inverso a lo que sería esperable de acuerdo a los estudios publicados acerca de la enfermedad del glaucoma.

6. Los resultados de nuestro estudio muestran una distribución no gaussiana de los valores de grosor en el sector horario H2, por lo que se ha decidido excluir este sector de los análisis posteriores de nuestro trabajo.
7. El gráfico de correlación entre valores Shapley y el grosor de la CFNR en el cuadrante inferior revela que el impacto en la predicción de glaucoma precoz es uniforme en los intervalos de grosor “menor a 95 μm ”, “entre 95 y 120 μm ” y “mayor de 120 μm ” de este cuadrante.
8. El gráfico de correlación entre valores Shapley y el grosor de la CFNR en el sector horario H1 muestra que la influencia para la predicción de glaucoma precoz es similar en los intervalos de grosor “menor o igual a 96 μm ” y “mayor de 96 μm ” de este sector horario.
9. El gráfico de correlación entre valores Shapley y el grosor promedio de la CFNR indica que el peso en la predicción de glaucoma precoz es semejante en los intervalos de grosor “menor a 78 μm ”, “entre 78 y 92 μm ” y “mayor de 92 μm ”.
10. En base a las predicciones de nuestro modelo, la combinación de un grosor superior a 120 μm en el cuadrante inferior junto con un grosor mayor a 96 μm en el sector horario H1 es un excelente predictor de ausencia de glaucoma precoz, con un 96% de clasificación correcta de observaciones. Por otra parte, un grosor en el cuadrante inferior menor a 95 μm junto a un grosor en el sector horario H1 menor a 96 μm predice en más de un 90% de los casos la presencia de glaucoma en estadio precoz.
11. Los resultados de nuestro estudio proporcionan fórmulas para el diagnóstico del glaucoma precoz a partir de los grosores de la CFNR con una utilidad clínica directa y amplia aplicabilidad. Para su uso en la práctica oftalmológica sería necesario desarrollar una herramienta informática de fácil uso, que podría incluirse dentro del software del instrumental de diagnóstico en oftalmología. Sería conveniente validar el modelo desarrollado con nuevas y más amplias bases de datos a fin de conseguir una mayor robustez.

BIBLIOGRAFIA

1. European Glaucoma Society Terminology and Guidelines for Glaucoma, 4th Edition. *Br J Ophthalmol*. 2017 May 1;101(5):73 LP – 127.
2. Prum BEJ, Rosenberg LF, Gedde SJ, Mansberger SL, Stein JD, Moroi SE, et al. Primary Open-Angle Glaucoma Preferred Practice Pattern Guidelines. *Ophthalmology*. 2016 Jan;123(1):P41–111.
3. Weinreb RN, Aung T, Medeiros FA. The pathophysiology and treatment of glaucoma: a review. *JAMA*. 2014 May;311(18):1901–11.
4. Hyman L, Wu SY, Connell AM, Schachat A, Nemesure B, Hennis A, et al. Prevalence and causes of visual impairment in The Barbados Eye Study. *Ophthalmology*. 2001 Oct;108(10):1751–6.
5. Resnikoff S, Pascolini D, Etya'ale D, Kocur I, Pararajasegaram R, Pokharel GP, et al. Global data on visual impairment in the year 2002. *Bull World Health Organ*. 2004 Nov;82(11):844–51.
6. Cedrone C, Nucci C, Scuderi G, Ricci F, Cerulli A, Culasso F. Prevalence of blindness and low vision in an Italian population: a comparison with other European studies. *Eye* . 2006 Jun;20(6):661–7.
7. Gordon MO, Beiser JA, Brandt JD, Heuer DK, Higginbotham EJ, Johnson CA, et al. The Ocular Hypertension Treatment Study: baseline factors that predict the onset of primary open-angle glaucoma. *Arch Ophthalmol*. 2002 Jun;120(6):714–30.
8. Anderson DR. Collaborative normal tension glaucoma study. *Curr Opin Ophthalmol*. 2003 Apr;14(2):86–90.
9. Heijl A, Leske MC, Bengtsson B, Hyman L, Bengtsson B, Hussein M. Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Arch Ophthalmol* . 2002 Oct;120(10):1268–79.
10. Quigley HA, Broman AT. The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol*. 2006 Mar;90(3):262–7.

11. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: A systematic review and meta-analysis. *Ophthalmology*. 2014;121(11):2081–90.
12. Cedrone C, Mancino R, Cerulli A, Cesareo M, Nucci C. Epidemiology of primary glaucoma: prevalence, incidence, and blinding effects. *Prog Brain Res*. 2008;173:3–14.
13. Friedman DS, Wolfs RCW, O’Colmain BJ, Klein BE, Taylor HR, West S, et al. Prevalence of open-angle glaucoma among adults in the United States. *Arch Ophthalmol*. 2004 Apr;122(4):532–8.
14. Antón A, Andrada MT, Mujica V, Calle MA, Portela J, Mayo A. Prevalence of primary open-angle glaucoma in a Spanish population: the Segovia study. *J Glaucoma*. 2004 Oct;13(5):371–6.
15. Friedman DS, Jampel HD, Muñoz B, West SK. The prevalence of open-angle glaucoma among blacks and whites 73 years and older: the Salisbury Eye Evaluation Glaucoma Study. *Arch Ophthalmol* . 2006 Nov;124(11):1625–30.
16. Casson RJ, Newland HS, Muecke J, McGovern S, Abraham L, Shein WK, et al. Prevalence of glaucoma in rural Myanmar: the Meiktila Eye Study. *Br J Ophthalmol*. 2007 Jun;91(6):710–4.
17. Song W, Shan L, Cheng F, Fan P, Zhang L, Qu W, et al. Prevalence of glaucoma in a rural northern china adult population: a population-based survey in kailu county, inner mongolia. *Ophthalmology*. 2011 Oct;118(10):1982–8.
18. Shaikh Y, Yu F, Coleman AL. Burden of undetected and untreated glaucoma in the United States. *Am J Ophthalmol*. 2014 Dec;158(6):1121-1129.
19. Tielsch JM, Sommer A, Katz J, Royall RM, Quigley HA, Javitt J. Racial variations in the prevalence of primary open-angle glaucoma. The Baltimore Eye Survey. *JAMA*. 1991 Jul;266(3):369–74.
20. Wensor MD, McCarty CA, Stanislavsky YL, Livingston PM, Taylor HR. The

