



MÁSTER DE FORMACIÓN PERMANENTE EN INTELIGENCIA ARTIFICIAL

Predicción de fibrosis hepática en pacientes con insuficiencia cardíaca mediante inteligencia artificial utilizando variables clínicas y el FibroScan como referencia diagnóstica.

Presentado por:

LUIS STUHARDO TERÁN BETANCO

Dirigido por:

ALEJANDRO PERDIGUERO O'LEARY

CURSO ACADÉMICO 2024-2025

Resumen

El presente Trabajo de Fin de Máster aborda la necesidad crítica de desarrollar métodos no invasivos para predecir la fibrosis hepática en pacientes con insuficiencia cardíaca, especialmente en escenarios donde el acceso al FibroScan es limitado. El estudio se propuso desarrollar un sistema basado en Machine Learning capaz de predecir la fibrosis hepática utilizando variables clínicas relevantes, complementadas con las mediciones del FibroScan como referencia diagnóstica. El proceso de desarrollo de estos modelos subraya la importancia fundamental de la experimentación y la validación iterativa. Se exploraron diversas vías, como la clasificación multiclase inicial, que, al revelar limitaciones por el desbalance de datos y la complejidad de grados intermedios, justificó la transición a un enfoque binario más aplicado. Este camino adaptativo permitió optimizar los modelos, priorizando el rendimiento en métricas específicas según el contexto clínico, como una alta sensibilidad para la detección de la enfermedad. La selección de variables de importancia se erigió como un paso crucial, no solo para mejorar la interpretabilidad de los modelos más sencillos, como la regresión logística, sino también para validar la relevancia clínica de los predictores identificados. En última instancia, este trabajo culmina en el desarrollo de herramientas basadas en inteligencia artificial que tienen el potencial de integrarse y enriquecer las investigaciones clínicas actuales. Resalta así la trascendencia de contar con especialistas que, trabajando activamente en estos modelos y explorando diversas estrategias, puedan traducir el potencial del Machine Learning en soluciones prácticas y significativas para el avance de la medicina.

Palabras clave: Fibrosis hepática, insuficiencia cardíaca, aprendizaje automático, FibroScan, clasificación binaria.

Abstract

This Master's Thesis addresses the critical need to develop non-invasive methods for predicting liver fibrosis in patients with heart failure, especially in settings where access to FibroScan is limited. The study aimed to develop a Machine Learning–based system capable of predicting liver fibrosis using relevant clinical variables, complemented by FibroScan measurements as the diagnostic reference. The development process of these models highlights the fundamental importance of experimentation and iterative validation. Various approaches were explored, such as the initial multiclass classification, which, upon revealing limitations due to data imbalance and the complexity of intermediate fibrosis stages, justified the shift to a more applicable binary approach. This adaptive path enabled the optimization of models, prioritizing performance in specific metrics according to the clinical context, such as high sensitivity for disease detection. The selection of important variables emerged as a crucial step, not only to improve the interpretability of simpler models, such as logistic regression, but also to validate the clinical relevance of the identified predictors. Ultimately, this work culminates in the development of artificial intelligence–based tools with the potential to integrate into and enhance ongoing clinical research. It thus underscores the significance of having specialists who, actively working on these models and exploring diverse strategies, can translate the potential of Machine Learning into practical and meaningful solutions for the advancement of medicine.

Keywords: Liver fibrosis, heart failure, machine learning, FibroScan, binary classification.

ÍNDICE

1. Introducción	9
2. Justificación	10
3. Objetivos	11
3.1. Objetivo General.....	11
3.2. Objetivos Específicos	11
4. Marco Teórico.....	12
4.1. Relación entre Insuficiencia Cardíaca y Enfermedad Hepática.....	12
4.1.1. Congestión hepática en insuficiencia cardíaca	13
4.1.2. Biomarcadores cardiovasculares y hepáticos en la evaluación de la congestión sistémica en insuficiencia cardíaca	15
4.2. FibroScan y Elastografía de Transición.....	16
4.2.1. Fundamentos técnicos del FibroScan	17
4.2.2. Tipos de sondas y su aplicación clínica	19
4.2.3. Interpretación de resultados: escala METAVIR	23
4.3. Inteligencia Artificial en Medicina	25
4.3.1. Aplicaciones específicas en cardiología	25
4.3.2. La IA en insuficiencia cardiaca.....	26
4.4. Modelos de predicción y clasificación aplicados al diagnóstico clínico.....	27
4.4.1 Métricas de evaluación de modelos predictivos	29

4.4.2. Modelos Predictivos Basados en Árboles de Decisión.....	31
5. Metodología	33
5.1. Población y Muestra	33
5.2. Variables a Analizar.....	34
5.3. Implementación Analítica y Técnicas de Inteligencia Artificial.....	39
5.3.1. Exploración de Datos (EDA)	39
5.3.2. Modelado	41
6. Resultados	52
6.1. Presentación de los Hallazgos del Análisis Exploratorio de Datos (EDA).....	52
6.2. Resultados de los Modelos de Machine Learning	57
6.2.1. Resultados del Modelo de Clasificación Multiclas (Inicial)	57
6.2.2 Resultados del Modelo de Clasificación Binaria (Principal - Gradient Boosting)	60
6.2.3. Resultados del Modelo de Refinamiento (Regresión Logística)	63
7. Discusión	66
8. Conclusiones	71
9. Bibliografía	74
10. Código Fuente y Repositorio GitHub.....	79

ÍNDICE DE FIGURAS

Figura 1 <i>Circulación hepática</i>	13
Figura 2 <i>Estructura del corazón humano</i>	14
Figura 3 <i>FibroScan® Mini+ 430</i>	16
Figura 4 <i>Prueba FibroScan</i>	19
Figura 5 <i>¿Qué sonda FibroScan® se adapta a su paciente?</i>	20
Figura 6 <i>Ejemplo de clasificación binaria</i>	28
Figura 7 <i>Clasificación desequilibrada</i>	29
Figura 8 <i>Comparación modelos basados en arboles</i>	32
Figura 9 <i>Distribución de pacientes por grado de fibrosis</i>	53
Figura 10 <i>Presencia/ausencia de fibrosis</i>	54
Figura 11 <i>Grado de Fibrosis</i>	55
Figura 12 <i>Presencia de Fibrosis</i>	56
Figura 13 <i>Matriz de Confusión - Grado de Fibrosis</i>	58
Figura 14 <i>Curvas Roc y AUC Clasificación Multiclasificación</i>	59
Figura 15 <i>Importancia de variables modelo multiclasificación</i>	60
Figura 16 <i>Matriz de Confusión - Presencia de Fibrosis</i>	61
Figura 17 <i>Curvas Roc y AUC Modelo Binario</i>	62
Figura 18 <i>Importancia de variables modelo binario</i>	63
Figura 19 <i>Matriz de Confusión - Presencia de Fibrosis Regresión Logística</i>	64
Figura 20 <i>Curvas Roc y AUC Regresión Logística</i>	65

ÍNDICE DE TABLAS

Tabla 1 <i>Criterios de selección</i>	22
Tabla 2 <i>Valores de rigidez hepática (kPa) según etiología y estadio de fibrosis</i>	24
Tabla 3 <i>Métricas de evaluación de modelos de clasificación</i>	30
Tabla 4 <i>Variables Dependientes</i>	34
Tabla 5 <i>Variables Independientes (Parámetros Demográficos y Fisiológicos)</i>	35
Tabla 6 <i>Variables Clínicas (Comorbilidades)</i>	36
Tabla 7 <i>Tratamiento Farmacológico Previo</i>	37
Tabla 8 <i>Parámetros Analíticos / Laboratorio</i>	38
Tabla 9 <i>Información general:</i>	52
Tabla 10 <i>Métricas de Rendimiento Modelo Multiclasificación</i>	57
Tabla 11 <i>Métricas de Rendimiento Modelo Binario</i>	61
Tabla 12 <i>Métricas de Rendimiento Modelo Regresión Logística</i>	64

1. Introducción

La insuficiencia cardíaca y la enfermedad hepática presentan una estrecha relación, dado que el corazón y el hígado son órganos que colaboran activamente en el mantenimiento de la circulación sanguínea. Cuando el corazón pierde la capacidad de bombear sangre de manera eficiente, puede provocar congestión: la sangre se acumula en las venas y órganos, generando congestión hepática. Esta situación afecta directamente al funcionamiento del hígado, provocando su inflamación, y en casos crónicos, conduciendo a procesos de fibrosis hepática. De manera bidireccional, un hígado enfermo ya sea por la presencia de fibrosis hepática o por enfermedad hepática metabólica (que abarca desde la esteatosis simple hasta la esteatohepatitis no alcohólica, NASH) puede impactar negativamente en la función cardíaca. La alteración de la estructura hepática aumenta la resistencia al flujo sanguíneo portal, afectando la reserva de precarga en el corazón y contribuyendo, además, a la liberación de toxinas que dañan directamente el tejido cardíaco. Todo ello disminuye progresivamente la calidad de vida del paciente (Xanthopoulos et al., 2019).

Según de Lédinghen y Vergniol (2008), la elastografía de transición es una técnica no invasiva que, utilizada como herramienta diagnóstica, permite identificar fenotipos específicos asociados a enfermedad hepática metabólica. La elastografía de transición mediante FibroScan ofrece una alternativa rápida, segura y estandarizada para la estimación de rigidez hepática, útil para detectar y monitorizar la fibrosis.

2. Justificación

En la actualidad, si bien el FibroScan representa una herramienta diagnóstica no invasiva, rápida y estandarizada para la estimación de rigidez hepática y la detección de fibrosis, su acceso no siempre está garantizado en todos los centros de salud, especialmente en contextos con recursos limitados o en pacientes con comorbilidades que dificultan su aplicación. Esta situación plantea la necesidad de desarrollar enfoques complementarios que permitan inferir la presencia de fibrosis hepática a partir de datos clínicos disponibles de forma más generalizada.

En este contexto, el uso de métodos basados en Inteligencia Artificial se presenta como una herramienta valiosa para apoyar el diagnóstico médico, mediante la construcción de modelos predictivos que integren múltiples variables clínicas relevantes. Esto puede facilitar la identificación de patrones ocultos en los datos, permitiendo, por ejemplo, reconocer factores con mayor influencia en la fibrosis hepática y contribuir así a una toma de decisiones más informada y eficiente en pacientes con insuficiencia cardíaca.

3. Objetivos

3.1. Objetivo General

Desarrollar modelos de inteligencia artificial capaces de predecir la presencia de fibrosis hepática en pacientes con insuficiencia cardíaca, utilizando variables clínicas y el FibroScan como referencia diagnóstica.

3.2. Objetivos Específicos

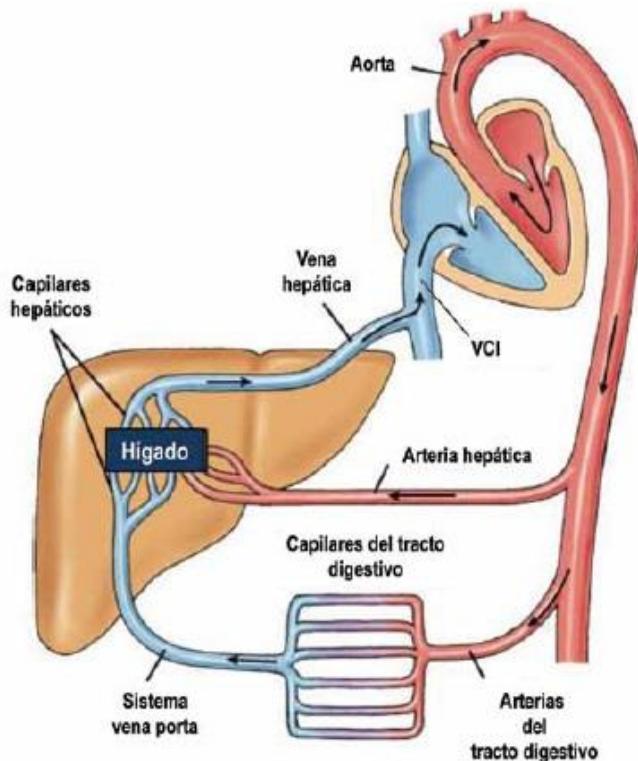
- Aplicar algoritmos de aprendizaje supervisado para construir modelos de predicción binaria (fibrosis: sí/no).
- Explorar el valor de la inteligencia artificial como herramienta complementaria en escenarios donde el acceso al FibroScan sea limitado.
- Identificar las variables clínicas con mayor influencia en la predicción de fibrosis hepática mediante análisis de importancia de características.

4. Marco Teórico

4.1. Relación entre Insuficiencia Cardíaca y Enfermedad Hepática

Entendiendo que el corazón y el hígado mantienen una relación funcional bidireccional, es decir, existen interacciones cardiohepáticas constantes (Figura 1), la comprensión de esta interdependencia adquiere especial relevancia en el abordaje clínico de pacientes con afectación en uno o ambos órganos. Las causas cardíacas de disfunción hepática incluyen pericarditis constrictiva, hipertensión arterial pulmonar (HAP) grave, estenosis mitral, insuficiencia tricuspídea (IT), corpulmonale, miocardiopatía isquémica y consecuencias posoperatorias de la cirugía de Fontan para la atresia pulmonar y el síndrome del corazón izquierdo hipoplásico (Fouad & Yehia, 2014).

De forma transversal, un hígado afectado por distintas condiciones, como la esteatosis hepática, la enfermedad hepática no alcohólica (NAFLD), entre otras, puede influir negativamente en la función cardíaca, contribuyendo al deterioro progresivo del rendimiento del corazón. "En una etapa avanzada, la enfermedad hepática produce vasodilatación esplánica y sistémica, alteración de la barrera inmunológica, translocación de bacterias intestinales y disfunción autonómica, evolucionando hacia un síndrome multiorgánico que afecta al corazón, pulmones, riñones y cerebro" (Brankovic et al., 2023).

Figura 1*Circulación hepática*

Nota. Interacciones cardiohepáticas, Tomado de Baialardo, A. G. (2015). Síndrome cardiohepático: ¿Qué debemos saber como cardiólogos? Insuficiencia cardíaca, 10(2), 66–77.

4.1.1. Congestión hepática en insuficiencia cardíaca

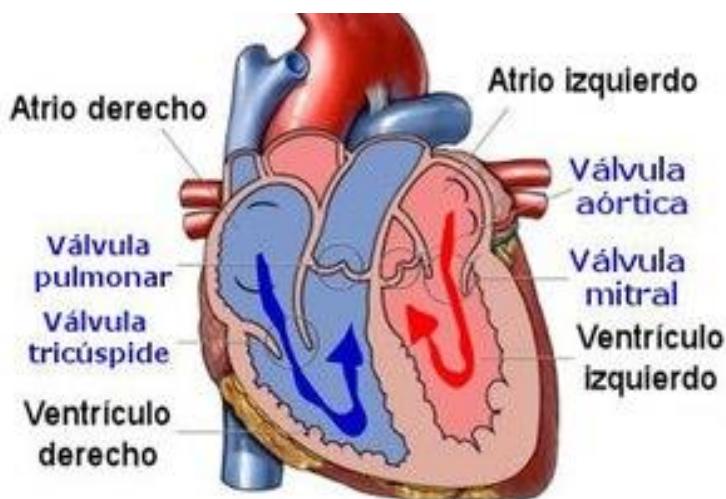
“Un efecto exclusivo de la insuficiencia cardíaca, a diferencia de otras causas de colapso circulatorio, es la congestión pasiva en el hígado” (Giallourakis et al., 2002). Gran parte del flujo sanguíneo que pasa por el hígado proviene del sistema cardiovascular, la insuficiencia cardíaca congestiva conduce inevitablemente a una acumulación de sangre en el hígado. Esta congestión puede estar provocada por factores como la disminución de la saturación arterial de oxígeno, presión venosa central elevada, entre otras.

“La congestión pasiva crónica conduce a hipertensión sinusoidal, fibrosis centrolobulillar y, en última instancia, cirrosis (“cirrosis cardíaca”) y carcinoma hepatocelular después de varias décadas de lesión continua” (Fortea et al., 2020, p. 3). Por ello, el diagnóstico y tratamiento oportuno de la hepatopatía congestiva resulta fundamental para prevenir la progresión hacia complicaciones graves, como la fibrosis hepática y sus consecuencias asociadas.

Es fundamental considerar las cardiopatías congénitas, especialmente aquellas que afectan el lado derecho del corazón, debido a su relación directa con el hígado. Según Hilscher y Sánchez (2016), “Cualquier causa de insuficiencia cardíaca ventricular derecha puede precipitar congestión hepática, incluyendo pericarditis constrictiva, insuficiencia tricuspídea, estenosis mitral, miocardiopatía y cor pulmonale”.

Figura 2

Estructura del corazón humano



Nota. Tomado de Ingrassia, V. (2018, mayo 25). Las nuevas válvulas cardíacas que se implantan sin operar a corazón abierto. Infobae. <https://www.infobae.com/salud/2018/05/25/las-nuevas-valvulas-cardiacas-que-se-implantan-sin-operar-a-corazon-abierto/>

4.1.2. Biomarcadores cardiovasculares y hepáticos en la evaluación de la congestión sistémica en insuficiencia cardíaca

En este contexto, resulta útil comprender la relevancia de ciertas variables fenotípicas, como los biomarcadores de congestión NT-proBNP (fragmento N-terminal del péptido natriurético tipo B) y CA125 (antígeno carbohidrato 125). Si bien estos ya están consolidados en el manejo de la insuficiencia cardíaca, es importante explorar su posible vinculación, directa o indirecta, con enfermedades hepáticas. El análisis multivariante mostró que el CA125 se asoció positiva e independientemente con presencia de edema periférico, derrame pleural y valores elevados de VCI. El NT-proBNP se relacionó con el derrame pleural y el diámetro de VCI, pero no con el edema (Llàcer et al., 2021).

Para evaluar la congestión sistémica y su posible impacto sobre el hígado, resulta fundamental considerar biomarcadores como el NT-proBNP y el CA125, ampliamente utilizados por los cardiólogos en el seguimiento de pacientes con insuficiencia cardíaca. Además de estos, otros parámetros como la gamma-glutamil transferasa (GGT) y la alanina aminotransferasa (GPT) pueden aportar información adicional. Aunque tradicionalmente se asocian con daño hepático, la elevación de estas enzimas hepáticas puede reflejar congestión venosa hepática secundaria a disfunción cardíaca. Según Lab Tests Online (2021), tanto la GGT como la ALP aumentan en las enfermedades hepáticas, pero la ALP se puede elevar también cuando existen lesiones óseas. Por lo tanto, si la ALP está aumentada, la determinación de GGT es útil para saber si el aumento de ALP es atribuible a una enfermedad hepática o a una enfermedad ósea. Del mismo modo, la presencia de edemas periféricos puede indicar un compromiso hemodinámico avanzado, lo que refuerza la importancia de un enfoque integrado que considere tanto variables cardiovasculares como hepáticas en la valoración de estos pacientes.

Estudiar la interrelación entre el corazón y el hígado, la relevancia clínica de los biomarcadores de congestión y la necesidad de una clasificación fenotípica más precisa en patologías como la IC, refuerzan la importancia de una visión integral para el abordaje diagnóstico y terapéutico de los síndromes cardiohepáticos.

4.2. FibroScan y Elastografía de Transición

El fibroscan es un instrumento que emplea técnicas de medición sofisticadas como la elastografía de transición, y más importante de manera no invasiva permite obtener una medida aproximada del grado de rigidez y esteatosis en el hígado.

Figura 3

FibroScan® Mini+ 430



Nota. FibroScan® Mini+ 430 es ligero y fácil de manejar (5 kg). Es un dispositivo alimentado por batería con pantalla táctil de 12,1 pulgadas. Adaptado de Echosens (2018), <https://www.choesens.com/es/products/fibroscan-mini-2/>

Antiguamente para conocer el estado de rigidez o fibrosis del hígado en pacientes era realizando una biopsia hepática, que consiste en una prueba ambulatoria que extrae un pequeño fragmento de tejido hepático que es analizado por el patólogo. Este método es completamente invasivo no está exento de riesgos, además de resultar incómodo para el paciente en cuestión. Actualmente, muchas de estas biopsias pueden evitarse utilizando un equipo de elastografía de transición(Fibroscan), mediante la cual se obtiene una estimación indirecta del grado de fibrosis hepática de una forma no invasiva y con una buena correlación con la biopsia hepática (*elastografía de transición FibroScan*, n.d.).

La fibrosis hepática es parte de las alteraciones estructurales y funcionales en la mayoría de las enfermedades hepáticas crónicas. Es uno de los principales factores de pronóstico ya que la cantidad de fibrosis se correlaciona con el riesgo de desarrollar cirrosis y complicaciones relacionadas con el hígado en enfermedades hepáticas crónicas virales y no virales (“EASL-ALEH Clinical Practice Guidelines,” 2015).

4.2.1. Fundamentos técnicos del FibroScan

La elastografía de transición mediante FibroScan utiliza una sonda transductora de ultrasonido que genera una onda transversal elástica a través de vibraciones de baja frecuencia (50 Hz) y amplitud leve, las cuales atraviesan el tejido hepático. Esta misma sonda emplea tecnología de pulso-eco para rastrear la propagación de la onda y medir su velocidad (en m/s), permitiendo así estimar la rigidez del hígado en un volumen de tejido aproximadamente 100 veces mayor que el obtenido por biopsia con aguja. La prueba es rápida y segura, con una duración aproximada de entre 5 y 10 minutos. Se recomienda que el paciente esté en ayunas entre 2 a 3 horas antes del procedimiento, debido a que la ingesta de alimentos puede aumentar temporalmente la rigidez hepática por el incremento del flujo sanguíneo posprandial (Patel &

Wilder, 2014). La sonda se aplica sobre la silueta hepática, entre las costillas, en la línea axilar media derecha y perpendicular al plano cutáneo. En la pantalla del FibroScan se genera una imagen bidimensional de la onda elástica, cuya morfología puede facilitar la obtención de registros válidos.

Además de los aspectos técnicos del procedimiento, es importante considerar tanto la posición del paciente como la del operador para garantizar la validez y reproducibilidad de la medición. El paciente debe encontrarse completamente recostado en decúbito supino sobre la camilla, con el brazo derecho extendido por detrás de la cabeza, lo que facilita el acceso intercostal al lóbulo hepático derecho. Se recomienda que adopte una posición ligeramente arqueada hacia la izquierda, similar a una forma de "banana", cruzando la pierna derecha sobre la izquierda, lo que contribuye a una mejor exposición del área a explorar.

Asimismo, la posición del operador es determinante para la precisión del estudio. Idealmente, debe ubicarse al lado derecho del paciente, sentado en una silla con ruedas que le permita libertad de movimiento. El operador debe tener acceso visual directo tanto a la zona de aplicación como a la pantalla del dispositivo, utilizando ambas manos de forma coordinada: la mano derecha sostiene la sonda y realiza los ajustes necesarios, mientras que la mano izquierda ayuda a mantener el ángulo y la presión adecuada para garantizar una correcta transmisión de la onda elástica.

Figura 4*Prueba FibroScan*

Nota. Evaluación avanzada y no invasiva de la salud hepática. Tomado de Alpharetta Internal Medicine (s.f.), <https://alpharettainternalmedicine.com/live-health-fibroscan/>

4.2.2. Tipos de sondas y su aplicación clínica

Para realizar la medición con FibroScan, se utilizan diferentes tipos de sondas. Las más comunes son la sonda M y la sonda XL, aunque en algunos casos también se usa una sonda especial para niños. La elección de la sonda depende del tipo de paciente. Por ejemplo, la sonda XL se utiliza en personas con sobrepeso, ya que permite obtener mejores resultados en tejidos más profundos. En cambio, en personas delgadas, se recomienda la sonda M, que es más precisa en esos casos. En conclusión, la elastografía de transición con la sonda XL obtiene una mayor proporción de exámenes fiables en pacientes obesos, pero no en pacientes no obesos (Herrero et al., 2014).

Figura 5

¿Qué sonda FibroScan® se adapta a su paciente?



Nota. FibroScan® ofrece una gama de tres sondas (S+, M+ y XL+) diseñadas para adaptarse a las características morfológicas del paciente. Tomado de Echosens (s.f.),
<https://www.echosens.com/es/products/sondas/>

En pacientes adultos, las sondas más comúnmente utilizadas son la sonda M y la sonda XL, y su elección depende principalmente del índice de masa corporal (IMC) y de la distribución del tejido adiposo, especialmente en la región abdominal. En pacientes con obesidad o acumulación significativa de grasa subcutánea, la sonda XL es la opción preferida, ya que permite alcanzar estructuras hepáticas más profundas y proporciona mediciones más confiables en estos casos. En cambio, en pacientes con constitución delgada o normopeso, la sonda M es más apropiada, ya que ofrece alta precisión y facilidad de uso.

Además, el tamaño y potencia del pulso mecánico generado por la sonda pueden influir en la tolerancia del procedimiento. La sonda M, al ser de menor tamaño, genera una estimulación más suave, lo que puede ser beneficioso en pacientes con hipersensibilidad, dolor abdominal previo o condiciones clínicas delicadas, ya que reduce la incomodidad durante la exploración. Por otro lado, aunque la sonda XL es indispensable en contextos de obesidad, su mayor energía y necesidad de repetición pueden resultar más molestas en ciertos casos, lo que refuerza la importancia de una selección individualizada según las características físicas y clínicas del paciente.

Tabla 1*Criterios de selección*

Característica	Sonda S+ VCTE™	Sonda M+ con Guided VCTE™	Sonda XL+ con Guided VCTE™
Uso principal	Para población pediátrica	Para pacientes con morfotipo estándar	Para pacientes con obesidad y obesidad extrema
Diseño	Diseñada para colocarse en un espacio intercostal estrecho	–	Diseñada para mejorar la penetración de la señal a través de tejidos más profundos
Profundidad de la medición	De 15 a 50 mm	De 25 a 70 mm	De 35 a 85 mm
Criterios de selección	Edad < 18 años y perímetro torácico \leq 75 cm	14 años \leq Edad < 18 años y perímetro torácico > 75 cm o Edad \geq 18 años y sonda XL+ no recomendada	Edad \geq 18 años y sonda M+ no recomendada

Nota. Características técnicas y criterios de selección para las sondas S+, M+ y XL+ de FibroScan®. Tomado de Echosens (s.f.), <https://www.echosens.com/es/products/sondas/>

4.2.3. Interpretación de resultados: escala METAVIR

La escala METAVIR es una de las más utilizadas para interpretar los resultados del FibroScan. Esta clasifica la fibrosis hepática en cinco etapas: F0 (sin fibrosis), F1 (fibrosis leve), F2 (fibrosis moderada), F3 (fibrosis avanzada) y F4 (cirrosis). El valor final se obtiene a partir de diez mediciones válidas, de las cuales se calcula la mediana, que es la que se utiliza para determinar el grado de rigidez hepática. Los resultados del FibroScan se expresan en kilopascales (kPa), con un rango que va desde 2,5 hasta 75 kPa. En personas sanas, los valores suelen rondar los 5,5 kPa. La correlación entre los valores del FibroScan y los grados de fibrosis según la escala METAVIR es la siguiente: menos de 7,6 kPa = F0–F1; entre 7,7 y 9,4 kPa = F2; entre 9,5 y 14 kPa = F3; y más de 14 kPa = F4 (Bartres & Lens, 2013).

Tabla 2*Valores de rigidez hepática (kPa) según etiología y estadio de fibrosis*

Enfermedad	F0 a F1	F2	F3	F4
Hepatitis B	2 a 7 kPa	8 a 9 kPa	8 a 11 kPa	18 kPa o más
Hepatitis C	2 a 7 kPa	8 a 9 kPa	9 a 14 kPa	14 kPa o más
Coinfección de HIV/HCV	2 a 7 kPa	7 a 11 kPa	11 a 14 kPa	14 kPa o más
Enfermedad colestática	2 a 7 kPa	7 a 9 kPa	9 a 17 kPa	17 kPa o más
Esteatohepatitis no alcohólica (NAFLD/NASH)	2 a 7 kPa	7.5 a 10 kPa	10 a 14 kPa	14 kPa o más
Enfermedad relacionada con el alcohol	2 a 7 kPa	7 a 11 kPa	11 a 19 kPa	19 kPa o más

Nota. Adaptado de Información sobre sus resultados de FibroScan, por Bexaray, 2018.

Recuperado el 24 de junio de 2025 de <https://www.bexaray.com/single-post/2018/11/29/informaci%C3%B3n-sobre-sus-resulstados-de-fibroscan>

4.3. Inteligencia Artificial en Medicina

La implementación de sistemas basados en algoritmos de inteligencia artificial abarca desde el diagnóstico hasta el tratamiento de enfermedades. Un ejemplo claro es el análisis de imágenes médicas, donde la IA puede detectar patrones o anomalías con mayor rapidez y precisión que un humano, facilitando así diagnósticos tempranos. Además, se aplica en el diseño de tratamientos personalizados, gracias al desarrollo de dispositivos inteligentes capaces de monitorear continuamente al paciente, detectar variaciones en su estado de salud e incluso anticipar posibles complicaciones. Asimismo, la IA puede utilizarse con fines predictivos, a través de algoritmos capaces de identificar el riesgo de desarrollar ciertas enfermedades antes de que se manifiesten clínicamente, lo que permite una intervención temprana y, en muchos casos, potencialmente salva vidas gracias a una prevención más eficaz.

Debido al reciente auge y desarrollo de la inteligencia artificial, se han logrado avances significativos en el campo de la medicina, introduciendo soluciones innovadoras en diversas áreas. Por ejemplo, se han desarrollado algoritmos de aprendizaje automático capaces de analizar datos, evaluar riesgos y, de esta manera, mejorar la eficiencia del trabajo clínico (Raraz-Vidal & Raraz-Vidal, 2022).

4.3.1. Aplicaciones específicas en cardiología

“En cardiología, el uso de la IA se ha estudiado en la predicción de hipertensión esencial, la detección de fibrilación auricular por relojes inteligentes, la clasificación de estenosis aórtica por análisis de señales cardiomecánicas en sensores inalámbricos portátiles, la clasificación de

arritmias mediante electrocardiograma de una sola derivación, etc" (Lanzagorta-Ortega et al., 2022).

La expansión de la inteligencia artificial en distintos ámbitos del conocimiento es cada vez más evidente, y el sector sanitario no es la excepción. En medicina, el interés por estas tecnologías ha crecido significativamente debido a su potencial para optimizar los procesos diagnósticos, terapéuticos y de seguimiento clínico, con el objetivo último de prolongar la vida y mejorar su calidad. Dentro de las especialidades médicas, la cardiología se posiciona como una de las más relevantes para la aplicación de la IA, no solo por la carga de enfermedad asociada a las patologías cardiovasculares, sino también por el papel central que desempeña el corazón en la homeostasis del organismo. El mal funcionamiento cardíaco compromete gravemente múltiples sistemas del cuerpo humano, afectando de forma directa la funcionalidad y el bienestar del paciente. Por ello, el desarrollo de soluciones basadas en IA para el abordaje de enfermedades cardíacas representa un campo de investigación prioritario cuando se busca maximizar la expectativa y calidad de vida.

4.3.2. La IA en insuficiencia cardiaca

El uso de modelos de *machine learning* en casos de insuficiencia cardíaca puede ser clave para desarrollar métodos que permitan una mejor comprensión de la enfermedad y de otras afecciones relacionadas, como la enfermedad hepática. Estas condiciones suelen estar interconectadas y presentan síntomas o alteraciones similares, lo que dificulta identificar con precisión las causas. Dado que los métodos convencionales pueden resultar limitados en estos casos, la creación de modelos que integren y analicen ambas enfermedades se vuelve una herramienta potencialmente útil y prometedora para mejorar tanto el diagnóstico como el tratamiento de los pacientes.

En el caso de la insuficiencia cardíaca, la inteligencia artificial puede ser de gran ayuda en el cuidado de pacientes que son propensos a sufrir descompensaciones. Es posible desarrollar sistemas capaces de detectar señales tempranas de alerta, lo que permite actuar antes de que la situación se vuelva grave. Además, la IA puede utilizarse como una herramienta de monitorización remota, facilitando un seguimiento más preciso del estado del paciente. También resulta útil en la interpretación de enfermedades cardíacas complejas, así como en la identificación de otras afecciones relacionadas con la insuficiencia cardíaca (“Aplicaciones de la inteligencia artificial en cardiología,” 2019).