- prevalence of glaucoma in the Melbourne Visual Impairment Project. *Ophthalmology*. 1998 Apr;105(4):733–9.
21. Mitchell P, Smith W, Attebo K, Healey PR. Prevalence of open-angle glaucoma in Australia. The Blue Mountains Eye Study. *Ophthalmology*. 1996 Oct;103(10):1661–9.
 22. Topouzis F, Coleman AL, Harris A, Koskosas A, Founti P, Gong G, et al. Factors associated with undiagnosed open-angle glaucoma: the Thessaloniki Eye Study. *Am J Ophthalmol*. 2008 Feb;145(2):327–35.
 23. Quigley HA, West SK, Rodriguez J, Munoz B, Klein R, Snyder R. The prevalence of glaucoma in a population-based study of Hispanic subjects: Proyecto VER. *Arch Ophthalmol*. 2001 Dec;119(12):1819–26.
 24. Varma R, Ying-Lai M, Francis BA, Nguyen BB-T, Deneen J, Wilson MR, et al. Prevalence of open-angle glaucoma and ocular hypertension in Latinos: the Los Angeles Latino Eye Study. *Ophthalmology*. 2004 Aug;111(8):1439–48.
 25. Bathija R, Gupta N, Zangwill L, Weinreb RN. Changing definition of glaucoma. *J Glaucoma*. 1998 Jun;7(3):165–9.
 26. Tuulonen A, Airaksinen PJ, Erola E, Forsman E, Friberg K, Kaila M, et al. The Finnish evidence-based guideline for open-angle glaucoma. *Acta Ophthalmol Scand*. 2003 Feb;81(1):3–18.
 27. Ringsdorf L, McGwin GJ, Owsley C. Visual field defects and vision-specific health-related quality of life in African Americans and whites with glaucoma. *J Glaucoma*. 2006 Oct;15(5):414–8.
 28. Parrish RK 2nd, Gedde SJ, Scott IU, Feuer WJ, Schiffman JC, Mangione CM, et al. Visual function and quality of life among patients with glaucoma. *Arch Ophthalmol*. 1997 Nov;115(11):1447–55.
 29. Mills RP. Correlation of quality of life with clinical symptoms and signs at the time of glaucoma diagnosis. *Trans Am Ophthalmol Soc*. 1998;96:753–812.

30. Jampel HD. Glaucoma patients' assessment of their visual function and quality of life. *Trans Am Ophthalmol Soc.* 2001;99:301–17.
31. Ramulu P. Glaucoma and disability: which tasks are affected, and at what stage of disease? *Curr Opin Ophthalmol.* 2009 Mar;20(2):92–8.
32. Skalicky SE, Goldberg I. Are we ready to assess quality of life routinely in our glaucoma patients? *Bulletin de la Societe belge d'ophtalmologie.* Belgium; 2010. p. 5–7.
33. Viswanathan AC, McNaught AI, Poinosawmy D, Fontana L, Crabb DP, Fitzke FW, et al. Severity and stability of glaucoma: patient perception compared with objective measurement. *Arch Ophthalmol.* 1999 Apr;117(4):450–4.
34. Turano KA, Rubin GS, Quigley HA. Mobility performance in glaucoma. *Invest Ophthalmol Vis Sci.* 1999 Nov;40(12):2803–9.
35. Haymes SA, Leblanc RP, Nicolela MT, Chiasson LA, Chauhan BC. Risk of falls and motor vehicle collisions in glaucoma. *Invest Ophthalmol Vis Sci.* 2007 Mar;48(3):1149–55.
36. Kotecha A, O'Leary N, Melmoth D, Grant S, Crabb DP. The functional consequences of glaucoma for eye-hand coordination. *Invest Ophthalmol Vis Sci.* 2009 Jan;50(1):203–13.
37. Weinreb RN, Khaw PT. Primary open-angle glaucoma. *Lancet (London, England).* 2004 May;363(9422):1711–20.
38. Huang D, Swanson EA, Lin CP, Schuman JS, Stinson WG, Chang W, et al. Optical coherence tomography. *Science.* 1991 Nov;254(5035):1178–81.
39. Buabbud JC, Al-latayfeh MM, Sun JK. Optical coherence tomography imaging for diabetic retinopathy and macular edema. *Curr Diab Rep.* 2010 Aug;10(4):264–9.
40. Domalpally A, Blodi BA, Scott IU, Ip MS, Oden NL, Lauer AK, et al. The

- Standard Care vs Corticosteroid for Retinal Vein Occlusion (SCORE) study system for evaluation of optical coherence tomograms: SCORE study report 4. *Arch Ophthalmol*. 2009 Nov;127(11):1461–7.
41. Li G, Fansi AK, Boivin J-F, Joseph L, Harasymowycz P. Screening for glaucoma in high-risk populations using optical coherence tomography. *Ophthalmology*. 2010 Mar;117(3):453–61.
 42. Sakata LM, Deleon-Ortega J, Sakata V, Girkin CA. Optical coherence tomography of the retina and optic nerve - a review. *Clin Experiment Ophthalmol*. 2009 Jan;37(1):90–9.
 43. Jaffe GJ, Caprioli J. Optical coherence tomography to detect and manage retinal disease and glaucoma. *Am J Ophthalmol*. 2004 Jan;137(1):156–69.
 44. Quigley HA, Dunkelberger GR, Green WR. Retinal ganglion cell atrophy correlated with automated perimetry in human eyes with glaucoma. *Am J Ophthalmol*. 1989 May;107(5):453–64.
 45. Kass MA, Heuer DK, Higginbotham EJ, Johnson CA, Keltner JL, Miller JP, et al. The Ocular Hypertension Treatment Study: a randomized trial determines that topical ocular hypotensive medication delays or prevents the onset of primary open-angle glaucoma. *Arch Ophthalmol* . 2002 Jun;120(6):701–30.
 46. Kerrigan-Baumrind LA, Quigley HA, Pease ME, Kerrigan DF, Mitchell RS. Number of ganglion cells in glaucoma eyes compared with threshold visual field tests in the same persons. *Invest Ophthalmol Vis Sci*. 2000 Mar;41(3):741–8.
 47. Mikelberg FS, Yidegiligne HM, Schulzer M. Optic nerve axon count and axon diameter in patients with ocular hypertension and normal visual fields. *Ophthalmology*. 1995 Feb;102(2):342–8.
 48. Quigley HA, Addicks EM, Green WR. Optic nerve damage in human glaucoma. III. Quantitative correlation of nerve fiber loss and visual field defect in glaucoma, ischemic neuropathy, papilledema, and toxic neuropathy. *Arch*

- Ophthalmol . 1982 Jan;100(1):135–46.
49. Ajtony C, Balla Z, Somoskeoy S, Kovacs B. Relationship between visual field sensitivity and retinal nerve fiber layer thickness as measured by optical coherence tomography. *Invest Ophthalmol Vis Sci.* 2007 Jan;48(1):258–63.
 50. Arintawati P, Sone T, Akita T, Tanaka J, Kiuchi Y. The applicability of ganglion cell complex parameters determined from SD-OCT images to detect glaucomatous eyes. *J Glaucoma.* 2013 Dec;22(9):713–8.
 51. Hood DC, Anderson SC, Wall M, Raza AS, Kardon RH. A test of a linear model of glaucomatous structure-function loss reveals sources of variability in retinal nerve fiber and visual field measurements. *Invest Ophthalmol Vis Sci.* 2009 Sep;50(9):4254–66.
 52. Gabriele ML, Ishikawa H, Wollstein G, Bilonick RA, Kagemann L, Wojtkowski M, et al. Peripapillary nerve fiber layer thickness profile determined with high speed, ultrahigh resolution optical coherence tomography high-density scanning. *Invest Ophthalmol Vis Sci.* 2007 Jul;48(7):3154–60.
 53. Heindl LM, Adler W, El-Malahi O, Schaub F, Hermann MM, Dietlein TS, et al. The Optimal Diameter for Circumpapillary Retinal Nerve Fiber Layer Thickness Measurement by SD-OCT in Glaucoma. *J Glaucoma.* 2018;27(12):1086–93.
 54. Han IC, Jaffe GJ. Evaluation of artifacts associated with macular spectral-domain optical coherence tomography. *Ophthalmology.* 2010 Jun;117(6):1177-1189.
 55. Ho J, Sull AC, Vuong LN, Chen Y, Liu J, Fujimoto JG, et al. Assessment of artifacts and reproducibility across spectral- and time-domain optical coherence tomography devices. *Ophthalmology.* 2009 Oct;116(10):1960–70.
 56. Mylonas G, Ahlers C, Malamos P, Golbaz I, Deak G, Schuetze C, et al. Comparison of retinal thickness measurements and segmentation performance of four different spectral and time domain OCT devices in neovascular age-related macular degeneration. *Br J Ophthalmol.* 2009 Nov;93(11):1453–60.

57. Sadda SR, Wu Z, Walsh AC, Richine L, Dougall J, Cortez R, et al. Errors in retinal thickness measurements obtained by optical coherence tomography. *Ophthalmology*. 2006 Feb;113(2):285–93.
58. Leung CK-S, Cheung CY-L, Weinreb RN, Qiu Q, Liu S, Li H, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: a variability and diagnostic performance study. *Ophthalmology*. 2009 Jul;116(7):1257–63.
59. Cheung CYL, Leung CKS, Lin D, Pang C-P, Lam DSC. Relationship between retinal nerve fiber layer measurement and signal strength in optical coherence tomography. *Ophthalmology*. 2008 Aug;115(8):1347–51, 1351.e1-2.
60. Mwanza J-C, Chang RT, Budenz DL, Durbin MK, Gendy MG, Shi W, et al. Reproducibility of peripapillary retinal nerve fiber layer thickness and optic nerve head parameters measured with cirrus HD-OCT in glaucomatous eyes. *Invest Ophthalmol Vis Sci*. 2010 Nov;51(11):5724–30.
61. González-García AO, Vizzeri G, Bowd C, Medeiros FA, Zangwill LM, Weinreb RN. Reproducibility of RTVue retinal nerve fiber layer thickness and optic disc measurements and agreement with Stratus optical coherence tomography measurements. *Am J Ophthalmol*. 2009 Jun;147(6):1067–74..
62. Kim JS, Ishikawa H, Sung KR, Xu J, Wollstein G, Bilonick RA, et al. Retinal nerve fibre layer thickness measurement reproducibility improved with spectral domain optical coherence tomography. *Br J Ophthalmol*. 2009 Aug;93(8):1057–63.
63. Budenz DL, Michael A, Chang RT, McSoley J, Katz J. Sensitivity and specificity of the StratusOCT for perimetric glaucoma. *Ophthalmology*. 2005 Jan;112(1):3–9.
64. Zhang X, Francis BA, Dastiridou A, Chopra V, Tan O, Varma R, et al. Longitudinal and Cross-Sectional Analyses of Age Effects on Retinal Nerve Fiber Layer and Ganglion Cell Complex Thickness by Fourier-Domain OCT.