4.4. Modelos de predicción y clasificación aplicados al diagnóstico clínico

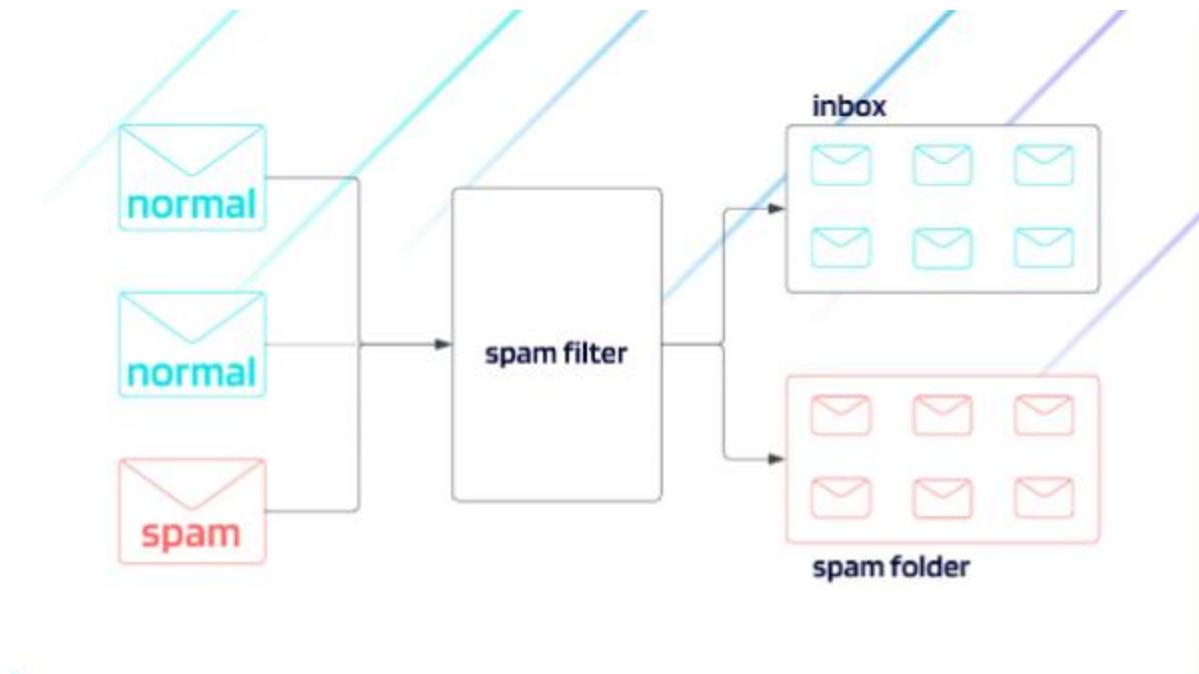
Según Mora Pineda (2022), los métodos de aprendizaje automático pueden usarse tanto para clasificación como para predicción, aunque la predicción suele aplicarse en contextos binarios (presente o ausente, sí o no, es o no es, etc.) y la clasificación en tareas con múltiples categorías, ósea determinar 1 resultado dentro de un grupo de 3 o más resultados posibles según la tarea a resolver. “Las técnicas de clasificación son una parte esencial de las aplicaciones de machine learning y minería de datos. Aproximadamente el 70 % de los problemas en ciencia de datos son problemas de clasificación” (DataCamp, 2024).

La distribución de Bernoulli predice la probabilidad de que un evento ocurra o no ocurra, utiliza 0 y 1, el trabajo basado en clases binarias suele utilizar algoritmos más sencillos y rápidos (regresión logística, arboles de decisión, etc). Por el contrario, la clasificación multiclas no suele tener una noción de normal o anormal, en este caso se refiere a las tareas de clasificación que tienen más de dos etiquetas de clase, es importante ser cuidadoso al momento del etiquetado, especialmente cuando se tiene un mayor número de clases (Por ejemplo, rojo = 0, azul = 1,

amarillo = 2), esta también puede utilizar una amplia gama de posibles algoritmos (Bosques aleatorios, Naïve-Bayes, etc) (Elearning Actual, 2022).

Figura 6

Ejemplo de clasificación binaria



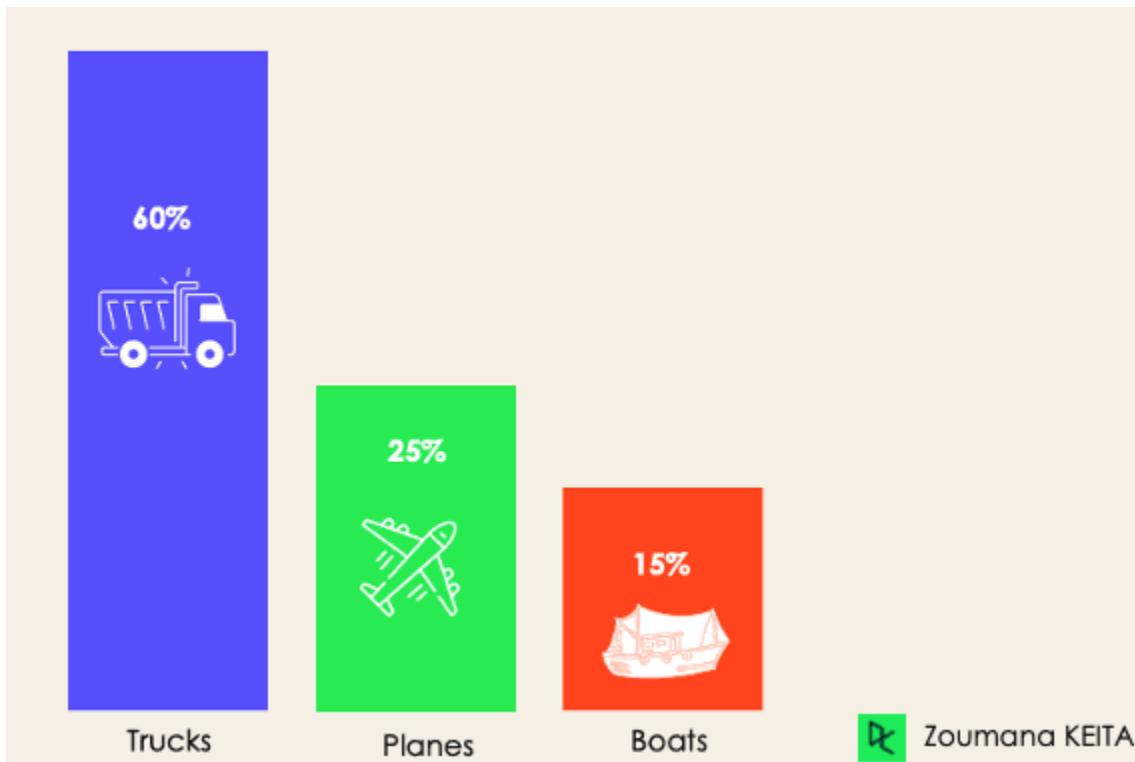
Nota. Filtros de correo electrónico (spam, no spam). Tomado de 4 tipos de tareas de clasificación en el machine learning, por elearningactual.com, 2022. <https://elearningactual.com/4-tipos-de-tareas-de-clasificacion-en-el-machine-learning/>

Un modelo de predicción clínica se desarrolla generalmente para estimar probabilidades, a menudo llamadas simplemente riesgo, para pacientes individuales con el objetivo de informarles a ellos, a sus familiares y a los profesionales de la salud sobre el diagnóstico o el pronóstico, para ayudar a tomar mejores decisiones (compartidas) sobre pruebas y tratamientos, o para realizar estratificaciones de riesgo para ensayos terapéuticos (van Smeden et al., 2021, p. 142). Según DataCamp (2024), el uso de modelos predictivos convencionales, como árboles de

decisión o regresión logística, puede no ser eficaz en conjuntos de datos desequilibrados, ya que tienden a favorecer la clase mayoritaria y a interpretar las clases minoritarias como ruido.

Figura 7

Clasificación desequilibrada



Nota. Problema de clasificación desequilibrada. Tomado de Clasificación en Machine Learning:

Una guía completa para principiantes, por DataCamp (2024),

<https://www.datacamp.com/es/blog/classification-machine-learning>

4.4.1 Métricas de evaluación de modelos predictivos

El rendimiento del modelo se evalúa típicamente mediante la precisión, la recuperación, la exactitud, F1 y el área bajo la curva ROC (característica operativa del receptor). Las métricas se calculan utilizando los resultados de la matriz de confusión para cada clase: TP

para verdaderos positivos , TN para verdaderos negativos, FN para falsos negativos y FP para falsos positivos (*Fan et al., 2021, p. 45*)

Tabla 3

Métricas de evaluación de modelos de clasificación

Métrica	¿Qué mide?	Fórmula
Precisión (Accuracy)	Proporción de predicciones correctas sobre el total de casos.	$(TP + TN) / (TP + TN + FP + FN)$
Precisión positiva (Precision)	Qué proporción de los casos predichos como positivos realmente lo son.	$TP / (TP + FP)$
Recuperación (Recall / Sensibilidad)	Qué proporción de los casos positivos reales fueron correctamente identificados.	$TP / (TP + FN)$
Especificidad	Qué proporción de los casos negativos reales fueron correctamente identificados.	$TN / (TN + FP)$
F1-score	Media armónica entre la precisión y la sensibilidad.	$2 \times (\text{Precision} \times \text{Recall}) / (\text{Precision} + \text{Recall})$

AUC - ROC	Capacidad del modelo para distinguir entre clases (positivo vs. negativo).	Área bajo la curva ROC basada en TPR y FPR.
------------------	--	---

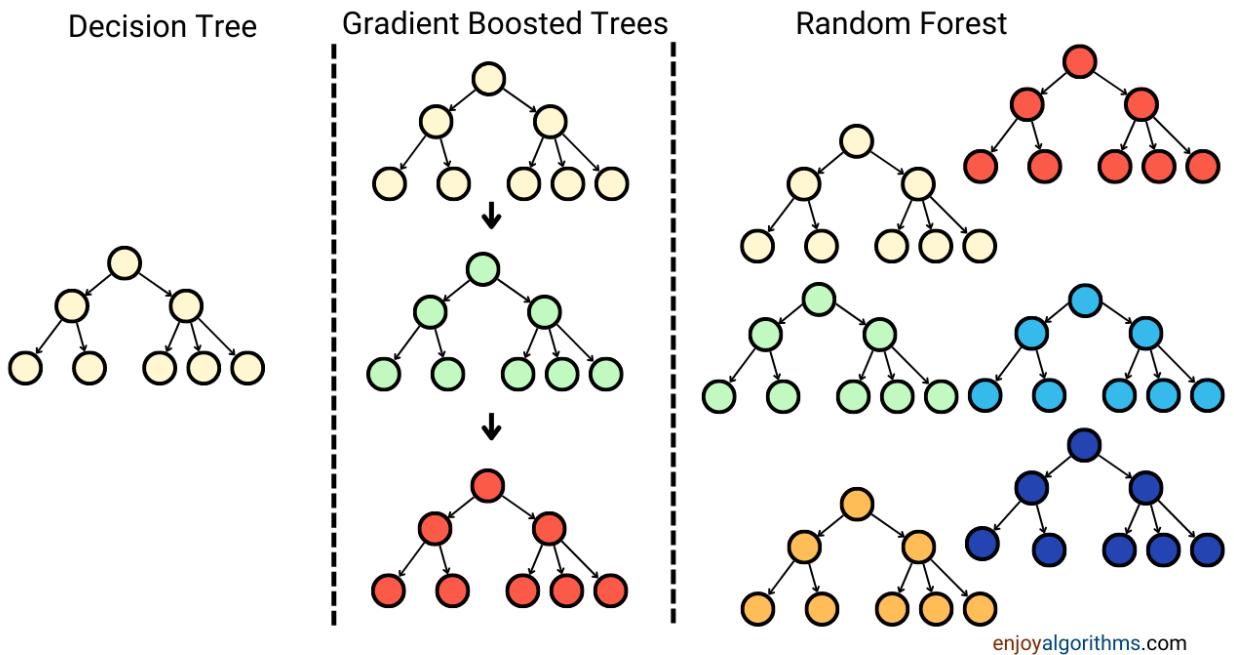
Nota. Elaboración propia, 2025.

4.4.2. Modelos Predictivos Basados en Árboles de Decisión

Según Ciencia de Datos (2017), los métodos basados en árboles han ganado relevancia dentro del ámbito predictivo debido a su capacidad para ofrecer buenos resultados en problemas de naturaleza diversa. Estos modelos, a diferencia de otros enfoques que utilizan una única ecuación válida para todo el espacio muestral, son especialmente útiles cuando el caso de uso involucra múltiples variables predictoras que interactúan entre sí de manera compleja y no lineal. En estos contextos, resulta difícil identificar un único modelo global que represente adecuadamente dichas relaciones. Los algoritmos como Random Forest abordan esta dificultad combinando múltiples árboles de decisión, cada uno entrenado con una muestra ligeramente distinta del conjunto original mediante técnicas de remuestreo. La predicción final se obtiene agregando las predicciones individuales de cada árbol, lo que mejora la robustez y la generalización del modelo. Por otro lado, el enfoque Gradient Boosting también emplea una combinación de árboles, pero lo hace de manera secuencial: cada nuevo árbol se entrena para corregir los errores cometidos por los árboles anteriores, afinando así progresivamente la precisión del modelo final.

Figura 8

Comparación modelos basados en arboles



Nota. Tomado de Brandon Wong, 2023, en Medium. Disponible en:

<https://medium.com/@brandon93.w/decision-tree-random-forest-and-xgboost-an-exploration-into-the-heart-of-machine-learning-90dc212f4948>

5. Metodología

Este estudio está enfocado en la aplicación de herramientas de Inteligencia Artificial para la investigación médica en el ámbito de la insuficiencia cardíaca y su relación con las enfermedades hepáticas. Se trata de un estudio cuantitativo, observacional y retrospectivo, centrado en el registro y análisis de datos clínicos de pacientes 101 evaluados con diagnóstico de insuficiencia cardíaca (IC).

La recolección de datos se lleva a cabo en el marco del proyecto de investigación “Utilidad de la elastografía de transición en el diagnóstico de la insuficiencia cardíaca estadio C”, desarrollado por el Instituto de Investigación Sanitaria INCLIVA en colaboración con el Hospital Clínico Universitario de Valencia, en el contexto de prácticas profesionales y cooperación investigadora.

Se incluirán en el estudio todos los pacientes con diagnóstico confirmado de IC que hayan sido sometidos a una exploración hepática mediante FibroScan. A partir de estos registros, se obtendrán variables relacionadas con la rigidez hepática y la esteatosis, así como variables clínicas y analíticas relevantes asociadas a la insuficiencia cardíaca.

5.1. Población y Muestra

- Pacientes con diagnóstico de IC confirmada por criterios clínicos y ecocardiográficos.
- Exploración hepática realizada con FibroScan.
- Datos clínicos relacionados a la congestión en IC, incluyendo biomarcadores y comorbilidades.

5.2. Variables a Analizar

En este estudio se analizarán diversas variables relevantes tanto desde el punto de vista clínico como diagnóstico. Las principales medidas provienen del FibroScan, que permite obtener valores para evaluar la rigidez hepática (medida en kPa) y la esteatosis hepática (medida mediante el parámetro CAP, en dB/m). Además, se incorporarán variables clínicas obtenidas de la unidad de insuficiencia cardíaca, especialmente aquellas asociadas con el desarrollo o progresión de la fibrosis hepática.

A continuación, se describen las variables consideradas en el análisis:

Tabla 4

Variables Dependientes

Nombre de la Variable	Utilidad / Qué calcula	Valores o Rangos típicos
Grado de fibrosis	Clasificación de fibrosis hepática derivada de la rigidez hepática (kPa) medida con FibroScan	F0–F4 (ej. F0 < 6.5; F4 > 14.5 kPa)

Nota: Elaboración propia, 2025.

Tabla 5*Variables Independientes (Parámetros Demográficos y Fisiológicos)*

Nombre de la	Utilidad / Qué calcula	Valores o Rangos
Variable		típicos
Edad (años)	Factor de riesgo general; puede influir en fibrosis	Continuo (ej. 18–90 años)
Sexo	Diferencias fisiológicas en el desarrollo de enfermedades hepáticas	Masculino / Femenino
Peso (kg)	Asociado a IMC y riesgo metabólico; influye en la esteatosis	Continuo (ej. 40–150 kg)

Nota. Elaboración propia, 2025.

Tabla 6*Variables Clínicas (Comorbilidades)*

Nombre de la	Utilidad / Qué calcula	Valores o Rangos típicos
Variable		
Grado de insuficiencia tricuspídea	Evalúa la severidad del reflujo de sangre a través de la válvula tricúspide	0 = No insuficiencia (no IT) , 1 = Ligera , 2 = Ligera-moderada, 3 = Moderada ,4 = Severa
Historia de fibrilación auricular	Comorbilidad frecuente en pacientes con IC	Sí / No
Diabetes mellitus	Comorbilidad clave en daño hepático metabólico	Sí / No
Enfermedad renal crónica (ERC)	Relacionada con inflamación sistémica y fibrosis avanzada	Sí / No
Cardiopatía valvular	Puede alterar la hemodinámica y asociarse a daño hepático	Sí / No
Prótesis valvular	Indicador de intervenciones cardíacas previas	Sí / No

Nota. Elaboración propia, 2025.

Tabla 7*Tratamiento Farmacológico Previo*

Nombre de la Variable	Utilidad / Qué calcula	Valores o Rangos típicos
Diuréticos de asa (furosemida oral)	Afectan el balance hídrico; relevantes en pacientes con IC	Sí / No – Dosis típica: 20–80 mg/día
Furosemida combinada (tiazida, Edemox)	Combinaciones comunes en IC avanzada para control de volumen	Sí / No – Edemox: 50–100 mg/día
ARM (antagonistas mineralocorticoides)	Moduladores del sistema RAAS; podrían influir en fibrosis	Sí / No – Espironolactona: 25–50 mg/día
iSGLT2 (inhibidores SGLT2)	Efecto beneficioso sobre función hepática y cardíaca	Sí / No – Ej.: Dapagliflozina 10 mg/día
Agonistas GLP-1 RA	Mejoran resistencia a insulina y pueden reducir esteatosis	Sí / No – Ej.: Semaglutida 0.5–2 mg/semana

Nota. Elaboración propia, 2025.