- Transl Vis Sci Technol. 2016 Mar;5(2):1.
65. Begum VU, Addepalli UK, Yadav RK, Shankar K, Senthil S, Garudadri CS, et al. Ganglion cell-inner plexiform layer thickness of high definition optical coherence tomography in perimetric and preperimetric glaucoma. *Invest Ophthalmol Vis Sci*. 2014 Jul;55(8):4768–75.
 66. Mwanza J-C, Durbin MK, Budenz DL, Sayyad FE, Chang RT, Neelakantan A, et al. Glaucoma diagnostic accuracy of ganglion cell-inner plexiform layer thickness: comparison with nerve fiber layer and optic nerve head. *Ophthalmology*. 2012 Jun;119(6):1151–8.
 67. Mwanza J-C, Budenz DL, Godfrey DG, Neelakantan A, Sayyad FE, Chang RT, et al. Diagnostic performance of optical coherence tomography ganglion cell-inner plexiform layer thickness measurements in early glaucoma. *Ophthalmology*. 2014 Apr;121(4):849–54.
 68. Nouri-Mahdavi K, Hoffman D, Tannenbaum DP, Law SK, Caprioli J. Identifying early glaucoma with optical coherence tomography. *Am J Ophthalmol*. 2004 Feb;137(2):228–35.
 69. Mwanza J-C, Warren JL, Budenz DL, Group S. Combining spectral domain optical coherence tomography structural parameters for the diagnosis of glaucoma with early visual field loss. *Invest Ophthalmol Vis Sci*. 2013 Dec;54(13):8393–400.
 70. Larrosa JM, Moreno-Montañés J, Martínez-de-la-Casa JM, Polo V, Velázquez-Villoria Á, Berrozpe C, et al. A Diagnostic Calculator for Detecting Glaucoma on the Basis of Retinal Nerve Fiber Layer, Optic Disc, and Retinal Ganglion Cell Analysis by Optical Coherence Tomography. *Invest Ophthalmol Vis Sci*. 2015 Oct;56(11):6788–95.
 71. Medeiros FA, Zangwill LM, Bowd C, Vessani RM, Susanna RJ, Weinreb RN. Evaluation of retinal nerve fiber layer, optic nerve head, and macular thickness measurements for glaucoma detection using optical coherence tomography. *Am J*

- Ophthalmol. 2005 Jan;139(1):44–55.
72. Güerri N, Polo V, Larrosa JM, Ferreras A, Fuertes I, Pablo LE. Performance of imaging devices versus optic disc and fiber layer photography in a clinical practice guideline for glaucoma diagnosis. *Eur J Ophthalmol.* 2012;22(4):554–62.
 73. Budenz DL, Anderson DR, Varma R, Schuman J, Cantor L, Savell J, et al. Determinants of normal retinal nerve fiber layer thickness measured by Stratus OCT. *Ophthalmology.* 2007 Jun;114(6):1046–52.
 74. Savini G, Zanini M, Carelli V, Sadun AA, Ross-Cisneros FN, Barboni P. Correlation between retinal nerve fibre layer thickness and optic nerve head size: an optical coherence tomography study. *Br J Ophthalmol.* 2005 Apr;89(4):489–92.
 75. Quigley HA, Coleman AL, Dorman-Pease ME. Larger optic nerve heads have more nerve fibers in normal monkey eyes. *Arch Ophthalmol.* 1991 Oct;109(10):1441–3.
 76. Varma R, Skaf M, Barron E. Retinal nerve fiber layer thickness in normal human eyes. *Ophthalmology.* 1996 Dec;103(12):2114–9.
 77. Garcia-Valenzuela E, Mori M, Edward DP, Shahidi M. Thickness of the peripapillary retina in healthy subjects with different degrees of ametropia. *Ophthalmology.* 2000 Jul;107(7):1321–7.
 78. Mwanza J-C, Oakley JD, Budenz DL, Anderson DR. Ability of cirrus HD-OCT optic nerve head parameters to discriminate normal from glaucomatous eyes. *Ophthalmology.* 2011 Feb;118(2):241-8.
 79. López-Peña MJ, Ferreras A, Polo V, Larrosa JM, Honrubia FM. [Relationship between standard automated perimetry and HRT, OCT and GDx in normal, ocular hypertensive and glaucomatous subjects]. *Arch Soc Esp Oftalmol.* 2007 Apr;82(4):197–208.

80. Jeoung JW, Choi YJ, Park KH, Kim DM. Macular ganglion cell imaging study: glaucoma diagnostic accuracy of spectral-domain optical coherence tomography. *Invest Ophthalmol Vis Sci.* 2013 Jul;54(7):4422–9.
81. Anton A, Moreno-Montañes J, Blázquez F, Alvarez A, Martín B, Molina B. Usefulness of optical coherence tomography parameters of the optic disc and the retinal nerve fiber layer to differentiate glaucomatous, ocular hypertensive, and normal eyes. *J Glaucoma.* 2007 Jan;16(1):1–8.
82. Na JH, Lee K, Lee JR, Baek S, Yoo SJ, Kook MS. Detection of macular ganglion cell loss in preperimetric glaucoma patients with localized retinal nerve fibre defects by spectral-domain optical coherence tomography. *Clin Experiment Ophthalmol.* 2013 Dec;41(9):870–80.
83. Orallo JH, Quintana MJR, Ramirez CF. *Introducción a la minería de datos.* Pearson Educación; 2004.
84. Molina Félix LC. *Data mining: torturando a los datos hasta que confiesen.* Univ Oberta Catalunya. 2002;
85. Frické MH. Data-Information-Knowledge-Wisdom (DIKW) Pyramid, Framework, Continuum. In: Schintler LA, McNeely CL, editors. *Encyclopedia of Big Data.* Springer International Publishing; 2018. p. 1–4.
86. IJCAI'95: Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2. In San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.; 1995.
87. Witten IH, Frank E, Hall MA. *Data Mining: Practical Machine Learning Tools and Techniques.* 3rd ed. Amsterdam: Morgan Kaufmann; 2011.
88. Fayyad UM, Piatetsky-Shapiro G, Smyth P, others. Knowledge Discovery and Data Mining: Towards a Unifying Framework. In: *KDD.* 1996. p. 82–8.
89. Adriaans P, Zantinge D. *Data Mining.* Addison Wesley Longman Limited. Edinburgh Gate, Harlow, Engl. 1996.

90. Pérez López C, Santín González D. Minería de datos. Técnicas y herramientas: técnicas y herramientas. Editorial Paraninfo; 2007.
91. Febles Rodríguez JP, González Pérez A. Aplicación de la minería de datos en la bioinformática. *Acimed*. 2002;10(2):69–76.
92. Millán DB, Boticario JG, Viñuela PI. Aprendizaje automático. Sanz y Torres; 2006.
93. Palma Méndez JT, Morales RM. Inteligencia artificial. Técnicas, métodos y aplicaciones. Ed McGraw–Hill, Madrid, España. 2008.
94. Breiman L, Friedman J, Stone CJ, Olshen RA. Classification and regression trees. CRC press; 1984.
95. Quinlan JR. C4. 5: programs for machine learning. Elsevier; 2014.
96. Breiman L. Bagging predictors. *Mach Learn*. 1996;24(2):123–40.
97. Friedman JH. Greedy function approximation: a gradient boosting machine. *Ann Stat*. 2001;1189–232.
98. Friedman J, Hastie T, Tibshirani R. Additive logistic regression: a statistical view of boosting. *Ann Stat*. 2000;28(2):337–407.
99. Breiman L. Some infinity theory for predictor ensembles. Technical Report 579, Statistics Dept. UCB; 2000.
100. Li P. Robust logitboost and adaptive base class (abc) logitboost. *arXiv Prepr arXiv12033491*. 2012.
101. Chen T, Guestrin C. Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016. p. 785–94.
102. Miller T. Explanation in artificial intelligence: Insights from the social sciences. *Artif Intell*. 2019;267:1–38.

103. Doshi-velez F, Kim B. Towards A Rigorous Science of Interpretable Machine Learning. 2017;(MI):1–13.
104. Ribeiro MT, Singh S, Guestrin C. Model-agnostic interpretability of machine learning. arXiv160605386. 2016;
105. Lundberg SM, Lee S-I. A unified approach to interpreting model predictions. In: Advances in neural information processing systems. 2017. p. 4765–74.
106. Lundberg S, Lee S. Shap. SHapley Addit Explanations. At: <https://github.com/slundberg/shap>. 2019;
107. Lundberg, S. M., & Lee, S. I. Consistent feature attribution for tree ensembles. *arXiv:1706.06060*. 2017.
108. Lundberg SM, Erion GG, Lee S-I. Consistent individualized feature attribution for tree ensembles. arXiv180203888. 2018.
109. Marsolo K, Twa M, Bullimore MA, Parthasarathy S. Spatial modeling and classification of corneal shape. *IEEE Trans Inf Technol Biomed.* 2007;11(2):203–12.
110. Schmidt GW, Broman AT, Hindman HB, Grant MP. Vision survival after open globe injury predicted by classification and regression tree analysis. *Ophthalmology.* 2008;115(1):202–9.
111. Quintana JM, Arostegui I, Alberdi T, Escobar A, Perea E, Navarro G, et al. Decision trees for indication of cataract surgery based on changes in visual acuity. *Ophthalmology.* 2010;117(8):1471–8.
112. Egejuru NC, Balogun JA, Mhambe PD, Asahiah FO, Idowu PA. Model for Prediction of Cataracts Using Supervised Machine Learning Algorithms. *Computing, Information Systems, Development Informatics.* 2017; 8(3):47-62.
113. Sacchetti M, Lambiase A, Mantelli F, Deligianni V, Leonardi A, Bonini S. Tailored approach to the treatment of vernal keratoconjunctivitis.

- Ophthalmology. 2010;117(7):1294–9.
114. Tufail A, Rudisill C, Egan C, Kapetanakis V V, Salas-Vega S, Owen CG, et al. Automated diabetic retinopathy image assessment software: diagnostic accuracy and cost-effectiveness compared with human graders. *Ophthalmology*. 2017;124(3):343–51.
 115. Abramoff MD, Folk JC, Han DP, Walker JD, Williams DF, Russell SR, et al. Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol*. 2013;131(3):351–7.
 116. Burlina PM, Joshi N, Pekala M, Pacheco KD, Freund DE, Bressler NM. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol*. 2017;135(11):1170–6.
 117. Schlegl T, Waldstein SM, Bogunovic H, Endstraßer F, Sadeghipour A, Philip A-M, et al. Fully automated detection and quantification of macular fluid in OCT using deep learning. *Ophthalmology*. 2018;125(4):549–58.
 118. Peng Y, Dharssi S, Chen Q, Keenan TD, Agrón E, Wong WT, et al. DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs. *Ophthalmology*. 2019;126(4):565–75.
 119. Huang M-L, Chen H-Y. Development and comparison of automated classifiers for glaucoma diagnosis using Stratus optical coherence tomography. *Invest Ophthalmol Vis Sci*. 2005;46(11):4121–9.
 120. Burgansky-Eliash Z, Wollstein G, Chu T, Ramsey JD, Glymour C, Noecker RJ, et al. Optical coherence tomography machine learning classifiers for glaucoma detection: a preliminary study. *Invest Ophthalmol Vis Sci*. 2005 Nov;46(11):4147–52.
 121. Naithani P, Sihota R, Sony P, Dada T, Gupta V, Kondal D, et al. Evaluation of