Tabla 8*Parámetros Analíticos / Laboratorio*

Nombre de la Variable	Utilidad / Qué calcula	Valores o Rangos típicos
GGT (Gamma-glutamil transferasa)	Enzima hepática elevada en daño hepático o colestasis	U/L (ej. 10–70 U/L)
ALT / GPT (Alanina aminotransferasa)	Enzima hepática; marcador de daño hepatocelular	U/L (ej. 7–56 U/L)
Recuento plaquetario	Plaquetas reducidas pueden asociarse a fibrosis avanzada	mil/mm ³ (ej. 150–400 mil/mm ³)
NT-proBNP	Biomarcador de estrés cardíaco; niveles elevados en IC avanzada	pg/mL (ej. <125 normal; >300 sugiere IC)
CA125	Marcador utilizado en oncología, pero elevado también en IC y ascitis	U/mL (ej. <35 U/mL normal)
CAP (dB/m)	Medición de esteatosis hepática mediante FibroScan	Continuo (ej. 100–400 dB/m)

Nota. Elaboración propia, 2025.

5.3. Implementación Analítica y Técnicas de Inteligencia Artificial

5.3.1. Exploración de Datos (EDA)

Resumen General del Conjunto de Datos

Para obtener una primera impresión de los datos, es útil revisar la información básica del DataFrame.

```
# Obtener información general sobre el DataFrame
print('Información general:')
data.info()
```

info(): Proporciona un resumen del DataFrame, incluyendo: número de entradas, tipo de datos, y cantidad de valores no nulos.

```
#Descripción de datos numéricos
print('Descripción:')
data.describe()
```

describe(): Genera estadísticas descriptivas (media, desviación estándar, mínimos, máximos, cuartiles) para las columnas numéricas.

Distribución por Grado de Fibrosis (Clasificación Multiclasa)

Esta visualización te permitirá entender la cantidad de pacientes en cada uno de los 5 grados de fibrosis (F0, F1, F2, F3, F4). Esto es crucial para identificar si hay un desequilibrio de clases, lo cual es común en datos médicos.

Tipo de Gráfica: Gráfico de barras (Count Plot).

```
#Conteo Grado de Fibrosis
```

```

plt.figure(figsize=(8, 5))
ax = sns.countplot(x='grado_fibrosis', data=data)
plt.title('Grado de Fibrosis por pacientes', fontsize=14)
plt.xlabel('Grado de Fibrosis')
plt.ylabel('Cantidad de Pacientes')

#Aregar etiquetas con la cantidad en cada barra
for p in ax.patches:
    height = p.get_height()
    if height > 0:
        ax.text(
            p.get_x() + p.get_width() / 2, # posición x (centro de barra)
            height + 0.5, # posición y (un poco arriba de
            la barra)
            int(height), # texto a mostrar (cantidad)
            ha='center', # alineación horizontal
            fontsize=10
        )
plt.show()

```

Ánalisis: Observar si alguna clase tiene significativamente menos muestras que otras (clases minoritarias).

Distribución Binaria de Fibrosis (Sí/No)

Visualizar la proporción de pacientes con y sin fibrosis.

Tipo de Gráfica: Gráfico de barras (Count Plot).

```

#Convertir grados 0 = ausencia de fibrosis(0), 1 a 4 = presencia de
fibrosis(1)
fibrosis_binaria = data['grado_fibrosis'].apply(lambda x: 1 if x > 0 else 0)

# Graficar countplot
plt.figure(figsize=(8, 6))
ax = sns.countplot(data=data, x=fibrosis_binaria)

plt.title("Presencia de fibrosis", fontsize=16)
plt.xlabel('Fibrosis(no=0 | si=1)', fontsize=14)
plt.ylabel('Cantidad de pacientes', fontsize=14)
plt.xticks(fontsize=12)
plt.yticks(fontsize=12)

#Aregar etiquetas con la cantidad en cada barra
for p in ax.patches:
    ax.annotate(f'{int(p.get_height())}', (p.get_x() + p.get_width() / 2,
    p.get_height()),
                ha='center', va='center', xytext=(0, 5), textcoords='offset
    points', fontsize=12)

```

```
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```

Análisis: Evalúa el nivel de desequilibrio entre las clases "Ausencia = 0" y "Presencia = 1".

Análisis de Correlación entre Variables

La matriz de correlación es una herramienta poderosa para entender las relaciones lineales entre las variables numéricas, incluyendo su relación con la variable objetivo.

Tipo de Gráfica: Mapa de calor (Heatmap) de la matriz de correlación.

```
#Matriz de correlación Grado de Fibrosis
#Excluir E [kPa] alta relación con grado de fibrosis
data_sin_E = data.drop(columns=['E[kPa]'])
# Visualización de la correlación entre variables
plt.figure(figsize=(12, 10))
sns.heatmap(data_sin_E.corr(), annot=True, fmt='.2f', cmap='coolwarm')
plt.title('Matriz de correlación Grado de Fibrosis')
plt.show()

#Matriz de correlación Presencia de Fibrosis
#Mostrar grados 0 = ausencia de fibrosis, 1 a 4 = presencia de fibrosis
data_copia = data.copy()
data_copia['fibrosis_binaria'] = data['grado_fibrosis'].apply(lambda x: 1 if
x > 0 else 0)
#Excluir E [kPa], y grado_fibrosis alta relación con presencia de fibrosis
data_sin_e = data_copia.drop(columns=['E[kPa]', 'grado_fibrosis'])
# Visualización de la correlación entre variables
plt.figure(figsize=(12, 10))
sns.heatmap(data_sin_e.corr(), annot=True, fmt='.2f', cmap='coolwarm')
plt.title('Matriz de correlación Presencia de Fibrosis')
plt.show()
```

Correlaciones entre predictores: Identifica si hay variables predictoras altamente correlacionadas entre sí (multicolinealidad).

5.3.2. Modelado

En esta sección, se detalla la metodología de aprendizaje automático supervisado empleada para abordar la tarea de clasificación de la fibrosis hepática. Inicialmente, se exploró

un enfoque de clasificación multiclasa, y tras evaluar sus resultados, se optó por una estrategia de clasificación binaria más enfocada y práctica.

Para la evaluación del rendimiento, se emplearán las siguientes métricas clave:

- Precisión (Accuracy): Proporción de predicciones correctas sobre el total de casos.
- Sensibilidad (Recall): Capacidad del modelo para identificar correctamente los casos positivos (presencia de fibrosis).
- Especificidad: Capacidad del modelo para identificar correctamente los casos negativos.
- ROC-AUC (Receiver Operating Characteristic - Area Under the Curve): Mide la capacidad del modelo para distinguir entre las clases, siendo un indicador robusto del rendimiento general del clasificador.

La implementación del modelo de clasificación se llevó a cabo utilizando el lenguaje de programación Python y diversas librerías especializadas en aprendizaje automático

Exploración Inicial: Clasificación Multiclasa de Grado de Fibrosis (F0-F4)

El primer enfoque de modelado supervisado tuvo como objetivo predecir directamente los diferentes grados de fibrosis hepática (F0, F1, F2, F3, F4) utilizando un algoritmo de Gradient Boosting. Esta aproximación buscaba una granularidad detallada en la detección de la severidad de la fibrosis basándose únicamente en parámetros clínicos y demográficos, excluyendo mediciones de FibroScan.

Implementación del Modelo Multiclasa

El proceso inició con la preparación de los datos y la división en conjuntos de entrenamiento y prueba, asegurando que la variable objetivo fuera el grado_fibrosis original.

```

from sklearn.model_selection import train_test_split
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.metrics import classification_report, confusion_matrix,
accuracy_score
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.metrics import ConfusionMatrixDisplay

#Separación de variables
# Predictores (X): Se eliminan 'E[kPa]' (rigidez hepática) y
# 'grado_fibrosis' para simular la ausencia de Fibroscan
X = data.drop(columns=['E[kPa]', 'grado_fibrosis'])
#Variable objetivo (y): El grado de fibrosis multiclase (0 a 4)
y = data['grado_fibrosis']

#División entrenamiento/prueba 30%
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42, stratify=y
)

```

X = Se define X (variables predictoras) excluyendo las mediciones de FibroScan (E[kPa]) y la propia variable objetivo.

y = La variable objetivo grado_fibrosis con sus categorías (0 a 4).

El modelo GradientBoostingClassifier fue configurado con parámetros específicos y entrenado con el conjunto de datos multiclase:

```

#Modelo Gradient Boosting multiclase
gb_model_multi = GradientBoostingClassifier(
    n_estimators=150,          # Número de árboles en el ensamble
    learning_rate=0.08,        # Contribución de cada árbol
    max_depth=4,              # Profundidad máxima de cada árbol
    min_samples_split=4,       # Mínimo de muestras para dividir un nodo
    subsample=0.9,             # Fracción de muestras para cada árbol
    random_state=42
)
gb_model_multi.fit(X_train, y_train)

# Realización de predicciones sobre el conjunto de prueba
y_pred = gb_model_multi.predict(X_test)

```

Parámetros del Modelo: n_estimators, learning_rate, max_depth, min_samples_split, y subsample fueron ajustados para optimizar el rendimiento del modelo en esta tarea multiclase.

Evaluación del Modelo Multiclasificación

```
#Evaluación del modelo
print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nReporte de clasificación:\n", classification_report(y_test, y_pred))

#Matriz de confusión
cm = confusion_matrix(y_test, y_pred)
disp = ConfusionMatrixDisplay(confusion_matrix=cm)
disp.plot(cmap='Blues', values_format='d')
plt.title("Matriz de Confusión - Grado de Fibrosis")
plt.show()

#Importancia de variables
importances = pd.Series(gb_model_multi.feature_importances_, index=X.columns)
importances.sort_values().plot(kind='barh', figsize=(8, 6),
title="Importancia de variables GradientBoostingClassifier multiclasificación")
# Las variables más importantes de este modelo multiclasificación también pueden servir como base para la siguiente etapa
#top_vars = importances.sort_values(ascending=False).head(10).index
plt.tight_layout()
plt.show()

Mostrar evaluación, matriz de confusión, importancia de variables.
```

Visualizar la curva ROC (Receiver Operating Characteristic) y el Área Bajo la Curva (AUC). Esto permite entender mejor la capacidad del modelo para distinguir entre una clase y las demás.

```
#Calcular curva ROC y AUC multiclasificación
#Binarizar etiquetas para el cálculo de ROC multiclasificación
from sklearn.preprocessing import label_binarize
from sklearn.metrics import roc_curve, auc
from sklearn.multiclass import OneVsRestClassifier
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import make_pipeline
#Grados son 0, 1, 2, 3, 4.
y_test_bin = label_binarize(y_test, classes=np.unique(y_test))
y_score = gb_model_multi.predict_proba(X_test)

#Calcular curva ROC y AUC para cada clase (enfoque One-vs-Rest)
fpr = dict() #False Positive
tpr = dict() #True Positive
roc_auc = dict() #Area Under the Curve
n_classes = y_test_bin.shape[1] #Número total de clases
```

```

for i in range(n_classes):
    fpr[i], tpr[i], _ = roc_curve(y_test_bin[:, i], y_score[:, i])
    roc_auc[i] = auc(fpr[i], tpr[i])

for i in range(n_classes):
    fpr[i], tpr[i], _ = roc_curve(y_test_bin[:, i], y_score[:, i])
    roc_auc[i] = auc(fpr[i], tpr[i])

# Graficar todas las curvas ROC
plt.figure(figsize=(8, 6))
colors = ['blue', 'orange', 'green', 'red', 'purple']
for i in range(n_classes):
    plt.plot(fpr[i], tpr[i], color=colors[i], label=f'Clase {i} (AUC = {roc_auc[i]:.2f})')

plt.plot([0, 1], [0, 1], 'k--')
plt.xlabel('Falsos Positivos')
plt.ylabel('Verdaderos Positivos')
plt.title('Curva ROC Multiclas - Grado de Fibrosis')
plt.legend(loc='lower right')
plt.grid()
plt.tight_layout()
plt.show()

```

Calcular los puntos de la curva ROC (TPR vs FPR) y el área bajo esa curva para cada clase.

Transición al Modelado Principal: Clasificación Binaria de Fibrosis

(Presencia/Ausencia)

Dada la complejidad y los resultados no óptimos de la clasificación multiclas (como evidenciaron las métricas y las curvas ROC por clase), se redefinió la variable objetivo para una tarea de clasificación binaria: predecir la presencia ($\text{fibrosis} > 0$) o ausencia ($\text{fibrosis} = 0$) de fibrosis. Este enfoque simplifica el problema, orientando el modelo a una pregunta de mayor relevancia clínica para la detección temprana. El siguiente apartado detallará la implementación de este modelo de Gradient Boosting binario.

Modelado Principal: Clasificación Binaria con Gradient Boosting

Para la tarea de clasificación binaria, se implementó un nuevo modelo de Gradient Boosting enfocado en la predicción de la presencia o ausencia de fibrosis. Los pasos de

preparación de datos, entrenamiento y evaluación se ajustaron a esta nueva variable objetivo binaria.

Preparación de Datos y División de Variables (Clasificación Binaria)

En esta fase, se transformó la variable de grado de fibrosis en una binaria (fibrosis_binaria).

```
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.ensemble import GradientBoostingClassifier
from sklearn.model_selection import train_test_split
from sklearn.metrics import classification_report, confusion_matrix,
accuracy_score
from sklearn.utils import compute_sample_weight
from sklearn.metrics import ConfusionMatrixDisplay

#Separación de variables
#Crear columna 'fibrosis_binaria' donde, Fibrosis 0 = ausencia, 1= presencia
data['fibrosis_binaria'] = data['grado_fibrosis'].apply(lambda x: 1 if x > 0
else 0)

#Predictores (X): Se eliminan 'E[kPa]' (rigidez hepática) y 'grado_fibrosis'
para simular la ausencia de Fibroscan
X = data.drop(columns=['fibrosis_binaria', 'E[kPa]', 'grado_fibrosis'])
#Variable objetivo (y): La nueva columna fibrosis binaria
y = data['fibrosis_binaria']

Se crea una nueva columna fibrosis_binaria (0 = ausencia, 1 = presencia)
X = variables predictoras) eliminando la columna objetivo (fibrosis_binaria), y las variables
relacionadas con FibroScan (E[kPa] y grado_fibrosis) para mantener el enfoque en variables clínicas.
y = fibrosis_binaria' variable objetivo binaria.
```

División del Conjunto de Datos y Ponderación de Clases

Para una evaluación robusta del modelo, el conjunto de datos se divide en subconjuntos de entrenamiento y prueba. Adicionalmente, se implementa una ponderación de clases para mitigar el sesgo que podría surgir de un desequilibrio entre las clases de "presencia" y "ausencia" de fibrosis.

```
#División entrenamiento/prueba 30%
X_train, X_test, y_train, y_test = train_test_split(
    X, y, test_size=0.3, random_state=42, stratify=y
)

#Ponderar clases para evitar sesgo
sample_weights = compute_sample_weight(class_weight='balanced', y=y_train)
Calcula pesos para cada muestra en el conjunto de entrenamiento, asignando un peso mayor a las
clases minoritarias para balancear su contribución durante el entrenamiento.
```

Entrenamiento del Modelo Binario

Se instancia y entrena el modelo `GradientBoostingClassifier` con parámetros ajustados para optimizar su rendimiento y evitar el sobreajuste, buscando un equilibrio entre capacidad de aprendizaje y generalización.

```
#Modelo Gradient Boosting con ajustes
gb_model = GradientBoostingClassifier(
    n_estimators=100,          # Número de árboles en el ensamble
    learning_rate=0.09,        # Contribución de cada árbol
    max_depth=3,              # Profundidad máxima de cada árbol
    min_samples_split=4,      # Número mínimo de muestras requeridas para
    dividir un nodo interno
    subsample=0.95,           # Fracción de muestras usadas para ajustar los
    estimadores individuales
    random_state=42
)

#Entrenamiento del modelo, aplicando los pesos de muestra calculados
gb_model.fit(X_train, y_train, sample_weight=sample_weights)
Los parámetros como n_estimators, learning_rate, max_depth, min_samples_split, y subsample se
seleccionan para controlar la complejidad del modelo y mejorar su capacidad de generalización.
```

Predicciones y Evaluación del Modelo

Una vez entrenado, el modelo se utiliza para realizar predicciones sobre el conjunto de prueba. Se ajusta un umbral de clasificación para priorizar una mayor sensibilidad, y finalmente, se evalúa el rendimiento del modelo utilizando métricas clave y visualizaciones.