- optical coherence tomography and heidelberg retinal tomography parameters in detecting early and moderate glaucoma. *Invest Ophthalmol Vis Sci.* 2007;48(7):3138–45.
122. Bizios D, Heijl A, Hougaard JL, Bengtsson B. Machine learning classifiers for glaucoma diagnosis based on classification of retinal nerve fibre layer thickness parameters measured by Stratus OCT. *Acta Ophthalmol.* 2010 Feb;88(1):44–52.
123. An G, Omodaka K, Tsuda S, Shiga Y, Takada N, Kikawa T, et al. Comparison of machine-learning classification models for glaucoma management. *J Healthc Eng.* 2018; vol.2018:1-8
124. Kim SJ, Cho KJ, Oh S. Development of machine learning models for diagnosis of glaucoma. *PLoS One.* 2017;12(5):e0177726.
125. Barella KA, Costa VP, Gonçalves Vidotti V, Silva FR, Dias M, Gomi ES. Glaucoma Diagnostic Accuracy of Machine Learning Classifiers Using Retinal Nerve Fiber Layer and Optic Nerve Data from SD-OCT. *J Ophthalmol.* 2013; vol. 2013:789129.
126. Christopher M, Belghith A, Weinreb RN, Bowd C, Goldbaum MH, Saunders LJ, et al. Retinal nerve fiber layer features identified by unsupervised machine learning on optical coherence tomography scans predict glaucoma progression. *Invest Ophthalmol Vis Sci.* 2018;59(7):2748–56.
127. Shigueoka LS, Vasconcellos JPC de, Schimiti RB, Reis ASC, Oliveira GO de, Gomi ES, et al. Automated algorithms combining structure and function outperform general ophthalmologists in diagnosing glaucoma. *PLoS One.* 2018;13(12):e0207784.
128. Sugimoto K, Murata H, Hirasawa H, Aihara M, Mayama C, Asaoka R. Cross-sectional study: Does combining optical coherence tomography measurements using the ‘ Random Forest ’ decision tree classifier improve the prediction of the presence of perimetric deterioration in glaucoma suspects? 2013;1–7.

129. Ting DSW, Cheung CY-L, Lim G, Tan GSW, Quang ND, Gan A, et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *Jama*. 2017;318(22):2211–23.
130. Medeiros FA, Jammal AA, Thompson AC. From machine to machine: an OCT-trained deep learning algorithm for objective quantification of glaucomatous damage in fundus photographs. *Ophthalmology*. 2019;126(4):513–21.
131. Maetschke S, Antony B, Ishikawa H, Wollstein G, Schuman J, Garnavi R. A feature agnostic approach for glaucoma detection in OCT volumes. *PLoS One*. 2019;14(7):e0219126.
132. Zhang L, Devalla SK, Cheng C-Y, Milea D, Nongpiur ME, Mani B, et al. A Multi-device, Multi-ethnicity Deep Learning Algorithm to Detect Glaucoma from A Single Optical Coherence Tomography Scan of the Optic Nerve Head. *Invest Ophthalmol Vis Sci*. 2019;60(9):2207.
133. Wang P, Shen J, Chang R, Moloney M, Torres M, Burkemper B, et al. Machine learning models for diagnosing glaucoma from retinal nerve fiber layer thickness maps. *Ophthalmol Glaucoma*. 2019;2(6):422–8.
134. Hemelings R, Elen B, Barbosa-Breda J, Lemmens S, Meire M, Pourjavan S, et al. Accurate prediction of glaucoma from colour fundus images with a convolutional neural network that relies on active and transfer learning. *Acta Ophthalmol*. 2020;98(1):e94–100.
135. Hothorn T, Lausen B. Bagging tree classifiers for laser scanning images: a data- and simulation-based strategy. *Artif Intell Med*. 2003;27(1):65–79.
136. Mikelberg FS, Parfitt CM, Swindale N V, Graham SL, Drance SM, Gosine R. Ability of the Heidelberg Retina Tomograph to detect early glaucomatous visual field loss. *J Glaucoma*. 1995;4(4):242–7.
137. Bathija R, Zangwill L, Berry CC, Sample PA, Weinreb RN. Detection of early

- glaucomatous structural damage with confocal scanning laser tomography. *J Glaucoma*. 1998;7(2):121–7.
138. Vidotti VG, Costa VP, Silva FR, Resende GM, Cremasco F, Dias M, et al. Sensitivity and specificity of machine learning classifiers and spectral domain OCT for the diagnosis of glaucoma. *Eur J Ophthalmol*. 2013; 23(1): 61-69.
139. Hodapp E, Parrish RK, Anderson DR. *Clinical decisions in glaucoma*. Mosby Incorporated; 1993.
140. Leske MC, Heijl A, Hyman L, Bengtsson B, Group EMGT. Early Manifest Glaucoma Trial: design and baseline data. *Ophthalmology*. 1999;106(11):2144–53.
141. Gordon MO, Kass MA. The Ocular Hypertension Treatment Study: design and baseline description of the participants. *Arch Ophthalmol*. 1999;117(5):573–83.
142. Mansberger SL, Demirel S. Early detection of glaucomatous visual field loss: why, what, where, and how. *Ophthalmol Clin North Am*. 2005;18(3):365-373.
143. Rao HL, Yadav RK, Begum VU, Addepalli UK, Choudhari NS, Senthil S, et al. Role of visual field reliability indices in ruling out glaucoma. *JAMA Ophthalmol*. 2015;133(1):40–4.
144. Mauschitz MM, Bonnemaier PWM, Diers K, Rauscher FG, Elze T, Engel C, et al. Systemic and ocular determinants of peripapillary retinal nerve fiber layer thickness measurements in the European eye epidemiology (E3) population. *Ophthalmology*. 2018;125(10):1526–36.
145. Kanai K, Abe T, Murayama K, Yoneya S. Retinal thickness and changes with age. *Nihon Ganka Gakkai Zasshi*. 2002 Mar;106(3):162–5.
146. Nieves-Moreno M, Martínez-de-la-Casa JM, Morales-Fernández L, Sánchez-Jean R, Sáenz-Francés F, García-Feijoó J. Impacts of age and sex on retinal layer thicknesses measured by spectral domain optical coherence tomography with Spectralis. *PLoS One*. 2018;13(3):e0194169.

147. Ooto S, Hangai M, Tomidokoro A, Saito H, Araie M, Otani T, et al. Effects of age, sex, and axial length on the three-dimensional profile of normal macular layer structures. *Invest Ophthalmol Vis Sci*. 2011 Nov;52(12):8769–79.
148. Bowd C, Zangwill LM, Berry CC, Blumenthal EZ, Vasile C, Sanchez-Galeana C, et al. Detecting early glaucoma by assessment of retinal nerve fiber layer thickness and visual function. *Invest Ophthalmol Vis Sci*. 2001 Aug;42(9):1993–2003.
149. Kanamori A, Nakamura M, Escano MFT, Seya R, Maeda H, Negi A. Evaluation of the glaucomatous damage on retinal nerve fiber layer thickness measured by optical coherence tomography. *Am J Ophthalmol*. 2003 Apr;135(4):513–20.
150. Chen TC, Hogue A, Junk AK, Nouri-Mahdavi K, Radhakrishnan S, Takusagawa HL, et al. Spectral-domain OCT: helping the clinician diagnose glaucoma: a report by the American Academy of Ophthalmology. *Ophthalmology*. 2018;125(11):1817–27.
151. Quigley HA, Varma R, Tielsch JM, Katz J, Sommer A, Gilbert DL. The relationship between optic disc area and open-angle glaucoma: the Baltimore Eye Survey. *J Glaucoma*. 1999 Dec;8(6):347–52.
152. Sekhar GC, Nagarajan R, Naduvilath TJ, Dandona L, Rao KM, Rao VD. Optic disc size in ocular hypertension. *Indian J Ophthalmol*. 1999 Dec;47(4):229–31.
153. Leung CKS, Choi N, Weinreb RN, Liu S, Ye C, Liu L, et al. Retinal nerve fiber layer imaging with spectral-domain optical coherence tomography: pattern of RNFL defects in glaucoma. *Ophthalmology*. 2010;117(12):2337–44.
154. Lisboa R, Leite MT, Zangwill LM, Tafreshi A, Weinreb RN, Medeiros FA. Diagnosing preperimetric glaucoma with spectral domain optical coherence tomography. *Ophthalmology*. 2012;119(11):2261–9.
155. Duan L, Xu L, Liu Y, Lee J. Cluster-based outlier detection. *Ann Oper Res*. 2009;168(1):151–68.

156. Chen T, Wang X, Chu Y, Wang Y, Jiang M, Wei D-Q, et al. T4SE-XGB: Interpretable Sequence-Based Prediction of Type IV Secreted Effectors Using eXtreme Gradient Boosting Algorithm. *Front Microbiol.* 2020;11:580382.
157. Ogunleye AA, Qing-Guo W. XGBoost Model for Chronic Kidney Disease Diagnosis. *IEEE/ACM Trans Comput Biol Bioinforma.* 2019 Apr;
158. Bowd C, Belghith A, Proudfoot JA, Zangwill LM, Christopher M, Goldbaum MH, et al. Gradient-Boosting Classifiers Combining Vessel Density and Tissue Thickness Measurements for Classifying Early to Moderate Glaucoma. *Am J Ophthalmol.* 2020 Sep;217:131–9.
159. Kuo DE, Wei MM, Armbrust KR, Knickelbein JE, Yeung IYL, Nussenblatt RB, et al. Gradient Boosted Decision Tree Classification of Endophthalmitis Versus Uveitis and Lymphoma from Aqueous and Vitreous IL-6 and IL-10 Levels. *J Ocul Pharmacol Ther Off J Assoc Ocul Pharmacol Ther.* 2017 May;33(4):319–24.
160. Huang M-L, Chen H-Y, Lin J-C. Rule extraction for glaucoma detection with summary data from StratusOCT. *Invest Ophthalmol Vis Sci.* 2007;48(1):244–50.
161. Baskaran M, Ong E-L, Li J-L, Cheung CY, Chen D, Perera SA, et al. Classification algorithms enhance the discrimination of glaucoma from normal eyes using high-definition optical coherence tomography. *Invest Ophthalmol Vis Sci.* 2012;53(4):2314–20.
162. Badalà F, Nouri-Mahdavi K, Raoof DA, Leeprechanon N, Law SK, Caprioli J. Optic disk and nerve fiber layer imaging to detect glaucoma. *Am J Ophthalmol.* 2007;144(5):724–32.

ANEXOS



D. Tomás Sebastián Viana, Secretario del Comité Ético de Investigación del Hospital Universitario de Fuenlabrada.

CERTIFICA

Que este Comité ha evaluado el Proyecto de Investigación titulado:

“Diagnóstico del glaucoma en estadio precoz mediante el estudio de la capa de fibra retinianas peripapilares: una aplicación de minería de datos en medicina.”

y considera que:

- Se cumplen los requisitos necesarios de idoneidad del protocolo en relación con los objetivos del estudio y están justificados los riesgos y molestias previsibles para el sujeto.
- La capacidad del investigador y los medios disponibles son apropiados para llevar a cabo el estudio.
- Son adecuados tanto el procedimiento para obtener el consentimiento informado como la compensación prevista para los sujetos por daños que pudieran derivarse de su participación en el ensayo.
- El alcance de las compensaciones económicas previstas no interfiere con el respeto a los postulados éticos.

Y que este Comité acepta dicho Estudio, y como investigador principal:

Dr. Carlos Salvador Fernández Escámez
del Servicio de **Oftalmología**
del Hospital Universitario de Fuenlabrada

Fuenlabrada, 28 de septiembre de 2018



Firmado: D. Tomás Sebastián Viana
Secretario Técnico CEI

HOJA DE INFORMACION AL PACIENTE

TITULO DEL ESTUDIO: “DIAGNOSTICO DEL GLAUCOMA EN ESTADIO PRECOZ MEDIANTE EL ESTUDIO DE LA CAPA DE FIBRA RETINIANAS PERIPAPILARES: UNA APLICACIÓN DE MINERIA DE DATOS EN MEDICINA”.

CODIGO DE ESTUDIO:

VERSION: 01

INVESTIGADOR PRINCIPAL: Dr. Carlos Salvador Fernández Escámez

CENTRO: Hospital Universitario de Fuenlabrada.

NOMBRE DEL PACIENTE:

El presente documento está destinado a informarle sobre la existencia un estudio en el que se le invita a participar. El estudio ha sido aprobado por el Comité de Etica e Investigación del Hospital Universitario de Fuenlabrada, de acuerdo a la legislación vigente, y se lleva a cabo respetando a las normas de buena práctica clínica

¿POR QUE SE REALIZA ESTE ESTUDIO?