```

#Predicciones con umbral ajustado
y_probs = gb_model.predict_proba(X_test)[:, 1] # Obtener probabilidades de
la clase positiva
threshold = 0.4                                # Umbral de decisión ajustado
(para mayor sensibilidad)
y_pred = (y_probs >= threshold).astype(int) # Convertir probabilidades a
predicciones binarias

#Evaluación del modelo
print("Accuracy:", accuracy_score(y_test, y_pred))
print("\nReporte de clasificación:\n", classification_report(y_test,
y_pred))

#Matriz de confusión
cm = confusion_matrix(y_test, y_pred)
disp = ConfusionMatrixDisplay(confusion_matrix=cm,
display_labels=gb_model.classes_)

#Visualizar matriz
plt.figure(figsize=(6, 6))
disp.plot(cmap='Blues', values_format='d')
plt.title('Matriz de Confusión')
plt.show()

#Importancia de variables
importances = pd.Series(gb_model.feature_importances_, index=X.columns)
importances.sort_values().plot(kind='barh', figsize=(8, 6),
title="Importancia de variables")
# Las variables 10 más importantes de este modelo binario se usan como base
para la siguiente etapa
top_vars = importances.sort_values(ascending=False).head(10).index
plt.tight_layout()
plt.show()

```

threshold = 0.4: Un umbral más bajo puede aumentar la sensibilidad del modelo.
 Genera y visualiza la matriz de confusión, que detalla los verdaderos positivos, verdaderos negativos, falsos positivos y falsos negativos.
 Calcula y visualiza la importancia de cada variable predictora en el modelo, lo que ayuda a entender qué características son más influyentes en la predicción.

Evaluación del Modelo Binario, Curva ROC y AUC.

```

from sklearn.metrics import roc_curve, roc_auc_score

#Obtener probabilidades de clase 1
y_probs = gb_model.predict_proba(X_test)[:, 1] # Probabilidad de clase
positiva

```

```

#Calcular curva ROC
fpr, tpr, thresholds = roc_curve(y_test, y_probs)
roc_auc = roc_auc_score(y_test, y_probs)

#Graficar
plt.figure(figsize=(8, 6))
plt.plot(fpr, tpr, color='darkorange', lw=2, label=f'Curva ROC (AUC = {roc_auc:.2f})')
plt.plot([0, 1], [0, 1], color='navy', lw=2, linestyle='--') # Línea base
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.05])
plt.xlabel('Tasa de Falsos Positivos (FPR)')
plt.ylabel('Tasa de Verdaderos Positivos (TPR)')
plt.title('Curva ROC - GradientBoostingClassifier Binario')
plt.legend(loc="lower right")
plt.grid(True)
plt.tight_layout()
plt.show()

```

Calcula directamente el área bajo la curva ROC, una métrica clave para clasificadores binarios que mide la capacidad del modelo para distinguir entre las clases.

Refinamiento y Validación: Clasificación Binaria con Regresión Logística y

Variables Importantes

Para refinar aún más el modelo binario y aumentar la interpretabilidad, se decidió utilizar un modelo de Regresión Logística. Este modelo, más simple y directamente interpretable que Gradient Boosting, se aplicó utilizando únicamente las variables 10 más importantes identificadas por el modelo de Gradient Boosting binario principal. Este enfoque permite evaluar la capacidad predictiva de un subconjunto de características clave con un modelo de base, validando su relevancia.

Preparación de Datos y Escalado para Regresión Logística

Se seleccionaron las variables importantes y se escalaron los datos, un paso crucial para algoritmos basados en distancias como la Regresión Logística, que son sensibles a la escala de las características.

```

from sklearn.model_selection import train_test_split
from sklearn.linear_model import LogisticRegression
from sklearn.metrics import classification_report, accuracy_score,
confusion_matrix, roc_auc_score, roc_curve
from sklearn.preprocessing import StandardScaler
import matplotlib.pyplot as plt
from sklearn.metrics import ConfusionMatrixDisplay

#Definir variables (usando top_vars del modelo binario)
X = data[top_vars] # top_vars se han extraido del modelo binario de Gradient
Boosting anterior
y = data['fibrosis_binaria']

#Escalar y separar datos
scaler = StandardScaler()
X_scaled = scaler.fit_transform(X) # Escalar las columnas predictoras
seleccionadas

X_train, X_test, y_train, y_test = train_test_split(
    X_scaled, y, test_size=0.3, random_state=42
) # Dividir datos en entrenamiento/prueba 30%
X = se compone únicamente de las variables previamente identificadas como top_vars por el modelo
de Gradient Boosting binario.
La escala de los datos es esencial para la Regresión Logística.

```

Entrenamiento del Modelo de Regresión Logística

El modelo de Regresión Logística se inicializó con un balance de pesos para las clases y un número máximo de iteraciones, y luego se entrenó con los datos escalados.

```

#Entrenar modelo RL
model_rl = LogisticRegression(class_weight='balanced', max_iter=1000)
model_rl.fit(X_train, y_train)

```

Se utiliza `class_weight='balanced'` para manejar posibles desequilibrios en las clases.

Predicciones y Evaluación del Modelo de Regresión Logística

Las predicciones se realizaron sobre el conjunto de prueba, aplicando un umbral ajustado para balancear la sensibilidad. Finalmente, se calcularon y visualizaron las métricas de rendimiento, incluyendo la matriz de confusión y la curva ROC.

```
#Predicción con umbral ajustado
y_probs = model_rl.predict_proba(X_test)[:, 1] # Probabilidad de la clase positiva

#Ajustar umbral
threshold = 0.38 # Un umbral ajustado puede priorizar la sensibilidad, dependiendo del contexto clínico
y_pred_thresh = (y_probs >= threshold).astype(int) # Convertir probabilidades a predicciones binarias

#Mostrar Métricas
print("Accuracy:", accuracy_score(y_test, y_pred_thresh))
print("\nReporte de clasificación:\n", classification_report(y_test, y_pred_thresh))

#Matriz de confusión
cm = confusion_matrix(y_test, y_pred_thresh)
disp = ConfusionMatrixDisplay(confusion_matrix=cm,
display_labels=model_rl.classes_)

#Visualizar matriz
plt.figure(figsize=(6, 6))
disp.plot(cmap='Blues', values_format='d')
plt.title('Matriz de Confusión - Regresión Logística Binaria')
plt.show()

#Curva ROC
fpr, tpr, _ = roc_curve(y_test, y_probs)
auc = roc_auc_score(y_test, y_probs)

plt.figure(figsize=(6, 4))
plt.plot(fpr, tpr, label=f"AUC = {auc:.2f}")
plt.plot([0, 1], [0, 1], linestyle='--', color='gray', label='Clasificador aleatorio')
plt.xlabel('Tasa de Falsos Positivos (1 - Especificidad)')
plt.ylabel('Tasa de Verdaderos Positivos (Sensibilidad)')
plt.title('Curva ROC - Regresión Logística Binaria')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
```

Se utilizan las mismas funciones para evaluar la capacidad discriminativa del modelo de Regresión Logística.

6. Resultados

6.1. Presentación de los Hallazgos del Análisis Exploratorio de Datos (EDA)

Tabla 9

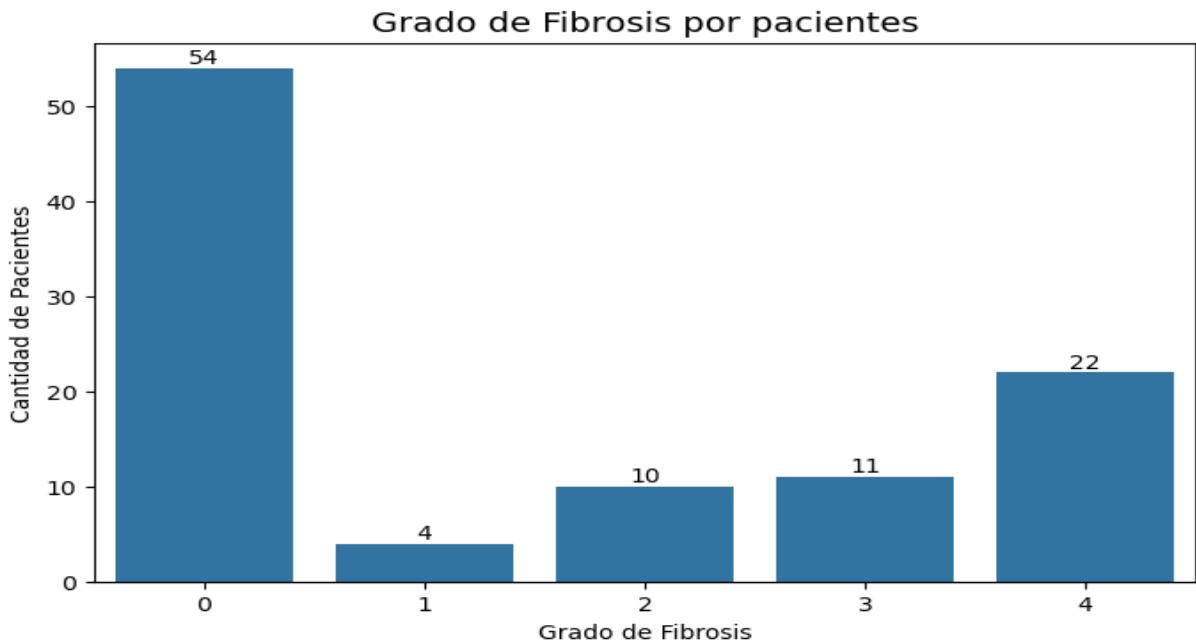
Información general:

#	Nombre en DataFrame	Nombre en Español	Tipo de Dato	Valores no nulos
0	sexo	Sexo	Entero (int64)	101
1	Edad	Edad	Entero (int64)	101
2	dm	Diabetes mellitus	Entero (int64)	101
3	cardiopatia_valvular	Cardiopatía valvular	Entero (int64)	101
4	protesis_valvular	Prótesis valvular	Entero (int64)	101
5	acxfa	Antecedentes de fibrilación auricular	Entero (int64)	101
6	irc	Enfermedad renal crónica (ERC)	Entero (int64)	101
7	furosemida_oral_dosis	Dosis de furosemida oral (mg/día)	Decimal (float64)	101
8	furo_otros	Furosemida combinada (tiazidas u otros)	Entero (int64)	101
9	arm	Antagonistas de mineralocorticoides (ARM)	Entero (int64)	101
10	isglt2	Inhibidores SGLT2	Entero (int64)	101
11	aGLP1	Agonistas del receptor GLP-1	Entero (int64)	101
12	fe	Fracción de eyección (FE)	Entero (int64)	101
13	it_grado	Grado de insuficiencia tricuspídea (leve/mod/sev)	Objeto (texto)	101
14	peso	Peso (kg)	Decimal (float64)	101
15	edemas	Presencia de edemas periféricos	Entero (int64)	101
16	gpt	ALT / GPT (Alanina aminotransferasa)	Entero (int64)	101
17	ggt	GGT (Gamma-glutamil transferasa)	Entero (int64)	101
18	plaquetas	Recuento plaquetario (mil/mm ³)	Entero (int64)	101
19	ca125	CA125 (U/mL)	Entero (int64)	101
20	ntprobnp	NT-proBNP (pg/mL)	Entero (int64)	101
21	CAP[dB/m]	CAP (dB/m) – Esteatosis por FibroScan	Entero (int64)	101
22	grado_fibrosis	Grado de fibrosis hepática (F0–F4)	Entero (int64)	101
23	E[kPa]	Rigidez hepática (E, en kPa) por FibroScan	Decimal (float64)	101

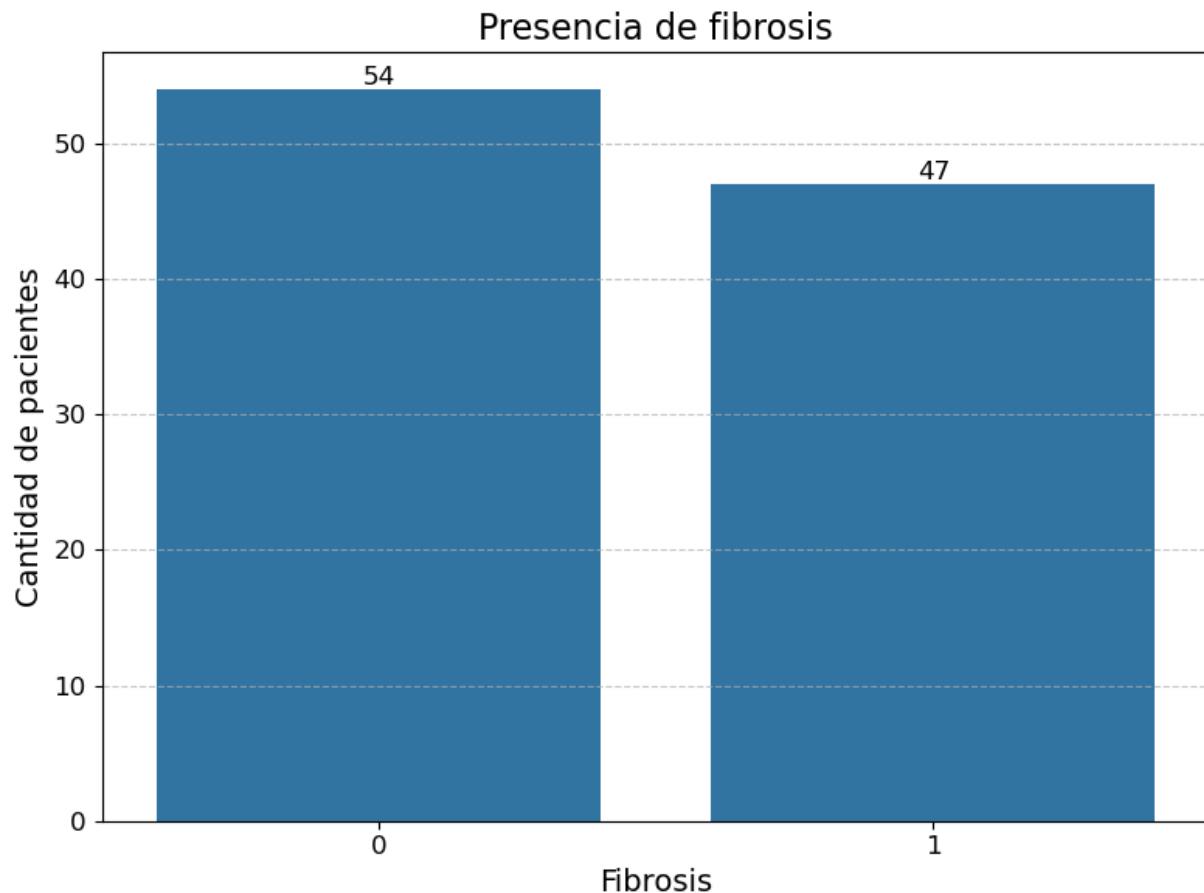
Nota. Exploración por columnas. Elaboración propia, 2025.

Distribución de Variables:**Figura 9**

Distribución de pacientes por grado de fibrosis



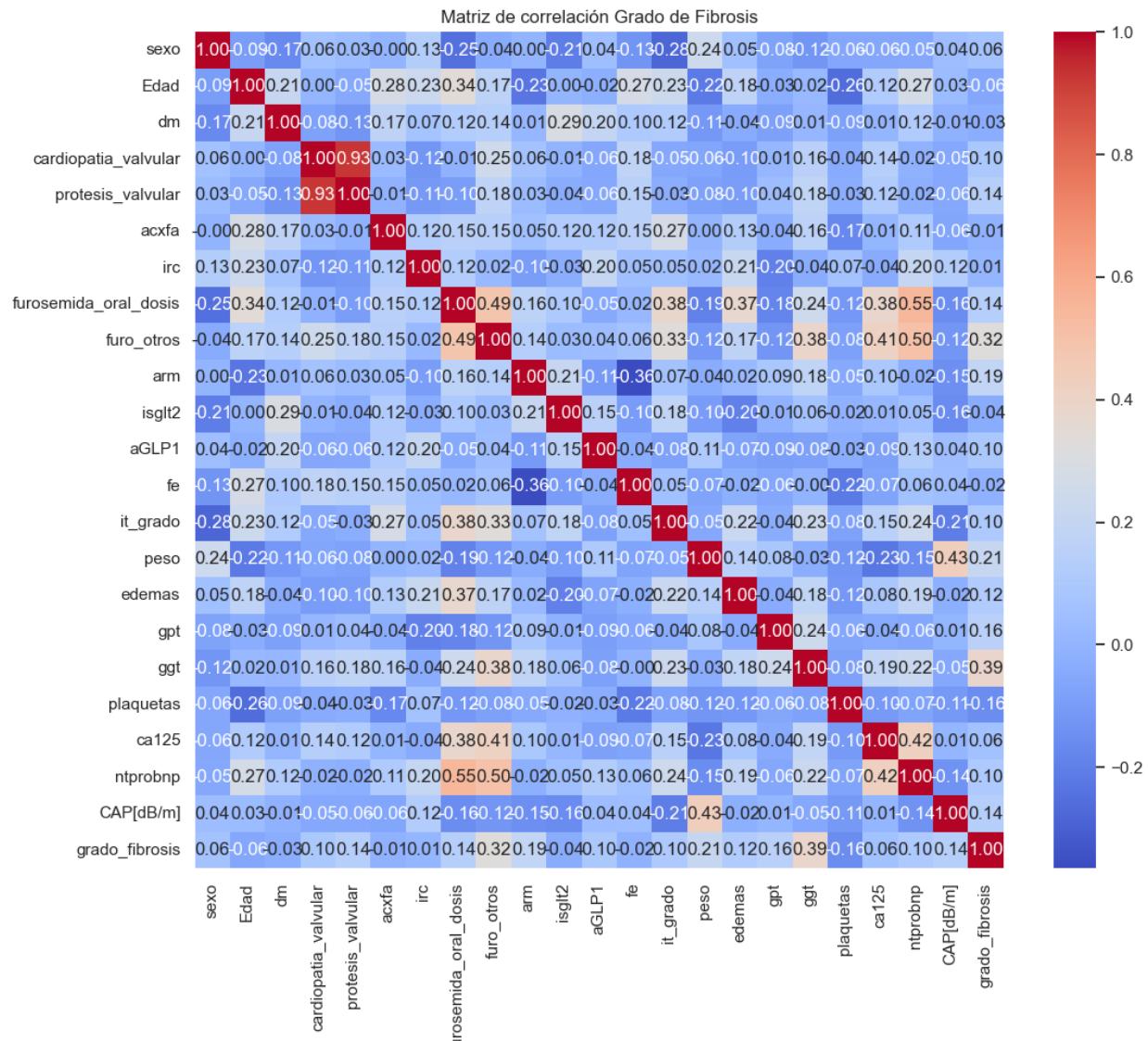
Nota: Tomado de código fuente, 2025.

Figura 10*Presencia/ausencia de fibrosis*

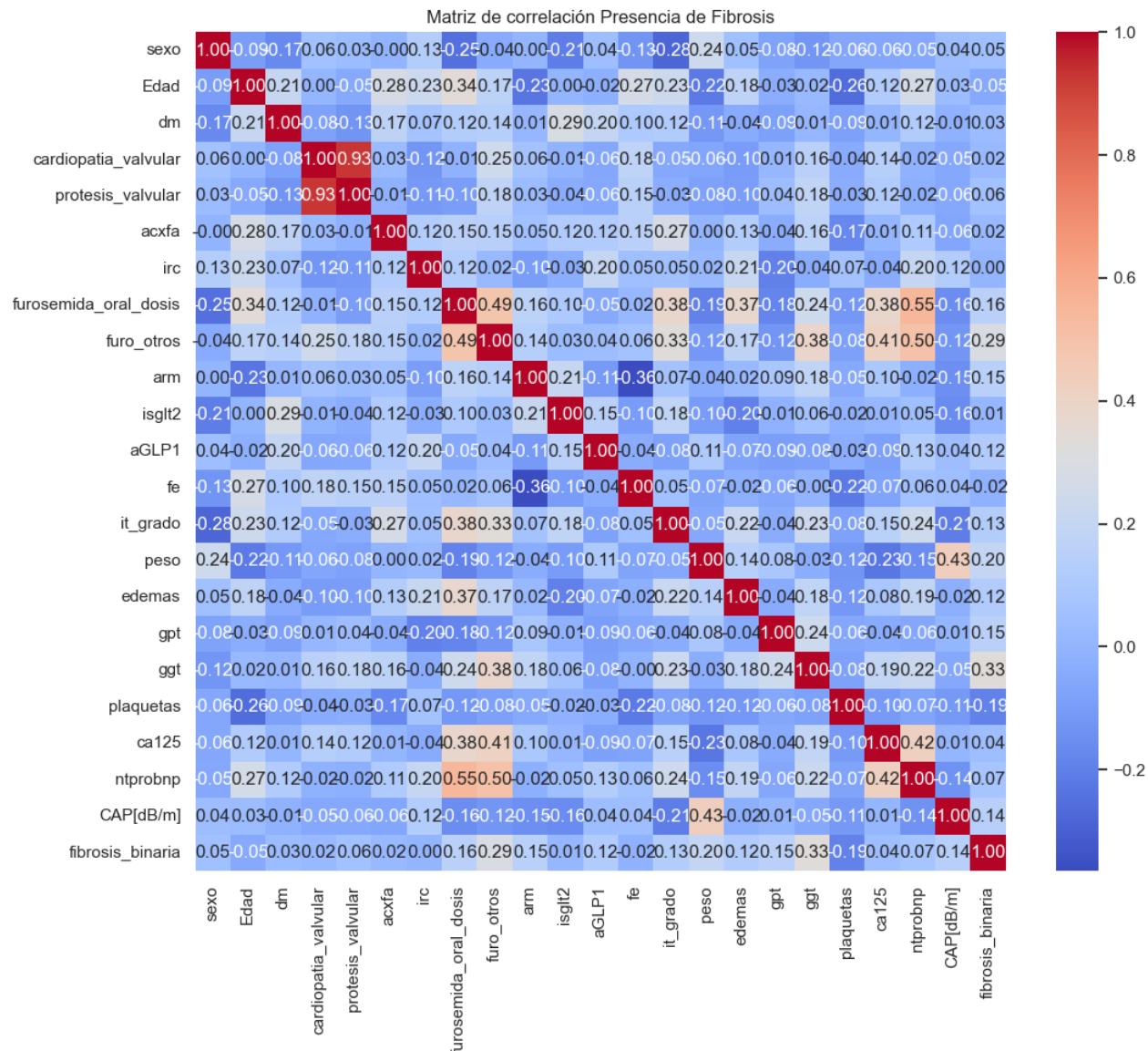
Nota. Tomado código fuente, 2025.

Correlaciones:

La matriz de correlación visualizada con un heatmap, destacando las relaciones entre las variables predictoras y la variable objetivo.

Figura 11*Grado de Fibrosis*

Nota. Tomado código fuente, 2025.

Figura 12*Presencia de Fibrosis*

Nota. Tomado de código fuente, 2025.

6.2. Resultados de los Modelos de Machine Learning

6.2.1. Resultados del Modelo de Clasificación Multiclas (Inicial)

Métricas de Rendimiento (Multiclas - Gradient Boosting)

Precisión (Accuracy), reporte de clasificación completo (Precision, Recall, F1-Score por clase).

- Exactitud (Accuracy) = 61%.

Tabla 10

Métricas de Rendimiento Modelo Multiclas

Clase	Precisión (Precision)	Sensibilidad (Recall)	F1-Score	Soporte (Support)
0	0.64	0.82	0.72	17
1	0.00	0.00	0.00	1
2	1.00	0.33	0.50	3
3	0.00	0.00	0.00	3
4	0.80	0.57	0.67	7

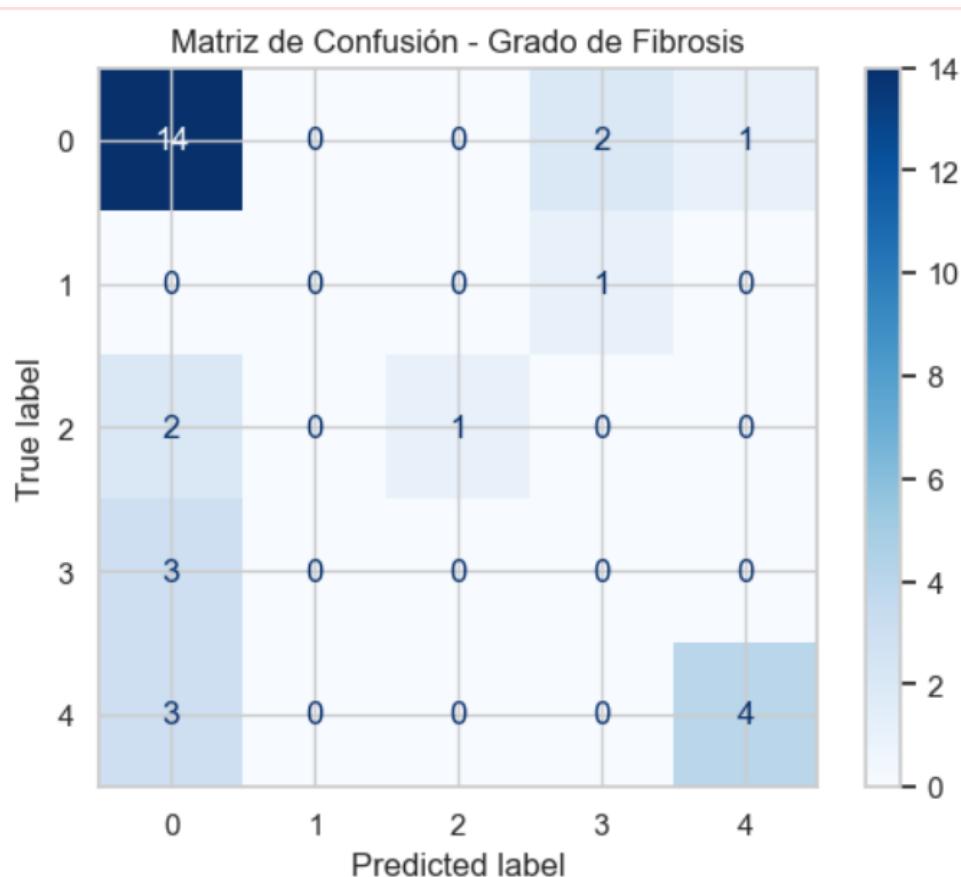
Nota. Tomado de código fuente, 2025.

Matriz de Confusión Multiclasé

La matriz de confusión para este modelo, que visualiza los aciertos y errores por cada clase de fibrosis.

Figura 13

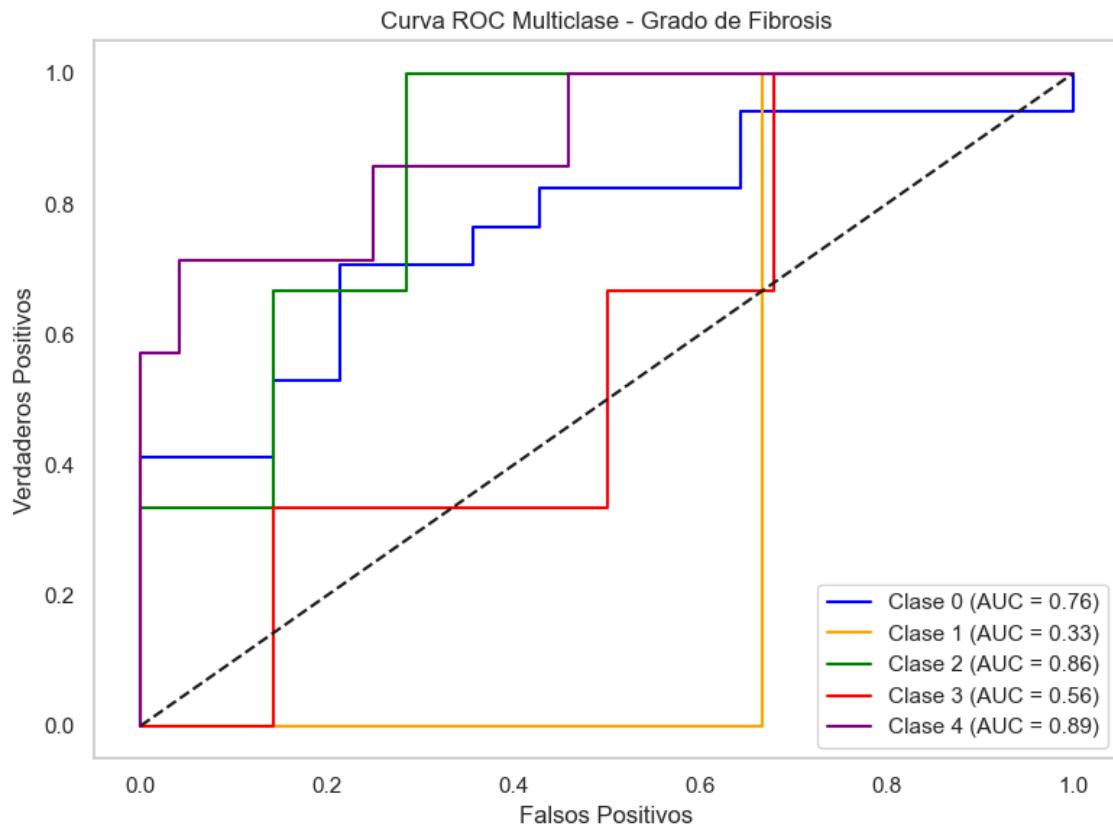
Matriz de Confusión - Grado de Fibrosis



Nota. Clasificación multiclasé (f0, f1, f2, f3, f4). Tomado de código fuente, 2025.

Curvas ROC y AUC Multiclasé

La gráfica de las curvas ROC multiclasé con los valores de AUC para cada clase.

Figura 14*Curvas Roc y AUC Clasificación Multiclas*

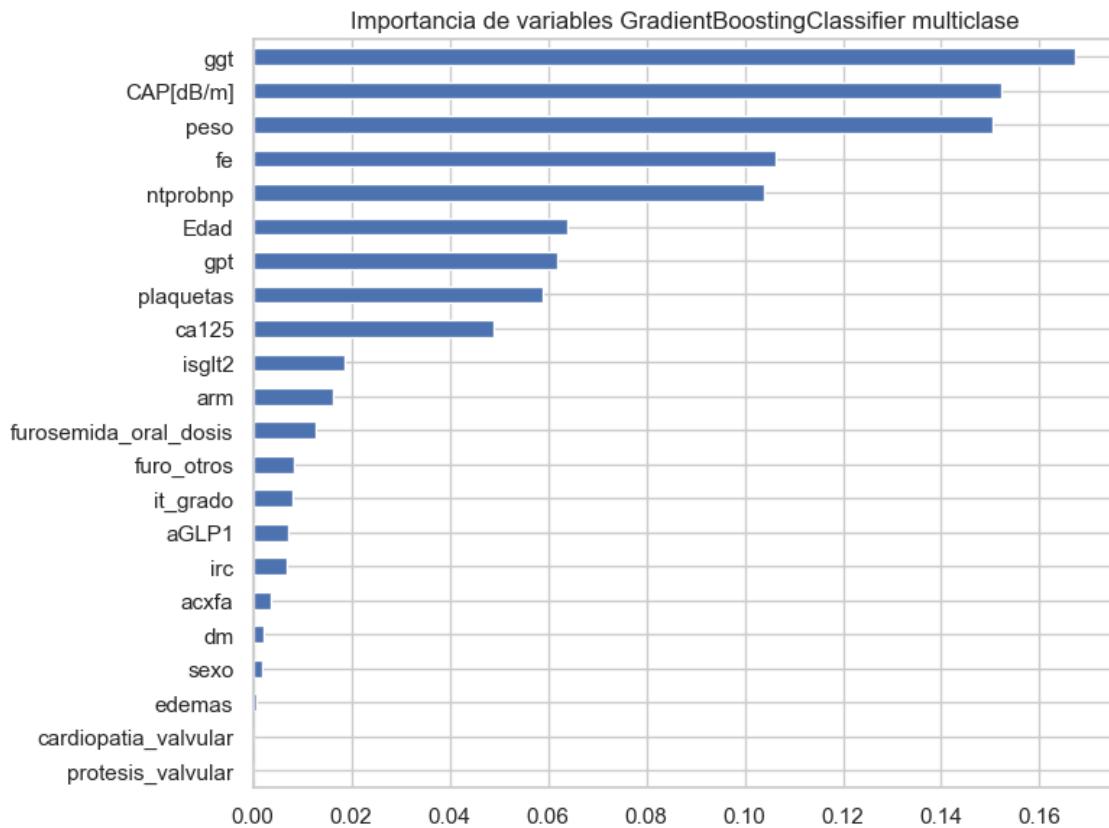
Nota. Tomado de código fuente, 2025.

Importancia de Variables Multiclas

La gráfica de barras de la importancia de las características para este modelo.

Figura 15

Importancia de variables modelo multiclas



Nota. Tomado de código fuente, 2025.

6.2.2 Resultados del Modelo de Clasificación Binaria (Principal - Gradient Boosting)

Métricas de Rendimiento (Binario - Gradient Boosting)

Precisión (Accuracy), reporte completo (Precision, Recall, F1-Score).

- Exactitud (Accuracy) = 74%.

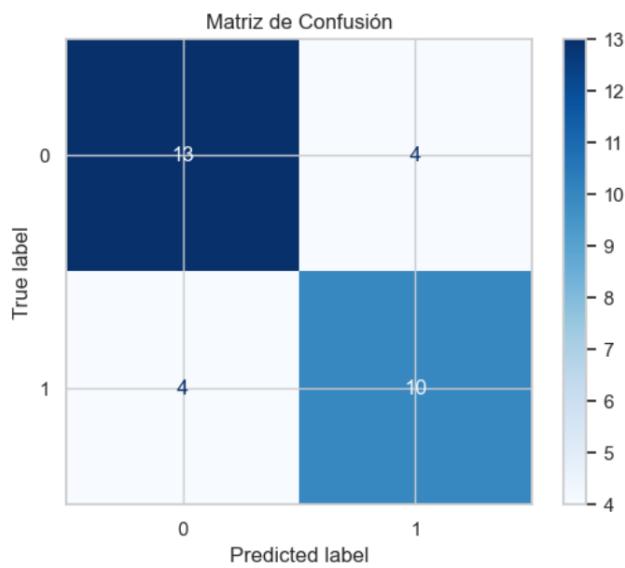
Tabla 11*Métricas de Rendimiento Modelo Binario*

Clase	Precisión (Precision)	Sensibilidad (Recall)	F1-Score	Soporte (Support)
0	0.76	0.76	0.76	17
1	0.71	0.71	0.71	14

Nota. Tomado de código fuente, 2025.

Matriz de Confusión Binario

La matriz de confusión para este modelo, que visualiza los aciertos y errores.

Figura 16*Matriz de Confusión - Presencia de Fibrosis*

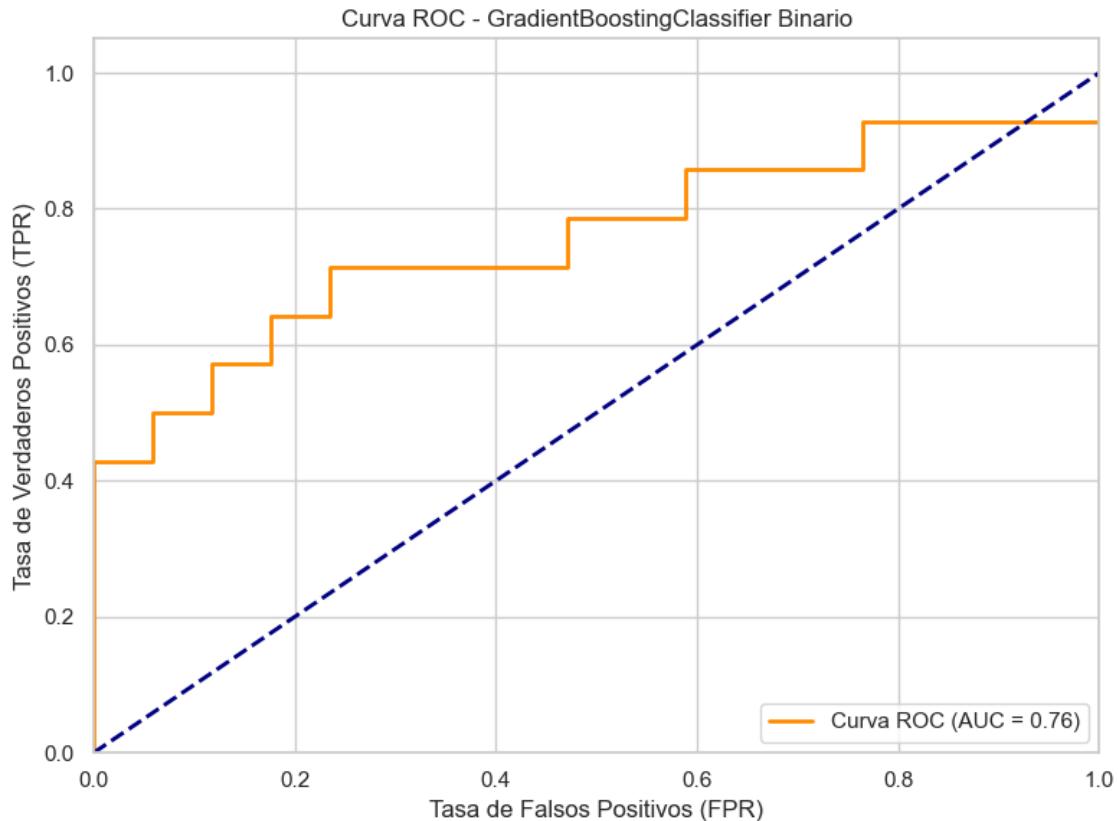
Nota. Tomado de código fuente, 2025.

Curvas ROC y AUC Modelo Binario

La gráfica de las curvas ROC modelo binario con el valor de AUC.

Figura 17

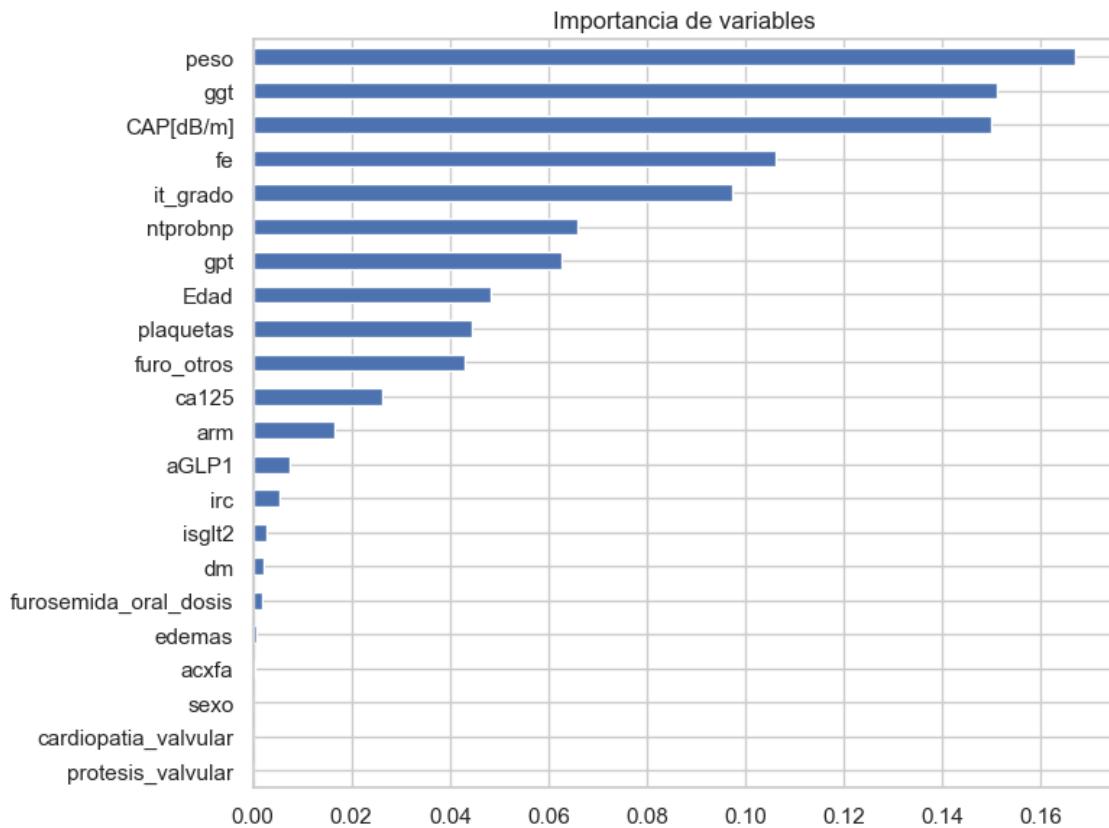
Curvas Roc y AUC Modelo Binario



Nota. Tomado de código fuente, 2025.

Importancia de Variables Modelo Binario

La gráfica de barras de la importancia de las características para este modelo.

Figura 18*Importancia de variables modelo binario*

Nota. Tomado de código fuente, 2025.

6.2.3. Resultados del Modelo de Refinamiento (Regresión Logística)

Métricas de Rendimiento (Regresión Logística)

Precisión, Sensibilidad, Especificidad, ROC-AUC.

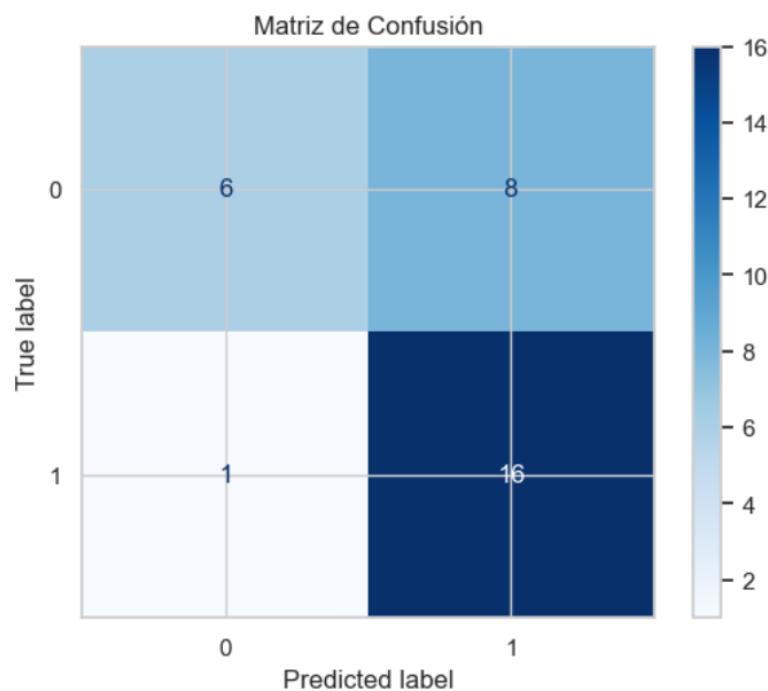
- Exactitud (Accuracy) = 71%.

Tabla 12*Métricas de Rendimiento Modelo Regresión Logística*

Clase	Precisión (Precision)	Sensibilidad (Recall)	F1-Score	Soporte (Support)
0	0.86	0.43	0.57	14
1	0.67	0.94	0.78	17

Nota. Tomado de código fuente, 2025.**Matriz de Confusión Regresión Logística**

La matriz de confusión para este modelo, que visualiza los aciertos y errores.

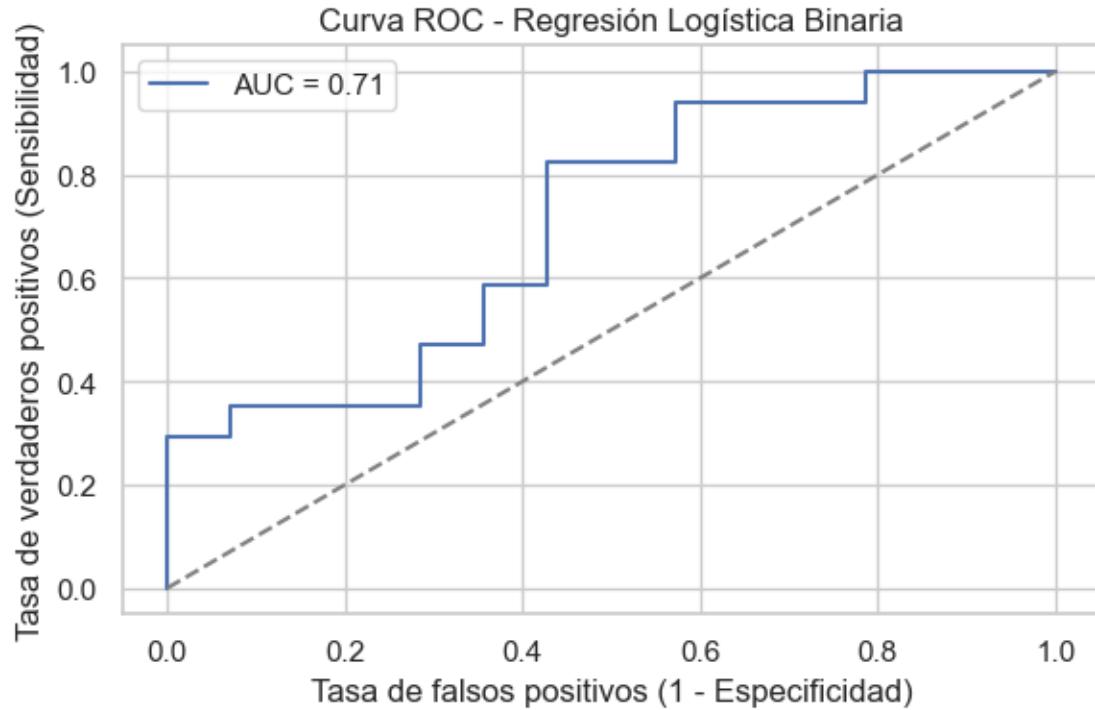
Figura 19*Matriz de Confusión - Presencia de Fibrosis Regresión Logística**Nota.* Tomado de código fuente, 2025.

Curvas ROC y AUC Regresión Logística

La gráfica de las curvas ROC modelo de Regresión Logística con el valor de AUC.

Figura 20

Curvas Roc y AUC Regresión Logística



Nota. Tomado de código fuente, 2025.

7. Discusión

La presente sección tiene como objetivo analizar en profundidad los resultados obtenidos del modelado predictivo de la fibrosis hepática en pacientes con insuficiencia cardíaca, contextualizándolos dentro del marco de la investigación previa y discutiendo sus implicaciones clínicas y teóricas, así como las limitaciones y futuras líneas de investigación.

Análisis del Desempeño del Modelo Multiclasificación y Justificación del Enfoque Binario

El estudio inició con una aproximación de clasificación multiclasificación para predecir los distintos grados de fibrosis hepática (F0 a F4) utilizando un modelo de Gradient Boosting. Aunque la precisión global alcanzó un 61%, el análisis detallado de las métricas de desempeño y la matriz de confusión revelaron limitaciones significativas. Específicamente, el modelo mostró una capacidad discriminativa deficiente en las clases intermedias (F1, F2, F3), con valores de recall de 0, 0.33 y 0 respectivamente, mientras que las clases extremas (F0 y F4) presentaron un rendimiento más aceptable, con AUCs de 0.76 para F0 y hasta 0.89 para F4 (y 0.86 para F2, aunque el recall es bajo para esta). Esta disparidad evidenció la dificultad del modelo para diferenciar entre estadios que clínicamente pueden tener superposiciones sutiles, así como el impacto del reducido número de muestras en algunas clases (clase 0=54, clase 1=4, clase 2=10, clase 3=11, clase 4=22). La matriz de confusión confirmó que el modelo aprendió a identificar bien la clase con mayor número de muestras (F0), mientras que su desempeño en otras clases fue equiparable a una predicción aleatoria.

Esta observación crítica justificó la transición metodológica hacia una clasificación binaria. Se argumentó que, si bien una predicción granular de los grados es deseable, la capacidad de detectar la presencia o ausencia de fibrosis es una pregunta clínicamente más

urgente y factible con los datos disponibles. Esta simplificación del problema busca ofrecer una herramienta de cribado no invasivo que priorice la identificación de pacientes que requieren mayor atención.

Rendimiento del Modelo de Clasificación Binaria (Gradient Boosting)

La reformulación del problema a una clasificación binaria (ausencia de fibrosis F0 vs. presencia de fibrosis F1-F4) mediante un modelo de Gradient Boosting Classifier, con ajustes de parámetros y ponderación de clases, resultó en una mejora sustancial del rendimiento y la aplicabilidad. El modelo alcanzó una accuracy del 0.74, con un recall del 0.71 para la clase de "presencia de fibrosis" (clase 1) y un recall del 0.76 para la clase de "ausencia de fibrosis" (clase 0). Los F1-scores de 0.76 y 0.71 respectivamente, y un AUC de 0.76, indican que el modelo es robusto y ofrece un rendimiento aceptable para la tarea de detección.

La matriz de confusión del modelo binario mostró un balance razonable, con 13 aciertos y 4 errores para la clase 0 (ausencia de fibrosis), y 10 aciertos y 4 errores para la clase 1 (presencia de fibrosis). Este rendimiento, aunque no es perfecto, es prometedor para su uso como herramienta de apoyo en la rápida observación de pacientes. La capacidad de detectar la fibrosis de manera automática, especialmente considerando las dificultades que algunos pacientes pueden experimentar con pruebas como FibroScan, resalta la ventaja clínica de este enfoque binario.

Importancia de las Características

El análisis de la importancia de variables en el modelo de Gradient Boosting binario reveló que el modelo se basaba en factores clínicamente relevantes. Destacaron variables relacionadas con el daño hepático (GGT, GPT), la esteatosis hepática medida por CAP, y otros factores antropométricos como el peso y la edad. Además, parámetros indicadores de congestión

y comorbilidades cardíacas, como la insuficiencia tricúspide, la fracción de eyección, NT-proBNP y CA125, así como el uso combinado de diuréticos, mostraron una alta relevancia. Este hallazgo se correlaciona directamente con la realidad clínica y los valores considerados en la unidad de insuficiencia cardíaca, lo que refuerza la validez interna del modelo. La relevancia de CAP también subraya la importancia de FibroScan en la detección de fibrosis, aunque en el modelo predictivo se busca prescindir de esta prueba en la fase de entrada de datos.

Rendimiento y Utilidad de la Regresión Logística con Variables Seleccionadas

Posteriormente, se construyó un modelo de Regresión Logística binaria utilizando las 10 variables más importantes identificadas por el Gradient Boosting, buscando una mayor simplicidad y explicabilidad. Los resultados obtenidos fueron una accuracy del 0.71 y un AUC del 0.71, lo cual es comparable al modelo de Gradient Boosting principal, validando la relevancia predictiva de las características seleccionadas. Cabe destacar un recall de 0.94 para la clase de "presencia de fibrosis" (clase 1) y una precision de 0.86 para la clase de "ausencia de fibrosis" (clase 0).

Este modelo de Regresión Logística, al priorizar la detección de la fibrosis con un alto recall (reduciendo significativamente los falsos negativos), se alinea con la necesidad clínica de localizar pacientes con fibrosis, incluso si esto implica un aumento de falsos positivos. Aunque un exceso de falsos positivos no es óptimo en la práctica diaria, este enfoque es valioso en un contexto de cribado, donde el sacrificio de especificidad por una mayor sensibilidad es a menudo preferible para no pasar por alto casos de enfermedad. La principal ventaja metodológica de la regresión logística radica en su interpretabilidad, ya que permite inferir la dirección y magnitud del efecto de cada variable, lo que es fundamental en contextos clínicos que requieren transparencia y justificación de las decisiones. Este modelo, por tanto, es más adecuado para su

potencial integración en protocolos clínicos, facilitando la comprensión por parte de los profesionales sanitarios.

Implicaciones Clínicas y Teóricas

Los hallazgos de este estudio tienen implicaciones significativas para la práctica clínica y el avance del conocimiento en la predicción no invasiva de la fibrosis hepática en pacientes con insuficiencia cardíaca. El modelo binario propuesto, basado en variables clínicas accesibles, demuestra el potencial de ser una herramienta de apoyo para el cribado inicial, identificando pacientes con alta probabilidad de fibrosis que podrían beneficiarse de una evaluación más exhaustiva. Esto es particularmente relevante dado que la fibrosis hepática puede ser asintomática en sus etapas iniciales y la prueba de FibroScan no siempre es factible.

Desde una perspectiva teórica, el estudio contribuye al cuerpo de evidencia que respalda la utilidad de algoritmos de aprendizaje automático para identificar patrones complejos en datos médicos, y valida la relevancia de un conjunto específico de variables clínicas en la predicción de la fibrosis en una población específica como la de insuficiencia cardíaca.

Limitaciones del Estudio y Futuras Líneas de Investigación

Es importante reconocer las limitaciones del presente estudio. El tamaño de la muestra actual (101 pacientes) es relativamente reducido, lo que puede afectar la generalizabilidad de los modelos. Un conjunto de datos más amplio y de múltiples centros sería crucial para validar y mejorar la robustez de los modelos, y abordar el desequilibrio de clases de manera más efectiva. Además, la definición de los grados de fibrosis se basó en mediciones de FibroScan, siendo la biopsia hepática el estándar de oro para el diagnóstico definitivo, lo que introduce una limitación inherente en la "verdad fundamental" del diagnóstico.

Basado en estas consideraciones, las futuras líneas de investigación incluyen:

- Aumento del Tamaño Muestral: La integración de más muestras, como se plantea con el avance de la investigación, será fundamental para aumentar la fiabilidad y capacidad predictiva de los modelos.
- Validación Externa: Realizar validaciones con conjuntos de datos independientes para evaluar la transferibilidad y robustez de los modelos en diferentes poblaciones.
- Optimización del Umbral: Explorar metodologías más avanzadas para la optimización del umbral de decisión en contextos clínicos donde el costo de los falsos negativos es muy alto.
- Desarrollo de Herramientas Interactivas: Crear una interfaz de usuario o una aplicación que permita a los profesionales sanitarios utilizar el modelo de forma sencilla para la evaluación rápida de pacientes.

8. Conclusiones

Las conclusiones de este trabajo se fundamentan en el riguroso proceso de modelado y análisis de datos, demostrando un progreso significativo en el abordaje de la predicción de fibrosis hepática en pacientes con insuficiencia cardíaca. Se ha validado la capacidad de los modelos desarrollados para cumplir con los objetivos generales y específicos propuestos, sentando las bases para futuras mejoras e implementaciones.

Cumplimiento del Objetivo General

El Objetivo General de "**Desarrollar modelos de inteligencia artificial capaces de predecir la presencia de fibrosis hepática en pacientes con insuficiencia cardíaca, utilizando variables clínicas y el FibroScan como referencia diagnóstica**" se ha logrado satisfactoriamente. A través de la aplicación de algoritmos de aprendizaje automático supervisado, se construyeron modelos predictivos que, si bien utilizan la medida de FibroScan ($E[kPa]$) como el estándar de oro para definir la condición de fibrosis (grados o presencia/ausencia), operan con variables clínicas como predictores de entrada. Esto es crucial, ya que permite que los modelos infieran la presencia de fibrosis sin requerir la realización del FibroScan en la fase de predicción, tal como se buscaba en la investigación.

Cumplimiento de los Objetivos Específicos

"Aplicar algoritmos de aprendizaje supervisado para construir modelos de predicción binaria (fibrosis: sí/no)."

Este objetivo se ha cumplido al desarrollar exitosamente un modelo de Gradient Boosting Classifier para la clasificación binaria de fibrosis (presencia/ausencia). A pesar de los desafíos inherentes al desbalance de clases iniciales ($F0=54$, $F1=4$, $F2=10$, $F3=11$, $F4=22$), el modelo

final binario demostró una accuracy del 0.74 y un recall del 0.71 para la clase de "presencia de fibrosis", con un AUC de 0.76. Aunque el rendimiento es aún mejorable y la investigación se encuentra en curso con el trabajo de extracción de más datos, estos resultados establecen una base robusta y prometedora para la detección no invasiva. La posterior aplicación de un modelo de Regresión Logística (con accuracy 0.71 y AUC 0.71) y un recall de 0.94 para la clase de "presencia de fibrosis", refuerza la consecución de este objetivo y valida la capacidad de los algoritmos para esta tarea.

"Explorar el valor de la inteligencia artificial como herramienta complementaria en escenarios donde el acceso al FibroScan sea limitado."

Este objetivo ha sido directamente abordado y su valor se ha puesto de manifiesto. Los modelos desarrollados utilizan únicamente variables clínicas accesibles, lo que implica el potencial desarrollo de una herramienta capaz de complementar o incluso, en el futuro, suplir la necesidad inmediata de un FibroScan en contextos donde su realización sea difícil o no se cuente con el equipo. Esto representa un avance significativo hacia la democratización del cribado de fibrosis, permitiendo una rápida identificación de pacientes en riesgo y facilitando decisiones terapéuticas tempranas sin depender de una tecnología específica.

"Identificar las variables clínicas con mayor influencia en la predicción de fibrosis hepática mediante análisis de importancia de características."

La investigación ha identificado con éxito variables de relevancia en la predicción de fibrosis. El análisis de importancia de características del modelo de Gradient Boosting binario destacó factores como el peso, indicadores de daño hepático (GGT, GPT), la esteatosis hepática (medida con CAP, que refuerza la importancia del FibroScan como referencia), comorbilidades

cardíacas (insuficiencia tricuspídea, fracción de eyección), indicadores de congestión (NT-proBNP, CA125), y el uso de diuréticos. Estas variables no solo concuerdan con la realidad clínica y la fisiopatología conocida de la fibrosis en el contexto de la insuficiencia cardíaca, sino que también son parámetros en los cuales se hace hincapié en la evaluación de pacientes en las unidades de insuficiencia cardíaca. Esta identificación robustece el conocimiento sobre los factores predictores y orienta futuras investigaciones y la atención clínica.

En síntesis, este TFM ha demostrado la viabilidad y el valor de la inteligencia artificial para predecir la fibrosis hepática a partir de datos clínicos, proporcionando una herramienta complementaria para contextos con limitaciones de acceso a tecnologías avanzadas y destacando las variables más influyentes en este proceso. La naturaleza continua de la investigación y la integración de más datos prometen una mejora sostenida en la robustez y aplicabilidad de los modelos.

9. Bibliografía

- Aplicaciones de la inteligencia artificial en cardiología: El futuro ya está aquí. (2019). *Revista Española de Cardiología*, 72(12), 1065–1075.
<https://doi.org/10.1016/j.recesp.2019.05.016>
- Bartres, C., & Lens, S. (2013). Elastografía hepática (Fibroscan®) en hepatología. *Revista Española de Enfermedades Digestivas*, 105(4), 235–235.
<https://doi.org/10.4321/S1130-01082013000400011>
- Brankovic, M., Lee, P., Pyrsopoulos, N., & Klapholz, M. (2023). Cardiac Syndromes in Liver Disease: A Clinical Conundrum. *Journal of Clinical and Translational Hepatology*, 11(4), 975–986. <https://doi.org/10.14218/JCTH.2022.00294>
- EASL-ALEH Clinical Practice Guidelines: Non-invasive tests for evaluation of liver disease severity and prognosis. (2015). *Journal of Hepatology*, 63(1), 237–264.
<https://doi.org/10.1016/j.jhep.2015.04.006>
- *Elastografía de transición FibroScan*. (n.d.). Hospital Universitario Quirónsalud Madrid. Retrieved May 29, 2025, from https://www.quironsalud.com/hospital-madrid/es/elastografia_de_transicion
- Fouad, Y. M., & Yehia, R. (2014). Hepato-cardiac disorders. *World Journal of Hepatology*, 6(1), 41–54. <https://doi.org/10.4254/wjh.v6.i1.41>
- Giallourakis, C. C., Rosenberg, P. M., & Friedman, L. S. (2002). The liver in heart failure. *Clinics in Liver Disease*, 6(4), 947–967. [https://doi.org/10.1016/S1089-3261\(02\)00056-9](https://doi.org/10.1016/S1089-3261(02)00056-9)
- Herrero, J. I., Iñarraiaegui, M., D'Avola, D., Sangro, B., Prieto, J., & Quiroga, J. (2014). Comparación de las sondas M y XL para estimar la rigidez hepática por medio de

elastografía de transición. *Gastroenterología y Hepatología*, 37(4), 233–239.

<https://doi.org/10.1016/j.gastrohep.2013.10.009>

- Lanzagorta-Ortega, D., Carrillo-Pérez, D. L., Carrillo-Esper, R., Lanzagorta-Ortega, D., Carrillo-Pérez, D. L., & Carrillo-Esper, R. (2022). Inteligencia artificial en medicina: Presente y futuro. *Gaceta médica de México*, 158, 17–21.

<https://doi.org/10.24875/gmm.m22000688>

- Llacer, P., Gallardo, M. Á., Palau, P., Moreno, M. C., Castillo, C., Fernández, C., de la Espriella, R., Mollar, A., Santas, E., Miñana, G., Manzano, L., Bayés-Genís, A., & Núñez, J. (2021). Comparación entre CA125 y NT-proBNP para valorar la congestión en insuficiencia cardíaca aguda. *Medicina Clínica*, 156(12), 589–594.

<https://doi.org/10.1016/j.medcli.2020.05.063>

- Patel, K., & Wilder, J. (2014). Fibroscan. *Clinical Liver Disease*, 4(5), 97–101.
- <https://doi.org/10.1002/cld.407>
- Raraz-Vidal, J., & Raraz-Vidal, O. (2022). Aplicaciones de la inteligencia artificial en la medicina. *Revista Peruana de Investigación en Salud*, 6(3), 131–133.
- <https://doi.org/10.35839/repis.6.3.1559>
- Xanthopoulos, A, Starling, R, Kitai, T. et al. Insuficiencia cardíaca y enfermedad hepática: interacciones cardiohepáticas. J Am Coll Cardiol HF. 2019 feb, 7 (2) 87–97.

<https://doi.org/10.1016/j.jchf.2018.10.007>

- de Lédinghen, V., & Vergniol, J. (2008). Transient elastography (FibroScan). *Gastroentérologie Clinique et Biologique*, 32(6, Suppl. 1), 58–67.

[https://doi.org/10.1016/S0399-8320\(08\)73994-0](https://doi.org/10.1016/S0399-8320(08)73994-0)

- Baialardo¹, Antonela Guadalupe. (2015). Síndrome cardio-hepático ¿Qué debemos saber como cardiólogos?. *Insuficiencia cardíaca*, 10(2), 66-77. Recuperado en 24 de junio de 2025, de https://www.scielo.org.ar/scielo.php?script=sci_arttext&pid=S1852-38622015000200003&lng=es&tlang=es.
- Fortea, JI, Puente, Á., Cuadrado, A., Huelin, P., Pellón, R., González Sánchez, FJ, Mayorga, M., Cagigal, ML, García Carrera, I., Cobreros, M., Crespo, J., & Fábrega, E. (2020). Hepatopatía congestiva. *Revista Internacional de Ciencias Moleculares* , 21 (24), 9420. <https://doi.org/10.3390/ijms21249420>
- Hilscher, Moira MD 1 ; Sánchez, William MD *,2 .Hepatopatia congestiva. Enfermedad Hepática Clínica 8(3):p 68-71, septiembre de 2016. | DOI: 10.1002/cld.573
- Ingrassia, V. (2018, mayo 25). Las nuevas válvulas cardíacas que se implantan sin operar a corazón abierto. *Infobae*. <https://www.infobae.com/salud/2018/05/25/las-nuevas-valvulas-cardiacas-que-se-implantan-sin-operar-a-corazon-abierto/>
- Lab Tests Online. (2021). *Gamma-glutamil transferasa (GGT)*. <https://www.labtestsonline.es/tests/gamma-gt-f>
- Bexaray. (2018, noviembre 29). *Información sobre sus resultados de FibroScan*. <https://www.bexaray.com/single-post/2018/11/29/informaci%C3%B3n-sobre-sus-resulstados-de-fibroscan>
- Echosens. (2018). *FibroScan® Mini+ 430* [Dispositivo médico]. <https://www.echosens.com/es/products/fibroscan-mini-2/>
- Echosens. (s.f.). *Sondas FibroScan®* [Imagen de dispositivo médico]. <https://www.echosens.com/es/products/sondas/>

- Alpharetta Internal Medicine. (s.f.). *Live health with FibroScan*.
<https://alpharettainternalmedicine.com/live-health-fibroscan/>
- Mora Pineda, J. (2022). Modelos predictivos en salud basados en aprendizaje de máquina (machine learning). *Revista Médica Clínica Las Condes, 33*(6), 583–590.
<https://doi.org/10.1016/j.rmclc.2022.11.002>
- van Smeden, M., Reitsma, J. B., Riley, R. D., Collins, G. S., & Moons, K. G. M. (2021). Clinical prediction models: Diagnosis versus prognosis. *Journal of Clinical Epidemiology, 132*, 142–145. <https://doi.org/10.1016/j.jclinepi.2021.01.009>
- Fan, G., Deng, Z., Ye, Q., & Wang, B. (2021). *Machine learning-based prediction models for patients no-show in online outpatient appointments*. Data Science and Management, 2, 45–52. <https://doi.org/10.1016/j.dsm.2021.06.002>
- Elearning Actual. (2022). *4 tipos de tareas de clasificación en el machine learning*.
<https://elearningactual.com/4-tipos-de-tareas-de-clasificacion-en-el-machine-learning/>
- DataCamp. (2024). *Clasificación en Machine Learning: Una guía completa para principiantes*. https://www.datacamp.com/es/blog/classification-machine-learning?dc_referrer=https%3A%2F%2Fwww.google.com%2F
- Ciencia de Datos. (2017). *Árboles de decisión, Random Forest, Gradient Boosting y C5.0*.
https://cienciadedatos.net/documentos/33_arboles_decision_random_forest_gradient_boosting_c50
- Wong, B. (2023). *Decision Tree, Random Forest, and XGBoost: An exploration into the heart of machine learning*. Medium. <https://medium.com/@brandon93.w/decision-tree-random-forest-and-xgboost-an-exploration-into-the-heart-of-machine-learning-5a2a2a2a2a2a>

[random-forest-and-xgboost-an-exploration-into-the-heart-of-machine-learning-90dc212f4948](#)

10. Código Fuente y Repositorio GitHub

El código fuente para la implementación de los modelos y análisis de este Trabajo de Fin de Máster ha sido subido a la plataforma GitHub. Este repositorio contiene el desarrollo completo, desde el preprocesamiento de datos hasta la evaluación de los modelos de Machine Learning.

Código fuente disponible en el siguiente enlace:

- <https://github.com/luisstuardo/Prediccion-Fibrosis-Hepatica>