El glaucoma es una enfermedad con potencial de causar ceguera irreversible. La cantidad de personas con glaucoma se estima en un 4% de la población a los 70 años de edad. Al disponer de tratamientos efectivos para el glaucoma, su detección precoz permite evitar la evolución de la enfermedad y la posible ceguera de la persona enferma. Por ello, es importante perfeccionar las técnicas que puedan mejorar la cantidad de personas con glaucoma que pueden ser detectadas y tratadas en las fases iniciales de la enfermedad.

FINALIDAD DEL ESTUDIO. ¿Qué exploraciones se realizaran en su ojo?

El estudio está destinado a mejorar la detección de personas afectas de glaucoma en las fases iniciales de la enfermedad. Para ello se necesita disponer de un grupo de personas con la enfermedad, y un grupo de personas sanas como Vd., con exploración ocular normal, con las que se pueda comparar a las personas enfermas.

Los datos que serán utilizados en el estudio son los obtenidos mediante la exploración ocular con la técnica de Tomografía de Coherencia Óptica y con el examen del campo visual mediante la realización de un Campo Visual. Ambas son técnicas totalmente indoloras, que se realizan en la consulta de Oftalmología, y que no necesitan ningún tipo de preparación previa. No son técnicas invasivas, por lo que no habrá que canalizar ninguna vena ni utilizar ningún tipo de dispositivo. Además, estas técnicas de exploración ocular en ningún caso pueden provocar pérdida visual ni ningún otro evento adverso ocular.

Vd., como sujeto sin enfermedad ocular, será examinado en uno de sus ojos mediante estas técnicas, incorporando los datos obtenidos al estudio.

¿Qué beneficios puedo obtener por participar en este estudio?

La participación en el estudio es totalmente voluntaria y altruista, no habiendo ningún tipo de compensación económica. El equipo investigador tampoco recibe ningún tipo de beneficio económico por llevar a cabo este estudio.

No obstante, con este estudio se pretende obtener nueva información que a día de hoy no tenemos, con el fin de poder detectar a tiempo al mayor número posible de personas con glaucoma y evitar así la posible ceguera provocada por la enfermedad. El beneficio para la sociedad es indudable, ya que podrá colaborar a que en el futuro haya menos personas con ceguera como consecuencia del glaucoma.

¿Qué riesgos y/o molestias puedo tener por participar en el estudio?

Las técnicas de exploración que realizamos en su ojo no tienen ningún tipo de contraindicación ni efecto secundario. No producen ningún tipo de perjuicio ocular ni necesitan canular ninguna vía venosa ni utilizar ningún dispositivo externo.

Por ello, la única molestia es el tiempo invertido por Vd. en la realización de las pruebas, que será aproximadamente de 15 minutos.

Anexos

¿Qué datos se van a recoger?

Los datos que serán utilizados en el estudio son los obtenidos mediante la exploración ocular con la Tomografía de Coherencia Óptica y la realización de un Campo Visual. También se incluirán datos de edad y sexo.

Todos los datos serán tratados de manera anónima, no pudiendo ser relacionados con Vd. una vez incorporados en la base de datos.

CONFIDENCIALIDAD DE LOS DATOS

Todos sus datos se tratarán confidencialmente por el investigador principal, obligado con el deber de secreto profesional. No se utilizarán su nombre ni apellidos para guardar la información registrada, utilizando un sistema de indexación que asegurará su anonimato.

De acuerdo con la Ley Orgánica de Protección de Datos 15/1999, debe saber que tiene derecho a acceder a los datos que de Vd. se guarden, a rectificarlos, a cancelarlos y a oponerse a su uso en cualquier momento sin tener que dar ninguna explicación.

La participación en el estudio es totalmente voluntaria. Cumpliendo con la Ley 41/2002 reguladora de la autonomía del paciente, Vd. podrá retirarse del estudio en cualquier momento si lo desea, sin tener que dar explicaciones y sin que por ello se produzca perjuicio alguno en la relación con su médico ni en los cuidados que se le deban administrar.

¿Con quien puedo contactar en caso de duda?

El investigador principal del presente estudio está a su disposición para informar y contestar a sus dudas y preguntas.

Nombre: Carlos Salvador Fernández Escámez

Tel Contacto: 91-6006129 /91-6006131

**CONSENTIMIENTO INFORMADO PARA INCLUSION EN ESTUDIO
“DIAGNOSTICO DEL GLAUCOMA EN ESTADIO PRECOZ MEDIANTE EL
ESTUDIO DE LA CAPA DE FIBRA RETINIANAS PERIPAPILARES: UNA
APLICACIÓN DE MINERIA DE DATOS EN MEDICINA”.**

CODIGO DE ESTUDIO:

VERSION: 01

NOMBRE:

Yo, D.
he sido informado por el Dr. acerca de la existencia
de un estudio que se desarrolla en el Hospital Universitario de Fuenlabrada destinado a
comparar personas sanas con pacientes afectos de la enfermedad de glaucoma,
solicitando mi autorización para incluir los datos obtenidos de la exploración de mis
ojos en dicho estudio, así como para realizar las técnicas de Campo Visual y
Tomografía de Coherencia Optica como parte de la exploración ocular, siendo
informado de que ambas técnicas son NO INVASIVAS y que en ningún caso
provocarían pérdida visual ni ninguna otra alteración ocular. Además he leído la hoja de
información que me ha sido entregada, he podido formular todas las preguntas que
deseara sobre el estudio y he recibido suficiente información sobre el estudio, por lo
que:

AUTORIZO

NO AUTORIZO

A que mis datos sean incorporados a la base datos para participar en el estudio y doy mi
consentimiento para el acceso y utilización de mis datos en las condiciones detalladas
en la hoja de información.

Comprendo que mi participación es voluntaria, pudiendo retirarme del estudio en
cualquier momento, sin tener que dar explicaciones, y sin que ello repercuta en mis
cuidados médicos.

Anexos

En Fuenlabrada, de del 2018

Fdo:

