

# Universidad de Salamanca



VNIVERSIDAD  
D SALAMANCA



## TESIS DOCTORAL

**Contribuciones de las técnicas machine learning a la cardiología.  
Predicción de reestenosis tras implante de stent coronario**

**Autor:** Pedro Ignacio Dorado Díaz

**Directores:** Purificación Galindo Villardón  
Pedro Luis Sánchez Fernández

**Salamanca, 2019**



## Resumen

**Antecedentes:** Existen pocos temas de actualidad equiparables a la posibilidad de la tecnología actual para desarrollar las mismas capacidades que el ser humano, incluso en medicina. Esta capacidad de simular los procesos de inteligencia humana por parte de máquinas o sistemas informáticos es lo que conocemos hoy en día como inteligencia artificial. Uno de los campos de la inteligencia artificial con mayor aplicación a día de hoy en medicina es el de la predicción, recomendación o diagnóstico, donde se aplican las técnicas *machine learning*. Asimismo, existe un creciente interés en las técnicas de medicina de precisión, donde las técnicas *machine learning* pueden ofrecer atención médica individualizada a cada paciente.

El intervencionismo coronario percutáneo (ICP) con *stent* se ha convertido en una práctica habitual en la revascularización de los vasos coronarios con enfermedad aterosclerótica obstructiva significativa. El ICP es asimismo patrón oro de tratamiento en pacientes con infarto agudo de miocardio; reduciendo las tasas de muerte e isquemia recurrente en comparación con el tratamiento médico. El éxito a largo plazo del procedimiento está limitado por la reestenosis del *stent*, un proceso patológico que provoca un estrechamiento arterial recurrente en el sitio de la ICP. Identificar qué pacientes harán reestenosis es un desafío clínico importante; ya que puede manifestarse como un nuevo infarto agudo de miocardio o forzar una nueva revascularización del vaso afectado, y que en casos de reestenosis recurrente representa un reto terapéutico.

**Objetivos:** Después de realizar una revisión de las técnicas de inteligencia artificial aplicadas a la medicina y con mayor profundidad, de las técnicas *machine learning* aplicadas a la cardiología, el objetivo principal de esta tesis doctoral ha sido desarrollar un modelo *machine learning* para predecir la aparición de reestenosis en pacientes con infarto agudo de miocardio sometidos a ICP con implante de un *stent*. Asimismo, han sido objetivos secundarios comparar el modelo desarrollado con *machine learning* con los *scores* clásicos de riesgo de reestenosis utilizados hasta la fecha; y desarrollar un *software* que permita trasladar esta contribución a la práctica clínica diaria de forma sencilla. Para desarrollar un modelo fácilmente aplicable, realizamos nuestras predicciones sin variables adicionales a las obtenidas en la práctica rutinaria.

**Material:** El conjunto de datos, obtenido del ensayo GRACIA-3, consistió en 263 pacientes con características demográficas, clínicas y angiográficas; 23 de ellos presentaron reestenosis a los 12 meses después de la implantación del *stent*. Todos los desarrollos llevados a cabo se han hecho en Python y se ha utilizado computación en la nube, en concreto AWS (Amazon Web Services).

**Metodología:** Se ha utilizado una metodología para trabajar con conjuntos de datos pequeños y no balanceados, siendo importante el esquema de validación cruzada anidada utilizado, así como la utilización de las curvas PR (*precision-recall*, exhaustividad-sensibilidad), además de las curvas ROC, para la interpretación de los modelos. Se han entrenado los algoritmos más habituales en la literatura para elegir el que mejor comportamiento ha presentado.

**Resultados:** El modelo con mejores resultados ha sido el desarrollado con un clasificador *extremely randomized trees*; que superó significativamente (0,77; área bajo la curva ROC a los tres scores clínicos clásicos; PRESTO-1 (0,58), PRESTO-2 (0,58) y TLR (0,62). Las curvas exhaustividad sensibilidad ofrecieron una imagen más precisa del rendimiento del modelo *extremely randomized trees* que muestra un algoritmo eficiente (0,96) para no reestenosis, con alta exhaustividad y alta sensibilidad. Para un umbral considerado óptimo, de 1,000 pacientes sometidos a implante de *stent*, nuestro modelo *machine learning* predeciría correctamente 181 (18%) más casos en comparación con el mejor *score* de riesgo clásico (TLR). Las variables más importantes clasificadas según su contribución a las predicciones fueron diabetes, enfermedad coronaria en 2 ó más vasos, flujo TIMI post-ICP, plaquetas anormales, trombo post-ICP y colesterol anormal. Finalmente, se ha desarrollado una calculadora para trasladar el modelo a la práctica clínica. La calculadora permite estimar el riesgo individual de cada paciente y situarlo en una zona de riesgo, facilitando la toma de decisión al médico en cuanto al seguimiento adecuado para el mismo.

**Conclusiones:** Aplicado inmediatamente después de la implantación del *stent*, un modelo *machine learning* diferencia mejor a aquellos pacientes que presentarán o no reestenosis respecto a los discriminadores clásicos actuales.

**Palabras clave:**

Inteligencia Artificial. *Machine learning*. *Deep learning*. Cardiología. Intervencionismo coronario percutáneo (ICP). Reestenosis. Infarto de miocardio. Curva ROC. Curva exhaustividad sensibilidad. *Extremely randomized trees*

## Summary

**Background:** There is currently no other hot topic comparable to the capacity of the current technology to develop similar capabilities as human beings, even in medicine. This ability to simulate the processes of human intelligence with computer systems is known as artificial intelligence. One of the fields of artificial intelligence with greater application today in medicine is that of prediction, recommendation or diagnosis, where machine learning techniques are applied. Furthermore, there is growing interest in precision medicine techniques as machine learning, which can deliver individually adapted medical care.

Percutaneous coronary interventions (PCI) with stent implantation has become routine practice for the revascularization of coronary vessels with significant obstructive atherosclerotic disease. Furthermore, PCI is the gold-standard of treatment in patients with acute myocardial infarction; reducing the rates of death and recurrent ischemia as compared to medical treatment. The long-term success of this procedure may be limited by stent restenosis, a pathological process leading to recurrent arterial narrowing at the site of PCI. Identifying which patients will do restenosis is an important clinical challenge; since it can manifest itself as a new acute myocardial infarction or force a new revascularization of the affected vessel, and that in cases of recurrent restenosis represents an important therapeutic dilemma.

**Objectives:** Starting with a review of artificial intelligence techniques applied to medicine and in greater depth, of machine learning techniques applied to cardiology, the main objective of this doctoral thesis has been to develop a machine learning model to predict restenosis in patients with acute myocardial infarction undergoing PCI with stent implantation. In addition, secondary objectives have been to compare the developed machine learning model with classical predictive clinical scores, and to develop software able to transfer this machine learning contribution to daily clinical practice. In order to develop an easily-applicable model, we performed our predictions without any additional variables than those obtained in daily practice.

**Material:** The dataset, obtained from the GRACIA-3 trial, consisted of 263 patients with demographic, clinical and angiographic characteristics; 23 of them presented restenosis at 12-months after stent implantation. All software development has been made in Python and cloud computing has been used, specifically AWS (Amazon Web Services).

**Results:** Our best performing model was developed with an extremely randomized trees classifier; which significantly outperformed (0.77; area under the ROC curve) three clinical scores; PRESTO-1 (0.58), PRESTO-2 (0.58), and TLR (0.62). Precision-

recall curves gave a more accurate picture of the extremely randomized trees model's performance showing an efficient algorithm (0.96) for non-restenosis, returning high precision as well as high recall. For a threshold considered as optimal, out of 1,000 patients undergoing stent implantation, our machine learning model would correctly predict 181 (18%) more cases compared to the best clinical score (TLR). The most important variables ranked according to their contribution to the predictions were diabetes,  $\geq 2$  vessel-coronary disease, post-PCI TIMI-flow, abnormal platelets, post-PCI thrombus, and abnormal cholesterol. Finally, a calculator has been developed to transfer the model to clinical practice. The calculator allows estimating the individual risk of each patient and placing it in a risk zone, facilitating the decision of the doctor regarding the appropriate follow-up.

**Conclusions:** Applied immediately after stent implantation, a machine learning model better differentiates those patients who will present or not restenosis over current classical discriminators.

**Key words:**

Artificial intelligence. Machine learning. Deep learning. Cardiology. Percutaneous Coronary Intervention (PCI). Stent Restenosis. Myocardial Infarction. ROC curve. Precision-recall (PR) curve. Extremely randomized trees

## **Fuentes de financiación**

El presente trabajo ha sido financiado por el Instituto de Salud Carlos III (Ministerio de Ciencia, Innovación y Universidades) dentro del CIBERCV, por el que está contratado Pedro Ignacio Dorado Díaz.

## **Agradecimientos**

**A** Mili, mi pareja, por su constante apoyo incondicional. Siempre juntos a pesar de las dificultades. Mili y Acho, Acho y Mili.

A mis hijos, María y Nacho, por todo el tiempo que no les he podido dedicar y que estoy intentando recuperar.

A mis padres. Esta tesis ha necesitado de un último empujón de mi padre. Papá, que la vieras acabada ha sido uno de los grandes estímulos que he tenido para conseguirlo. Gracias Mamá por tu esfuerzo y entrega a la familia.

A Purificación Galindo, por sus buenos consejos y ánimos. Un gran valor para esta Universidad. Sin tu ayuda no habría podido seguir adelante con este proyecto.

A Pedro Luis Sánchez, por su apoyo, su orientación y sus ánimos. Tutor, jefe y sobre todo amigo. Gracias amigo por estar a mi lado en momentos tan complicados y apoyarme con tanta energía.

A Jesús Manuel Sampedro, Víctor Vicente y Antonio Sánchez, gracias por vuestras aportaciones y vuestra paciencia. Una suerte haber compartido este proceso con personas como vosotros.

A mi hermana Ana, por su ayuda, por sus ganas de ayudarme.



A Mili  
A Nacho y María  
A mis padres

**El momento es ahora**

# ÍNDICE

Glosario de acrónimos .....	15
Índice de figuras.....	17
Índice de tablas .....	18
1. Introducción.....	19
1.1. Ciencia de datos.....	19
1.2. Inteligencia artificial.....	24
1.2.1. Reconocimiento automático de voz y procesamiento de lenguaje natural .....	24
1.2.2. Algoritmos de predicción, recomendación o diagnóstico .....	25
1.2.3. Visión artificial y análisis de imagen .....	25
1.2.4. Robótica e inteligencia artificial.....	25
1.2.5. Sistemas expertos .....	25
1.3. <i>Machine learning</i> : técnicas .....	26
1.4. <i>Machine learning</i> : metodología.....	26
1.5. <i>Deep learning</i> y la imagen médica .....	31
1.6. Limitaciones del <i>machine learning</i> y del <i>deep learning</i> .....	33
1.7. <i>Machine learning</i> y la estadística clásica: similitudes y diferencias .....	34
1.8. Aspectos legales y éticos de la inteligencia artificial .....	34
2. Hipótesis y objetivos .....	37
2.1. Hipótesis.....	37
2.2. Objetivos .....	37
2.2.1. Objetivo Principal.....	37
2.2.2. Objetivos Secundarios.....	38
3. Material.....	40
3.1. Datos .....	40
3.2. Software utilizado .....	44
3.3. Hardware utilizado.....	45
4. Metodología.....	46
4.1. Preparación para construir el modelo <i>machine learning</i> .....	48
4.1.1. Definición del objetivo del modelo .....	48
4.1.2. Sujetos del estudio.....	48

---

4.1.3.	Variables de entrada .....	48
4.1.4.	Variable resultado .....	48
4.1.5.	Preparación de datos .....	49
4.2.	Entrenando el modelo <i>machine learning</i> .....	49
4.2.1.	Modelos <i>machine learning</i> y técnicas de selección de variables.....	51
4.2.2.	Esquema de validación cruzada ( <i>cross-validation</i> ) y ajuste de hiperparámetros ( <i>hyperparameter tuning</i> ) .....	55
4.3.	Evaluando el modelo <i>machine learning</i> .....	58
4.3.1.	Esquema de evaluación .....	58
4.3.2.	Comparación del modelo <i>machine learning</i> propuesto con los <i>scores</i> clásicos de predicción de riesgo de sufrir reestenosis .....	60
4.3.3.	Métricas utilizadas en la evaluación .....	61
4.4.	Trasladando el modelo <i>machine learning</i> a la práctica clínica. Desarrollo de una calculadora de riesgo de reestenosis coronaria .....	62
4.4.1.	Calculadora con Google Colab .....	63
4.4.2.	Calculadora con Tkinter .....	64
4.4.3.	Representación del paciente en las curvas ROC y PR .....	64
5.	Resultados.....	66
5.1.	Características de la población de estudio .....	66
5.2.	Comparación de los modelos <i>machine learning</i> con los discriminadores predictivos inferenciales desarrollados con modelos estadísticos clásicos .....	73
5.3.	Interpretación de las curvas ROC y PR en diferentes escenarios de aplicación clínica .....	79
5.4.	Variables con mayor peso en la predicción .....	85
5.5.	Aplicando el modelo a la práctica clínica. Propuesta de calculadora de riesgo de reestenosis del <i>stent</i> tras ICP.....	87
5.5.1.	Calculadora con Google Colab .....	87
5.5.2.	Calculadora con Tkinter .....	90
6.	Discusión .....	92
6.1.	Reestenosis coronaria.....	92
6.2.	Utilizando <i>machine learning</i> para predicción de reestenosis coronaria .....	94
6.3.	Conjunto de datos pequeños y no balanceados para entrenar algoritmos <i>machine learning</i> .....	96
6.4.	Los <i>scores</i> de riesgo desarrollados con modelos estadísticos clásicos.....	100
6.5.	Declaración TRIPOD .....	102

---

6.6.	Curvas ROC y PR como medida de evaluación para un sistema clasificador binario .....	103
6.7.	Variables predictoras en modelos <i>machine learning</i> .....	104
6.8.	Integrando los modelos <i>machine learning</i> en la práctica clínica diaria .....	106
6.9.	Inteligencia artificial en Cardiología.....	107
6.9.1.	Arritmias cardiacas.....	109
6.9.2.	Cardiopatía isquémica.....	110
6.9.3.	Insuficiencia cardiaca .....	110
6.9.4.	Imagen cardiaca .....	111
6.9.5.	Otras aplicaciones .....	112
7.	Limitaciones .....	117
8.	Conclusiones .....	119
	Bibliografía .....	122
	Anexos.....	130
	Anexo 1. Artículo publicado. “Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí” .....	131
	Anexo 2. Artículo enviado. “Machine Learning to Predict Stent Reestenosis Based on Daily Demographic, Clinical and Angiographic Characteristics” .....	143
	Anexo 3. Código fuente desarrollado (Python) para la construcción del modelo <i>machine learning</i> de predicción de reestenosis.....	169
	Anexo 4. Código fuente calculadoras desarrolladas en Python para predecir reestenosis del stent tras ICP .....	202
	Anexo 4.1 Calculadora Google Colab.....	202
	Anexo 4.2 Calculadora Tkinter .....	202
	Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed) .....	210
	Anexo 6. Aprobación CEIm proyecto ANGIO .....	307

# Glosario de acrónimos

**ML:** aprendizaje automático (*Machine Learning*)

**BBDD:** Base de Datos

**ROC:** Característica Operativa del Receptor (*Receiver Operating Characteristic*)

**PR:** Exhaustividad-Sensibilidad (*Precision-Recall*)

**AUC:** Área bajo la curva (*Area Under the Curve*)

**TRIPOD:** *Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis*

**MDA:** Disminución media en exactitud (*Mean Decrease Accuracy*)

**STEMI:** Infarto agudo de miocardio con elevación del ST (*ST Elevation Acute Myocardial Infarction*)

**GUI:** interfaces gráficas de usuario (*Graphical User Interface*)

**AWS:** Amazon Web Services

**GPU:** Unidad de procesamiento gráfico (*Graphics Processing Unit*)

**ICP, PCI:** intervención coronaria percutánea (*Percutaneous Coronary Intervention*)

**QCA:** Análisis coronario cuantitativo (*Quantitative Coronary Angiography*)

**RF:** *Random forest*

**ERT:** *Extremely randomized trees*

**GB:** *Gradient boosting*

**SVC, SVM:** *Support vector machine classifier*

**LR:** *L2-regularized logistic regression*

**LR\_NOREG:** *Non-regularized logistic regression*

**ANOVA:** Análisis de la varianza (*ANalysis Of VAriance*)

**FI:** *Feature importance*

**VP:** Verdadero positivo

**FP:** Falso positivo

**FN:** Falso negativo

**VN:** Verdadero negativo

**VPN:** Valor predictivo negativo

**IECA:** Inhibidores enzima convertidora de la angiotensina

**ARA II:** Antagonistas de los receptores de la angiotensina II

**LAD:** Arteria descendente anterior izquierda (*Left Anterior Descending artery*)

**LCX:** Arteria circunfleja izquierda (*Left Circumflex Artery*)

**RCA:** Arteria coronaria derecha (*Right Coronary Artery*)

**OR:** *Odds-ratio*

**IC:** Intervalo de confianza

# Índice de figuras

<i>Figura 1.</i> Diagrama de Venn de los términos más comúnmente empleados en las disciplinas de la ciencia de datos.....	21
<i>Figura 2A:</i> Búsquedas totales (fuente Google Trends), en los últimos 5 años completos, de términos relacionados con inteligencia artificial y ciencia de datos. ....	22
<i>Figura 2B:</i> Términos más buscados en cada país en el mismo periodo.....	23
<i>Figura 3.</i> Metodología de aplicación de un modelo machine learning. ....	28
<i>Figura 4.</i> Métricas comúnmente empleadas en la evaluación de algoritmos de clasificación supervisada. Entre paréntesis se indican los términos en inglés.....	30
<i>Figura 5.</i> Ejemplo de segmentación e identificación automáticas del ventrículo izquierdo y derecho. ....	32
<i>Figura 6.</i> Flujo de pacientes ensayo GRACIA-3. ....	43
<i>Figura 7.</i> Resumen de las fases que se han seguido para construir el modelo machine learning.....	47
<i>Figura 8.</i> Esquema de validación cruzada anidada seguido para entrenar y validar el modelo. ....	50
<i>Figura 9.</i> Ejemplo de uno de los extremely randomized trees (ERT) utilizado en nuestro modelo para predecir la reestenosis del stent. ....	59
<i>Figura 10.</i> Curvas ROC para los modelos machine learning entrenados.....	74
<i>Figura 11.</i> Curvas ROC comparando el mejor modelo machine learning desarrollado con los scores de riesgo clásicos.....	75
<i>Figura 12.</i> Curvas PR (Precision-Recall) comparando la predicción de reestenosis (precision) y no reestenosis (valor predictivo negativo) de los distintos modelos machine learning entrenados. ....	77
<i>Figura 13.</i> Curvas PR (Precision-Recall) comparando la predicción de reestenosis (precision) y no reestenosis (valor predictivo negativo) del modelo machine learning y los scores de riesgo clásicos. ....	78
<i>Figura 14.</i> Curvas ROC comparando el mejor modelo machine learning desarrollado con los scores de riesgo clásicos en 3 umbrales de uso clínico.....	82
<i>Figura 15.</i> Curvas PR (Precision-Recall) comparando la predicción de reestenosis (precision) y no reestenosis (valor predictivo negativo) del modelo machine learning y los scores de riesgo clásicos en tres umbrales de uso clínico. ....	83
<i>Figura 16.</i> Zonas de riesgo definidas por los umbrales. Paciente de alto riesgo.....	84
<i>Figura 17.</i> Zonas de riesgo definidas por los umbrales. Paciente de bajo riesgo. ....	84
<i>Figura 18.</i> Gráfico de barras con la importancia de las variables. Importancia de cada una de las variables en la predicción de una reestenosis del stent en el seguimiento a 12 meses con un clasificador extremely randomized trees classifier (ERT). ....	86
<i>Figura 19.</i> Calculadora de riesgo de reestenosis tras angioplastia desarrollada en Google Colab (Google Colab Calculadora Reestenosis) .....	87
<i>Figura 20.</i> Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. Entrada de variables. ....	88
<i>Figura 21.</i> Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. Score y métricas calculados por el modelo.....	88
<i>Figura 22.</i> Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. Estratificación del riesgo del paciente basado en curvas ROC.....	89
<i>Figura 23.</i> Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. Estratificación del riesgo del paciente basado en curvas PR para reestenosis y no reestenosis.....	90
<i>Figura 24.</i> Calculadora de riesgo de reestenosis tras ICP desarrollada con Tkinter. ....	90
<i>Figura 25.</i> Calculadora de riesgo de reestenosis tras ICP desarrollada en Tkinter. Estratificación del riesgo del paciente basado en curvas ROC y curvas PR para reestenosis y no reestenosis.....	91
<i>Figura 26.</i> Hitos en la evolución de la ICP <sup>57</sup> . ....	93
<i>Figura 27.</i> Evolución en las publicaciones referenciadas en PubMed sobre inteligencia artificial, machine learning y deep learning según el área de interés en cardiología. ....	108



## Índice de tablas

<i>Tabla 1. Algoritmos más utilizados en machine learning.....</i>	<i>29</i>
<i>Tabla 2. Lista de los hiperparámetros ajustados en los modelos machine learning utilizados para construir nuestro modelo final. ....</i>	<i>57</i>
<i>Tabla 3. Tablas de contingencia con 4 categorías y métricas asociadas. ....</i>	<i>65</i>
<i>Tabla 4. Diferencias en las variables de entrada continuas de la población.....</i>	<i>67</i>
<i>Tabla 5. Diferencias en las variables de entrada categóricas de la población.....</i>	<i>69</i>
<i>Tabla 6. Valores para ERT y TRL en tres umbrales diferentes de aplicación clínica.....</i>	<i>81</i>
<i>Tabla 7. Información utilizada en el Desarrollo de los scores clásicos de riesgo de reestenosis.....</i>	<i>102</i>
<i>Tabla 8. Factores relacionados con la aparición de reestenosis coronaria en diferentes estudios.....</i>	<i>105</i>
<i>Tabla 9. Contribuciones relevantes de la inteligencia artificial en las diferentes áreas de aplicación de la cardiología.....</i>	<i>113</i>

# Capítulo 1

## 1. Introducción

Existen pocos temas de actualidad equiparables a la posibilidad de la tecnología actual para desarrollar las mismas capacidades que el ser humano, incluso en medicina. Esta capacidad de simular los procesos de inteligencia humana por parte de máquinas o sistemas informáticos es lo que conocemos hoy en día como inteligencia artificial. El objetivo de este capítulo es aclarar diferentes términos como inteligencia artificial, *machine learning* (aprendizaje automático), *deep learning* (aprendizaje profundo), *data science* o *big data*; describir el concepto de inteligencia artificial y sus tipos; las técnicas de aprendizaje y metodología que se utiliza en el *machine learning*; el análisis en imagen con *deep learning*; la aportación de esta revolución tecnológica a la estadística clásica; sus limitaciones actuales y; sus aspectos legales.

En todos los puntos se analizan los conceptos más relevantes y se ponen ejemplos de su utilización en el ámbito de la medicina y, en concreto, de la cardiología.

### 1.1. Ciencia de datos

Los términos ciencia de datos, *big data*, inteligencia artificial y *machine learning* son diferentes, aunque pertenezcan a una misma disciplina: el análisis y extracción de información de los datos. Sin embargo, son conceptos que en ocasiones se utilizan indistintamente y de forma ambigua.

El término *big data* fue acuñado por primera vez en 2005 por R. Magoulas, que lo describió como un volumen masivo de datos que por tamaño supera al *software* tradicional de almacenaje y tratamiento. El *big data* gravita en torno a 7 uves: volumen, velocidad, variedad, veracidad, validez, volatilidad y valor de los datos. Es decir, el *big*

*data* genera información a gran velocidad, controlando la integridad de los datos y con mucha variedad en cuanto al tipo y naturaleza de los mismos (texto, imágenes, vídeos, diferentes sistemas, diferentes proveedores, etc.). En medicina, el *big data* es viable desde un marco de colaboración nacional, europeo o mundial en el que entidades hospitalarias homogeneizan y comparten datos para la creación de repositorios de gran volumen. Estos grandes repositorios de datos sirven de apoyo en la práctica diaria en el establecimiento de protocolos, diagnóstico estandarizado y temprano de enfermedades, pronóstico de la evolución de enfermedades y planificación del tratamiento de los pacientes<sup>1</sup>.

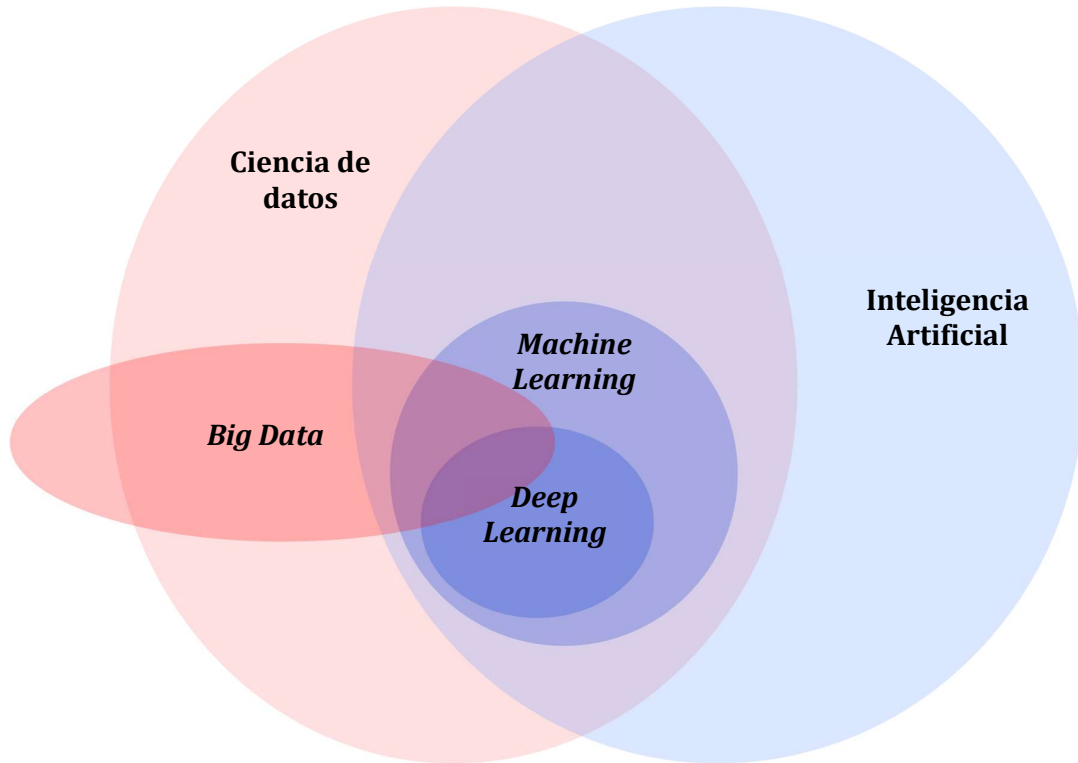
Este tipo de tratamiento de la información no sería posible sin los recientes avances tecnológicos en inteligencia artificial. La inteligencia artificial utiliza herramientas capaces de identificar los patrones existentes en los datos. La inteligencia artificial se define por primera vez en la década de los 50, y engloba múltiples subdisciplinas, desde los sistemas expertos o la robótica hasta el *machine learning*. Es justamente este último término sobre el que vamos a hacer énfasis en esta tesis debido a los grandes avances experimentados en los últimos años.

El término *machine learning* fue acuñado por primera vez en 1959 por A.L. Samuel<sup>2</sup>, y su objetivo es desarrollar algoritmos que permitan que sistemas informáticos tomen decisiones y aprendan del resultado de las mismas; que estos sistemas sean capaces de aprender a hacer algo sin haber sido programados explícitamente para ello.

Dentro de las técnicas de *machine learning* existe cada día mayor interés por los modelos de *deep learning*<sup>3</sup>. Estos basan su poder predictivo en redes neuronales artificiales y se caracterizan por disponer de múltiples capas de procesamiento (transformaciones) de la información, lo que les permite analizar conjuntos de datos con patrones de mayor complejidad. Estos algoritmos están permitiendo avances muy importantes en el reconocimiento de voz y de imágenes, aspecto este último muy relevante en medicina.

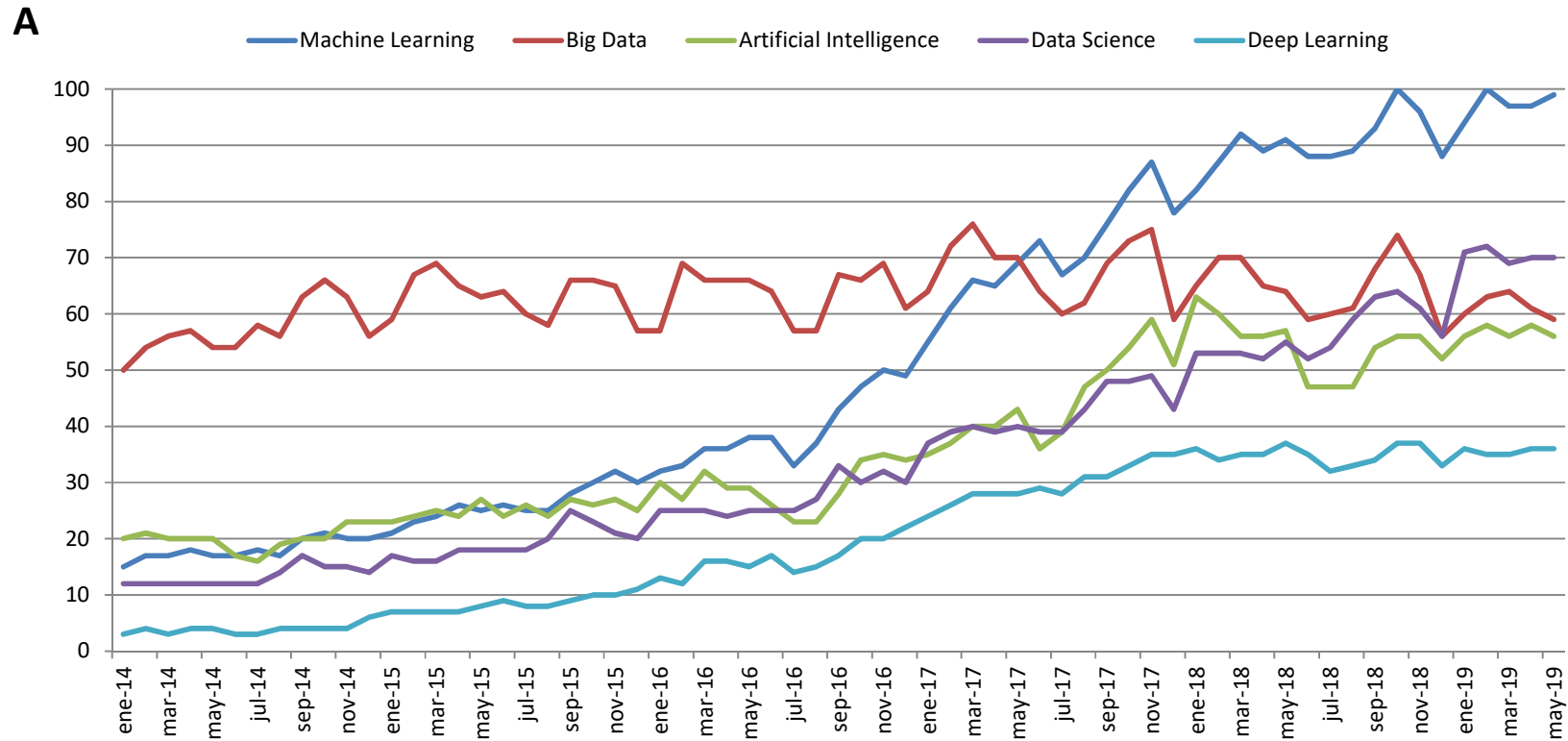
La ciencia de datos es el campo en el que podemos englobar todos los aspectos relacionados con la obtención del conocimiento existente en los datos. Involucra métodos científicos, procesos y sistemas para extraer conocimiento o un mejor entendimiento de los datos, normalmente aplicando técnicas novedosas en el tratamiento como la inteligencia artificial. Aunque es frecuente utilizar como sinónimos inteligencia artificial y *machine learning*; no lo son. La inteligencia artificial cubre un ámbito más amplio de esta revolución tecnológica incluyendo al *machine learning* y al

*deep learning*. En la **figura 1** se presenta un esquema que pone en relación todos los términos que hemos introducido.



**Figura 1.** Diagrama de Venn de los términos más comúnmente empleados en las disciplinas de la ciencia de datos.

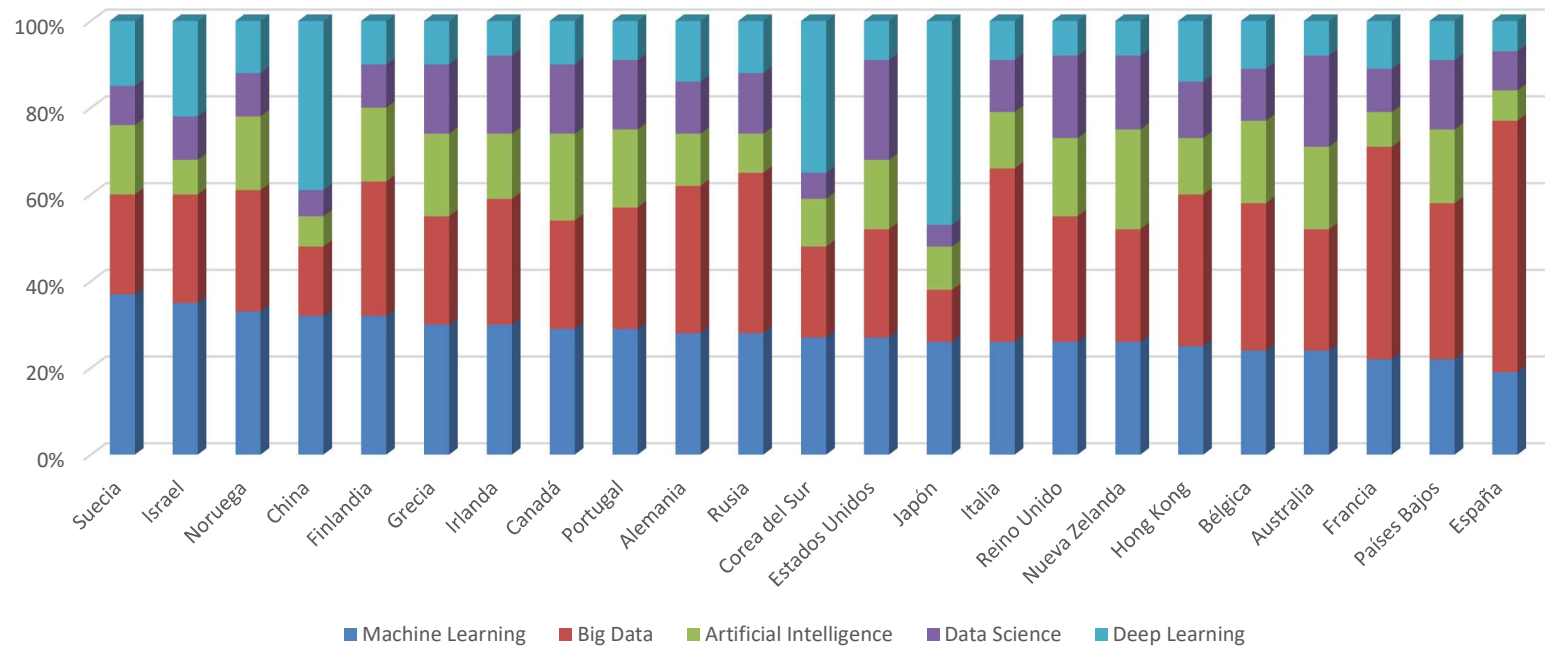
En la **figura 2A** se muestra el interés que despiertan estos conceptos durante los últimos años, con un cada vez más alto interés por el *machine learning*, la inteligencia artificial, la ciencia de datos y el *deep learning*. En la **figura 2B** se representa el término más buscado en cada país. En España, sur de Europa y Sudamérica sigue apareciendo *big data* como el término más relevante; en EE.UU., Canadá, Reino Unido, norte de Europa y Australia es *machine learning*; en China y Japón, *deep learning*.



**Figura 2A:** Búsquedas totales (fuente Google Trends), en los últimos 5 años completos, de términos relacionados con inteligencia artificial y ciencia de datos.

El eje vertical del diagrama representa la proporción de un tema con respecto al total de búsquedas sobre los temas.

**B**



**Figura 2B: Términos más buscados en cada país en el mismo periodo.**

## **1.2. Inteligencia artificial**

Hoy en día es difícil encontrar una definición universal de lo que se conoce como inteligencia artificial. A menudo se refiere al campo de las ciencias de la computación que trata de imitar los procesos cognitivos humanos, la capacidad de aprendizaje y el almacenamiento de conocimiento. Otras definiciones lo estudian en un sentido más amplio, a través de la compresión y construcción de entidades inteligentes, generalmente entendidas como software informático. Centrándonos en el tipo de tareas, podríamos considerar la inteligencia artificial como sistemas automáticos capaces de, por ejemplo, traducir un documento, reconocer a una persona por sus rasgos faciales, o conducir un automóvil. La inteligencia artificial no se limita a imitar tareas humanas, sino que en algunos casos es capaz de superar al mejor experto en la materia, tomando decisiones con menores tasas de error que una persona o identificando patrones imperceptibles para el ojo humano<sup>4</sup>. Por tanto, la inteligencia artificial nos permite analizar información con una aproximación diferente a la tradicional. Ya no nos limitamos a describir la información de la que disponemos dando respuesta a lo sucedido. La inteligencia artificial nos permite responder otras preguntas: ¿qué sucedió? (diagnosticar), ¿qué pasará? (predecir), y ¿qué debería hacer? (prescribir).

Dentro de las áreas de aplicación de la inteligencia artificial más relevantes, con aportación en el ámbito de la salud, podemos citar las siguientes: el reconocimiento automático de voz y el procesamiento de lenguaje natural, los algoritmos de predicción, recomendación o diagnóstico, la visión artificial y el análisis de imagen, la robótica y la inteligencia artificial y los sistemas expertos.

### **1.2.1. Reconocimiento automático de voz y procesamiento de lenguaje natural**

El objetivo de estas disciplinas es desarrollar mecanismos para la comunicación entre personas y máquinas utilizando lenguajes naturales. En el ámbito médico el reconocimiento automático de voz ya se está empleando para el registro de información clínica de pacientes<sup>5</sup>. Asimismo, el procesamiento de lenguaje natural nos está permitiendo, como ya se ve en varios ejemplos en nuestro país en el ámbito de la medicina, clasificar patologías o elegir la cohorte más adecuada para un estudio clínico analizando registros de historial médico<sup>6</sup>.

### **1.2.2. Algoritmos de predicción, recomendación o diagnóstico**

Esta es probablemente el área más madura en el mundo de la medicina, *machine learning* y *deep learning*. El mayor interés reside tanto en la automatización de tareas repetitivas, como puede ser la evaluación de una prueba diagnóstica; como en la generación de conocimiento a través del análisis de datos clínicos. En los próximos puntos se describe la metodología a seguir en este tipo de estudios, así como algunos casos de uso.

### **1.2.3. Visión artificial y análisis de imagen**

En esta disciplina científica se incluyen los métodos para adquirir, procesar, analizar y comprender las imágenes del mundo real con el fin de producir información numérica o simbólica para que puedan ser tratados por un ordenador. Estas técnicas han experimentado una auténtica revolución en los últimos años gracias a la aplicación de algoritmos de *deep learning* y es una de las disciplinas que más están aportando a la medicina hoy en día.

### **1.2.4. Robótica e inteligencia artificial**

La robótica y la inteligencia artificial es un área cuyo objetivo es construir sistemas físicos con comportamiento inteligente. Lleva desarrollándose durante mucho tiempo, pero está teniendo un nuevo auge apoyada en las distintas áreas descritas (reconocimiento automático de voz, visión artificial, etc.). En el ámbito médico se está desarrollando desde hace años en el ámbito de la cirugía; el sistema quirúrgico Da Vinci es un claro ejemplo<sup>7</sup>.

### **1.2.5. Sistemas expertos**

En inteligencia artificial, un sistema experto es un sistema computacional (*software*) que emula la capacidad de tomar decisiones de un humano experto. Estos sistemas están basados en reglas o incluso en casos clínicos. Desde hace tiempo se están desarrollando sistemas expertos médicos que tratan de simular el razonamiento del especialista y de proporcionar el diagnóstico probable y el manejo del paciente<sup>8</sup>.



### 1.3. *Machine learning*: técnicas

El *machine learning* es un área de la ciencia de datos que forma parte de lo que hoy conocemos como inteligencia artificial, creando sistemas que aprenden automáticamente, entendiendo por «aprender» como la capacidad de reconocer patrones complejos de forma autónoma, sin necesidad de intervención humana, en conjuntos de datos de cualquier tipo, ya sean numéricos, visuales, texto, etc. A medida que la experiencia de estos sistemas aumenta, es decir, se les suministra nuevos datos, su desempeño mejora hasta un límite que puede o no ser superior a la capacidad humana en dicha tarea.

Aunque existen varias técnicas de aprendizaje *machine learning*, suelen agruparse en 2 tipos: aprendizaje supervisado y no supervisado.

Las técnicas de aprendizaje supervisado son, sin lugar a duda, los métodos más empleados en *machine learning* y aquellos con mejores resultados. En estos procedimientos contaremos con un conjunto de datos del que conocemos la variable respuesta que queremos predecir (un diagnóstico, parámetro, segmentación, etc.) a través del correcto etiquetado de cada registro. En función del tipo de predicción, se emplean o algoritmos de clasificación o algoritmos regresivos. En el primer caso la meta es identificar 2 o más clases a través de una serie de variables. Por contra, los algoritmos regresivos buscan aproximarse lo más posible a un valor continuo.

En las técnicas de aprendizaje no supervisado, no contamos con la información sobre la variable que queremos predecir. Estas técnicas deben aprender de las relaciones entre los elementos de un conjunto de datos y clasificarlos sin apoyarse en etiquetas o categorías. Para ello, estos procedimientos buscan estructuras, patrones o características en los datos de origen que puedan ser reproducibles en nuevos conjuntos de datos. Los métodos de agrupamiento o *clustering* son los más empleados para esta tarea.

### 1.4. *Machine learning*: metodología

Construir un modelo *machine learning* no se reduce solo a utilizar un algoritmo de aprendizaje sobre una base de datos, sino que es todo un proceso<sup>9,10</sup>, que suele involucrar los pasos que se muestran en la **figura 3**.

Las primeras etapas son comunes a la estadística convencional; partiendo de datos crudos (*raw data*) pasamos a una información con datos estructurados (preprocesado), construyendo una base de datos inicial (paso 1). A partir de esta base de datos realizamos un análisis descriptivo y exploratorio, con el objetivo de identificar y

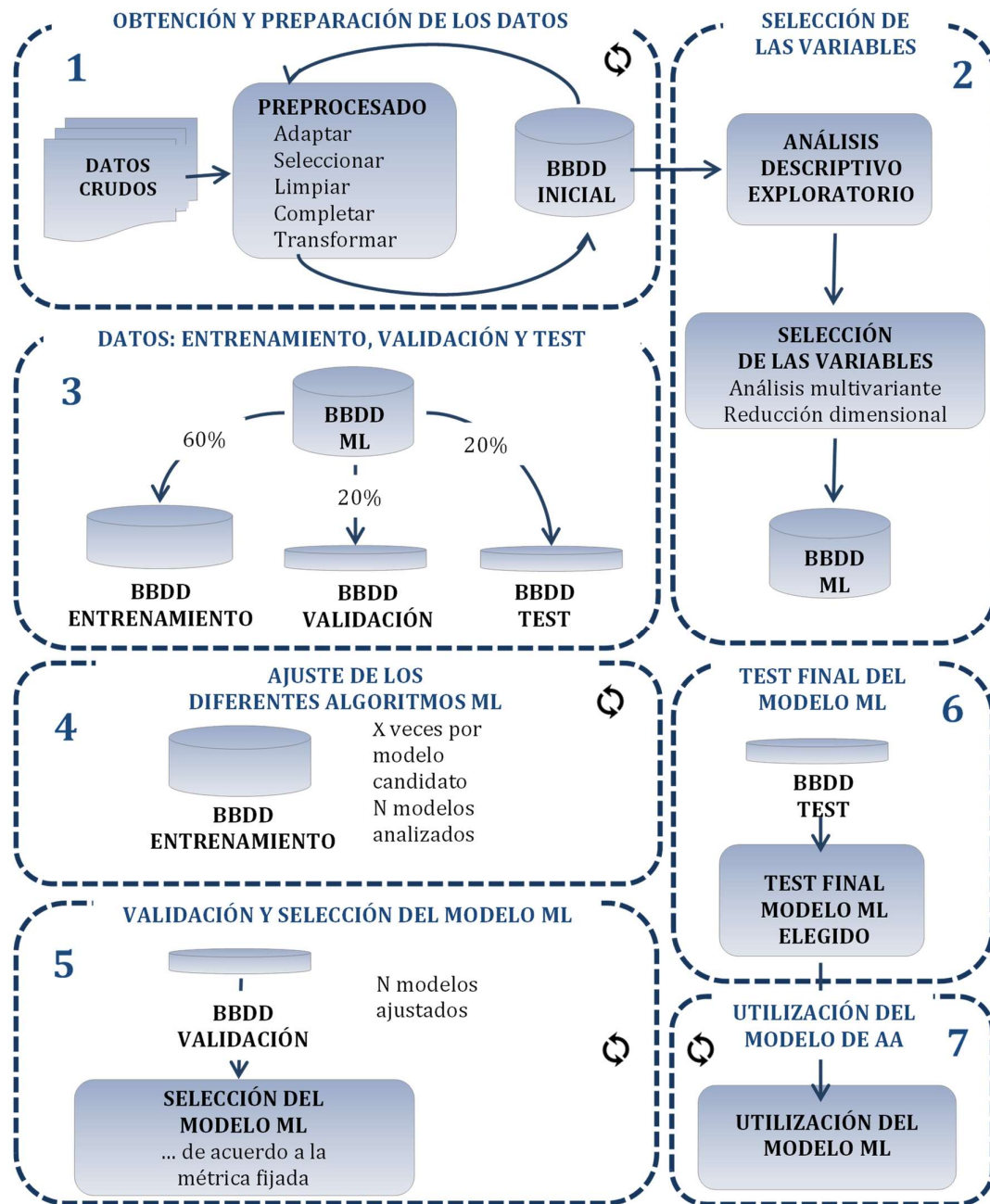
seleccionar las variables más significativas, sobre las que aplicaremos directamente los algoritmos *machine learning* (paso 2).

La siguiente etapa, específica de las técnicas *machine learning*, es la división del conjunto de datos en 3 subconjuntos: entrenamiento, validación y *test*; en porcentajes habituales del 60%, 20% y 20%, respectivamente (paso 3).

El conjunto de entrenamiento es el que usaremos para ajustar los distintos algoritmos *machine learning* seleccionados (paso 4). Existe una amplia gama de algoritmos de clasificación y regresión<sup>11</sup> (**tabla 1**), desde técnicas lineales más clásicas como la regresión logística o el análisis discriminante lineal, a otras más modernas como los algoritmos de tipo agregación o *bagging* (*random forest*) y de empaquetado o *boosting* (*xgboost*). Además, se pueden utilizar técnicas de submuestreo y sobremuestreo cuando, como a menudo ocurre en datos de salud, tenemos una categoría con muchos más casos que otra (muestras desbalanceadas). Ningún algoritmo es en principio mejor que otro, su capacidad para realizar un buen ajuste dependerá de las características de nuestros datos (número de variables, linealidad, normalidad, valores faltantes, variables continuas o categóricas, etc.).

Realizado el ajuste, se utiliza el subconjunto de datos de validación para evaluar la calidad del modelo (paso 5). Para ello trataremos de maximizar la métrica que más interese para nuestro caso particular: área bajo la curva ROC (característica operativa del receptor), exhaustividad, precisión, exactitud, u otras (**figura 4**). Es común que este proceso de entrenamiento-validación se repita un número de veces aleatorizando ambos subconjuntos, lo que se conoce como *k-folds*. El objetivo es optimizar los parámetros internos del algoritmo empleado, evaluar la robustez del modelo y saber si nuestro modelo está sub- o sobre-ajustando nuestros datos, tratando de encontrar un equilibrio entre ambos escenarios.

Construido un modelo final, se emplea el subconjunto de datos de *test* para comprobar que el modelo *machine learning* final se comporta según lo previsto con datos que no han sido utilizados para su construcción y validación (paso 6). Si este resultado difiere del obtenido en el conjunto de validación, probablemente el conjunto de datos utilizado para el entrenamiento es insuficiente y debe ser ampliado si queremos un estimador fiable antes de generalizar su utilización (paso 7). Para la implementación del *machine learning*, los lenguajes de programación de código abierto más usados en la actualidad son Python y R. Ambas plataformas disponen de librerías como Scikit-learn (Python) o Caret (R) con implementaciones de las técnicas y algoritmos más utilizadas. En esta tesis los desarrollos se han llevado a cabo utilizando Python.



**Figura 3.** Metodología de aplicación de un modelo *machine learning*. ML: *machine learning*; BBDD: base de datos.

Tabla 1. Algoritmos más utilizados en *machine learning*.

Algoritmo	Tipo	Descripción
<i>Random forest</i>	Métodos combinados	Combinación de árboles predictores no correlacionados
<i>Gradient boosting</i>	Métodos combinados	Combinación de árboles predictores escalonados
Regresión logística	Regresión	Análisis de regresión utilizado para predecir el resultado de una variable categórica
Máquinas de vectores de soporte	Clasificador supervisado	Clasificador a través de construcción de hiperplanos disociadores
K vecinos más próximos	Clasificador supervisado	Estimación de la función de densidad de las variables predictoras en función de las clases
Análisis discriminante lineal	Discriminante lineal	Generalización del discriminante lineal de Fisher
Clasificador bayesiano ingenuo	Clasificador supervisado probabilístico	Clasificador probabilístico fundamentado en el teorema de Bayés

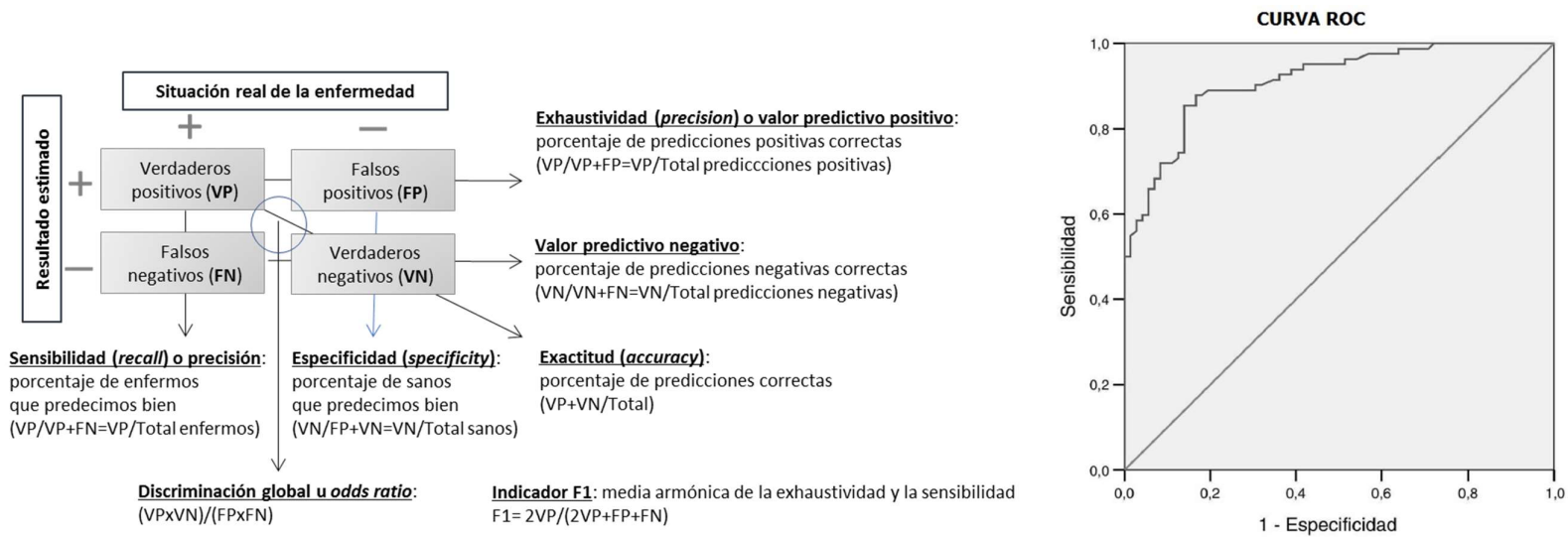


Figura 4. Métricas comúnmente empleadas en la evaluación de algoritmos de clasificación supervisada. Entre paréntesis se indican los términos en inglés

## 1.5. *Deep learning* y la imagen médica

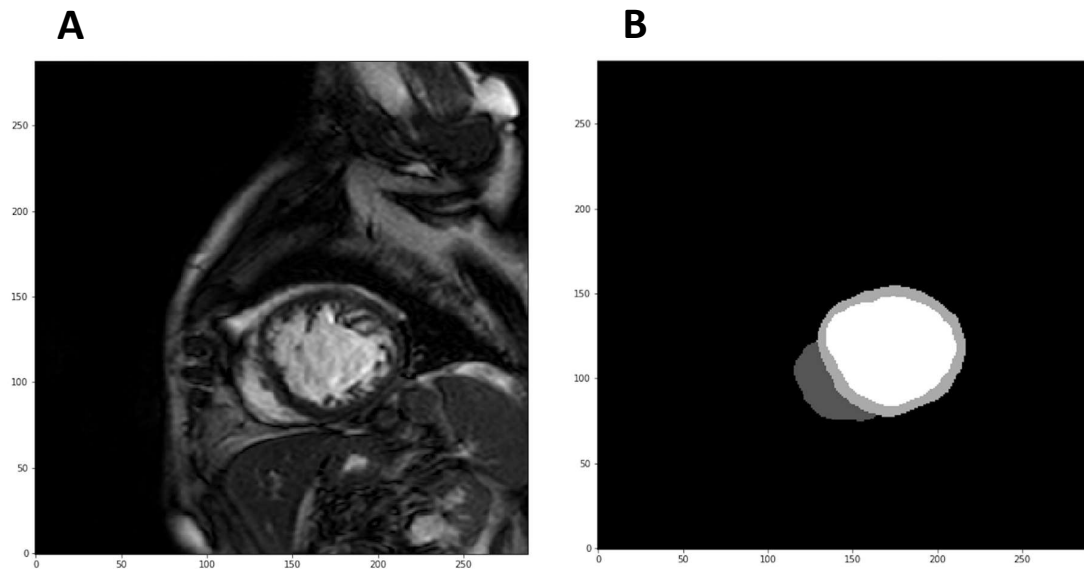
El análisis de imagen es el campo en el que la inteligencia artificial está avanzando a mayor velocidad, y por tanto de gran interés en medicina. El análisis que se hace de las diferentes imágenes médicas puede ser tedioso y consume gran parte de tiempo. Existen ya herramientas para el tratamiento automático de imágenes, basados en redes neuronales, que podrían mejorar el rendimiento en el análisis de imágenes médicas.

Las redes neuronales<sup>3</sup> son un tipo de algoritmos de inteligencia artificial que guardan analogía con el proceso de aprendizaje que se da en las propias neuronas del cerebro. Desde su formulación como modelo computacional han ido surgiendo algoritmos de entrenamiento y arquitecturas de redes que han mejorado considerablemente la precisión y eficiencia del aprendizaje con cada vez menores cantidades de datos para su entrenamiento. Cuando estas arquitecturas constan de numerosas capas de neuronas, es cuando actualmente utilizamos el término *deep learning*. En la actualidad, entre las redes neuronales más empleadas se encuentran las *convolutional neural networks*, las *recursive neural networks*, las *generative adversarial networks* o las *U-nets*, cada una con diferentes usos y subtipos de arquitecturas. Las redes neuronales son además muy flexibles a la hora de su utilización en contextos supervisados, no supervisados o de aprendizaje por refuerzo. Para su implementación, se dispone de diferentes plataformas de código abierto como TensorFlow, Pytorch, Keras o Caffe. Como desventajas de las técnicas de *deep learning*, se puede destacar su alto coste computacional. Además, se requiere un alto grado de pericia para su ajuste correcto y, en el caso del aprendizaje supervisado, se necesita contar con un conjunto de imágenes informadas manualmente, lo cual en ocasiones puede resultar muy costoso.

Dentro del campo de la imagen cardiaca existen diferentes objetivos a la hora de emplear técnicas de inteligencia artificial. Estos pasan por la segmentación e identificación de las diferentes estructuras del corazón (**figura 5**)<sup>12,13</sup>, la clasificación de imágenes catalogadas con diferentes patologías, la detección y segmentación de lesiones, el registro de imágenes o la clasificación de tejidos a partir de imagen histológica. Otro tipo de tareas se relacionan con la generación de imágenes artificiales lo más realistas posibles.

La investigación con técnicas *deep learning* aplicadas en la imagen cardiológica ha crecido considerablemente en los últimos 3 años. Sin embargo, el uso de técnicas de *deep learning* en imagen médica ha calado algo más en otras disciplinas, como la neurología o la neumología<sup>14</sup>. Entre los retos que probablemente se explorarán en los

próximos años se encuentran el análisis combinado con diferentes fuentes de imagen, la incorporación de datos clínicos e informes médicos o el estudio de la evolución temporal de imágenes médicas.



**Figura 5.** Ejemplo de segmentación e identificación automáticas del ventrículo izquierdo y derecho.

Realizado en el Servicio de Cardiología del Hospital Universitario de Salamanca a partir de imágenes obtenidas con una resonancia Philips Achieva 1.5 T12 mediante *deep learning*. A partir de imágenes originales (A), la red neuronal libre empleada es capaz de identificar y segmentar ventrículo izquierdo y derecho (B)<sup>13</sup>. El color blanco identifica el ventrículo izquierdo, el color gris claro el miocardio ventricular izquierdo y, el color gris oscuro el ventrículo derecho.

## 1.6. Limitaciones del *machine learning* y del *deep learning*

La principal característica de la inteligencia artificial en el contexto de los modelos *machine learning* o *deep learning* es el hecho de que el aprendizaje se basa en la identificación de patrones en conjuntos de datos. Esta filosofía es a la vez su punto fuerte, pues los ordenadores son extremadamente eficientes y precisos para encontrar dichos patrones cuando estos existen, pero es a la vez una limitación. En primer lugar, porque la cantidad de datos que se requieren para obtener un modelo fiable puede llegar a ser muy elevada. Esto puede ser un problema en medicina, donde los sistemas de recogida automática de datos están comenzando a implantarse, deben cumplir los criterios legales y éticos establecidos, y algunas patologías poco frecuentes necesariamente cuentan con pocos estudios.

Por otro lado, aun contando con algoritmos con buenos resultados para un conjunto de datos, los modelos *machine learning* y *deep learning* son incapaces de detectar y clasificar correctamente casos que no hayan visto previamente. En relación con esto, la fiabilidad y calidad de la fuente de datos es fundamental para que un algoritmo sea realista y correcto. Si existe sesgo en el propio conjunto de datos empleado para ajustar el modelo, es posible que éste no generalice bien a otras poblaciones<sup>15</sup>. Este problema no es nuevo en el ámbito médico donde se han aceptado escalas de riesgo basadas en datos de otras poblaciones, por ejemplo, Framingham para riesgo cardiovascular, dando lugar a resultados erróneos sobreestimando o infraestimando el riesgo al aplicarse a otras poblaciones<sup>16</sup>. Existen ya iniciativas para estandarizar los requisitos que debe cumplir un algoritmo de inteligencia artificial para su posible aplicación práctica<sup>17</sup>. Por todas estas razones, parece claro que se va a necesitar un especial esfuerzo en la integración de conjuntos de datos procedentes de diferentes poblaciones.

Otra limitación importante es la opacidad e interpretabilidad de los modelos *machine learning* y en especial, del *deep learning*. Estas técnicas se emplean como «cajas negras» a las que damos unos *inputs* para obtener una predicción como *output*. Así, las máquinas nos ofrecen respuestas en forma de predictores, pero no una explicación a los mismos. Aunque existen diferentes metodologías para poder interpretar los resultados de un modelo y verificar su correcto funcionamiento a través del análisis de los pesos específicos de parámetros o variables, resaltar las partes más discriminantes o la explicación individualizada de cada predicción<sup>18</sup>; a día de hoy no es posible saber de forma clara por qué los modelos *machine learning* y *deep learning* más complejos toman una determinada decisión, lo que dificulta la generación de



conocimiento y la aplicabilidad de la inteligencia artificial en un ámbito tan crítico como es la medicina.

## **1.7. *Machine learning* y la estadística clásica: similitudes y diferencias**

Aunque los campos de los que provienen son diferentes, ambas especialidades guardan rasgos comunes, aunque con finalidad diferente.

El pilar de la estadística clásica es la inferencia. Extender lo observado en una muestra a toda la población, generalmente creando un modelo matemático que define la relación entre las variables. *Machine learning* basa su fortaleza en la predicción basada en la información disponible, sin necesidad de tener que conocer cuáles son los mecanismos que relacionan entre sí las variables. Estas características no son exclusivas ya que en mayor o menor medida ambas disciplinas usan técnicas inferenciales para mejorar sus resultados (*machine learning*) o algoritmos predictivos para corroborar conclusiones inferenciales (estadística).

Las técnicas *machine learning*, comparadas con las estadísticas, no requieren de suposiciones previas sobre las variables, tienen la capacidad de manejar casos en los que tenemos datos faltantes y mejoran su fiabilidad cuando disponemos de grandes volúmenes de datos detectando relaciones complejas entre variables.

## **1.8. Aspectos legales y éticos de la inteligencia artificial**

Existen cuestiones a la hora de aplicar la inteligencia artificial en el mundo de la medicina no estrictamente relacionadas con la calidad de los algoritmos de *machine learning* y *deep learning*. Estas tienen que ver con los problemas legales del tratamiento automático de datos personales o cómo aplicar *machine learning* en el día a día de la práctica clínica.

Prácticamente todas las técnicas de inteligencia artificial estudiadas hasta la fecha requieren de una determinada cantidad de datos para entrenar y validar los modelos predictivos. En medicina, la información empleada es especialmente sensible, ya que generalmente consiste en datos personales y clínicos de pacientes. Por ello, el uso adecuado de esta información desde un punto de vista legal es un aspecto crítico.

La protección de los datos de carácter personal es un derecho fundamental recogido en la Constitución Española; sin embargo, se reconoce la posibilidad de utilización de dichos datos bajo ciertas condiciones si con ello se da un servicio a la sociedad. Si tanto la organización que recoge los datos como la que los trata, tienen residencia en España, desde 2018 deben ajustarse a la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales<sup>19</sup>. Según esta, se permite el uso de datos personales con fines estadísticos, siempre y cuando los resultados sean de carácter agregado. Además, también es necesario adoptar medidas para anonimizar y proteger dichos datos y se reconocen una serie de derechos de los interesados como el de ser informado, el de acceso o el de rectificación. Por todo ello, dado que la utilización de servicios de computación en la nube es muy habitual en inteligencia artificial por su flexibilidad y mayor potencia computacional, los proveedores más importantes han adaptado sus políticas de procesado de datos para suministrar un servicio completamente seguro. Además de suscribirse al Código de Conducta CISPE (*Cloud Infrastructure Services Providers in Europe*) para garantizar los estándares de protección de datos apropiados asociados con la legislación. Asimismo, en situaciones en las que se emplee un modelo de inteligencia artificial como herramienta diagnóstica o de toma de decisiones, es obligado informar al paciente y obtener su consentimiento.

La discriminación sutil inherente a la prestación de servicios de salud puede significar también un problema a la hora de desarrollar y aplicar modelos de inteligencia artificial<sup>15</sup>. Por ejemplo, consideramos la edad al racionar recursos escasos como el trasplante cardiaco. Tal consideración puede llevar a profecías autocumplidas: si los médicos retiramos la atención en pacientes por su edad avanzada, los sistemas *machine learning* pueden concluir que la atención en pacientes mayores es siempre fatal. Por otro lado, también es posible que los modelos *machine learning* permitan ayudar a resolver disparidades en la prestación de asistencia médica si se pudieran construir algoritmos para compensar sesgos conocidos o identificar áreas de investigación necesarias.

Existe una amplia discusión sobre cómo la inteligencia artificial afectará al flujo de trabajo del personal médico. La inteligencia artificial no tiene por qué eliminar puestos de trabajo, sino más bien desplazar tareas de la práctica médica. Trabajos rutinarios y fatigantes pueden ser llevados a cabo por máquinas para liberar tiempo y permitir a los profesionales médicos su desempeño en las tareas más complejas y sensibles. Sin embargo, el empleo de herramientas diagnósticas automáticas puede conllevar problemas de sesgo en la toma de decisiones y en la depuración de responsabilidades en caso de error. Un médico puede inclinarse a revisar una prueba diagnóstica de forma más relajada si un algoritmo previamente ha catalogado el caso como negativo. El

diseño de sistemas de toma de decisiones que involucren máquinas y humanos es un aspecto crucial para la inteligencia artificial, existen modelos intermedios entre la práctica clínica convencional y los sistemas totalmente automatizados, cada uno con diferentes características de eficiencia en coste y tiempo, riesgo o interpretabilidad<sup>20</sup>. La implementación de estos sistemas será un área en la que el papel del médico como validador será fundamental.

## Capítulo 2

### 2. Hipótesis y objetivos

#### 2.1. Hipótesis

Las técnicas *machine learning* pueden aplicarse al procesamiento con conjuntos de datos pequeños y no balanceados en medicina para la predicción de reestenosis tras implante de stent coronario; permitiendo automatizar las decisiones y mejorar la predicción de discriminadores clásicos elaborados con estadística inferencial.

#### 2.2. Objetivos

##### 2.2.1 Objetivo Principal

Desarrollar un modelo *machine learning* para predecir la aparición de reestenosis en pacientes con infarto agudo de miocardio sometidos a intervencionismo coronario percutáneo con implante de un stent (ICP). La angioplastia coronaria con implante de stent se ha convertido en práctica habitual en la revascularización de los vasos coronarios con enfermedad aterosclerótica obstructiva significativa. El éxito a largo plazo de este procedimiento está limitado por la reestenosis del stent, un proceso patológico que provoca estrechamiento arterial recurrente en el sitio de la angioplastia.

La incidencia de reestenosis se cifra en la actualidad en torno al 10%. Identificar qué pacientes harán reestenosis es un desafío clínico importante; ya que puede manifestarse como un nuevo infarto agudo de miocardio o forzar una nueva resvascularización del vaso afectado, que en casos de reestenosis recurrente representa un importante dilema terapéutico.

### 2.2.2 Objetivos Secundarios

1. Construir un conjunto de datos para entrenamiento, validación y test que pueda ser utilizado para determinar la técnica de *machine learning* a utilizar en el clasificador y que a su vez permita evaluar otras técnicas, ahora no disponibles, en el futuro.
2. Aplicar y analizar la metodología *cross-validation* y de ajuste de hiperparámetros en el procesamiento de datos pequeños y no balanceados como el caso de reestenosis coronaria donde la realización de angiografía coronaria en el seguimiento no es rutinaria y donde 1 de cada 10 pacientes presenta reestenosis.
3. Entrenar y validar los clasificadores *machine learning* más empleados en la literatura, y escoger el que ofrezca mejores resultados.
4. Comparar el valor predictivo del mejor clasificador *machine learning* con discriminadores predictivos inferenciales desarrollados con modelos estadísticos clásicos y validados en la literatura para el ejemplo de reestenosis tras implante de un stent coronario; *scores* de riesgo PRESTO-1, PRESTO-2 y TLR.
5. Aplicar la declaración TRIPOD, en la validación externa de los *scores* de riesgo para reestenosis clásicos (PRESTO-1, PRESTO-2 y TLR); utilizando los cocientes de probabilidades calculados en sus respectivos estudios de cohortes originales.
6. Incorporar al análisis con curvas ROC del resultado predictivo de los modelos *machine learning* y *scores* de riesgo clásico, el resultado de dibujar la gráfica entre exhaustividad (*precision*) y sensibilidad (*recall*) (curva PR); sabiendo que el objetivo ideal en una curva PR es que se acerque lo máximo posible a la esquina superior derecha (alta exhaustividad y alta sensibilidad).
7. Calcular y establecer el rendimiento de clasificación del mejor modelo *machine learning* para diferentes umbrales de corte basados en las curvas ROC y PR de

acuerdo con la sensibilidad, especificidad, exhaustividad y valor predictivo negativo; en comparación con el mejor *score* clásico.

8. Aplicar metodología MDA (*Mean Decrease Accuracy*) para identificar qué variables tienen mayor impacto en el modelo predictivo *machine learning* desarrollado.
9. Construir un software que permita una interpretación directa e individualizada del modelo *machine learning* desarrollado.

## Capítulo 3

### 3. Material

#### 3.1. Datos

Los datos utilizados para entrenar y validar los modelos *machine learning* provienen del ensayo GRACIA-3<sup>21</sup>. De forma resumida, el ensayo GRACIA-3 fue un ensayo clínico aleatorizado, abierto, multicéntrico, 2x2 que compara la eficacia del *stent* fármaco liberador de paclitaxel con el *stent* metálico convencional. Los pacientes con infarto agudo de miocardio con elevación del ST (STEMI) se reclutaron en 20 hospitales españoles entre octubre 2004 y agosto 2006.

Los criterios de inclusión del estudio fueron:

- Edad igual o superior a 18 años.
- Sintomatología compatible con isquemia coronaria aguda con una duración entre 30 minutos y 12 horas hasta la aleatorización.
- Elevación del segmento ST en el electrocardiograma estándar de 12 derivaciones de al menos 0,1 mV en dos o más derivaciones de miembros o de 0,2 mV en dos o más derivaciones continuas precordiales, o bloqueo completo de rama izquierda, o ritmo de marcapasos.
- No presentar síntomas ni signos de insuficiencia cardiaca grave (clase Killip <3).

Los criterios de exclusión del estudio fueron:

- Shock cardiogénico definido como presión arterial sistólica <90 mm Hg sin mejoría tras la infusión de volumen o <100 mm Hg en pacientes en tratamiento inotrópico y sin bradicardia.

- Sospecha de complicación mecánica postinfarto agudo de miocardio.
- Revascularización coronaria quirúrgica previa.
- Enfermedad no cardíaca que condicione la esperanza de vida a corto plazo.
- Mujeres que pudiesen estar embarazadas a menos que se realice un test de embarazo y su resultado sea negativo.
- Sangrado activo o cirugía reciente (<2 semanas) que contraindique el uso de heparina, tirofiban, o antiagregantes orales.
- Contraindicación mayores o menores para recibir tratamiento fibrinolítico.
- Historia de hipersensibilidad a aspirina, ticlopidina, clopidogrel, heparina, tirofiban, o acero inoxidable.
- Enfermedad hepática grave que contraindique el uso de clopidogrel.
- Trombopenia conocida <100.000.
- Enfermedad coronaria multivaso conocida y no susceptible de revascularizar.
- Enfermedad vascular periférica grave conocida que dificulte la realización de la angiografía.

Los objetivos principales del estudio fueron dos: comparar la incidencia de reestenosis coronaria entre el *stent* recubierto de paclitaxel frente al convencional y, evaluar el beneficio del uso de tirofiban intravenoso administrado 120 minutos después de administrar el tratamiento fibrinolítico y antes de realizar la revascularización coronaria percutánea sobre el flujo epicárdico y miocárdico de la arteria responsable del infarto.

Para determinar la incidencia de reestenosis, se realizó una angiografía coronaria antes e inmediatamente después de la intervención y después de 12 meses de seguimiento. Todas las cuantificaciones angiográficas para la evaluación de las dimensiones de la arteria coronaria antes y después de la implantación del *stent* se analizaron en un laboratorio central de angiografía independiente (ICICOR, Valladolid, España) con un sistema de medida informático bien validado para la investigación (Medis, Leesburg, Va). La tasa de reestenosis fue evaluada de forma ciega por un lector experimentado y no directamente involucrado en el estudio.



La **figura 6** muestra el flujo de pacientes del ensayo, donde finalmente fueron incluidos 436 pacientes de los cuales 299 (69%) se sometieron a un seguimiento angiográfico a los 12 meses. El conjunto de datos utilizado en esta tesis para el desarrollo del modelo *machine learning* está compuesto por 263 de 299 (88%) pacientes en los que estaban disponibles 68 variables (*features*) relacionadas con características demográficas, clínicas, características del *stent* y angiográficas. Esta cohorte de desarrollo del modelo se identificó antes de fusionar los datos de entrada con el resultado para evitar el sesgo de selección.

El comité ejecutivo del GRACIA-3 y el Comité Ético del Hospital Universitario de Salamanca aprobaron el uso retrospectivo de los datos identificados del ensayo para este estudio (PI201902178).

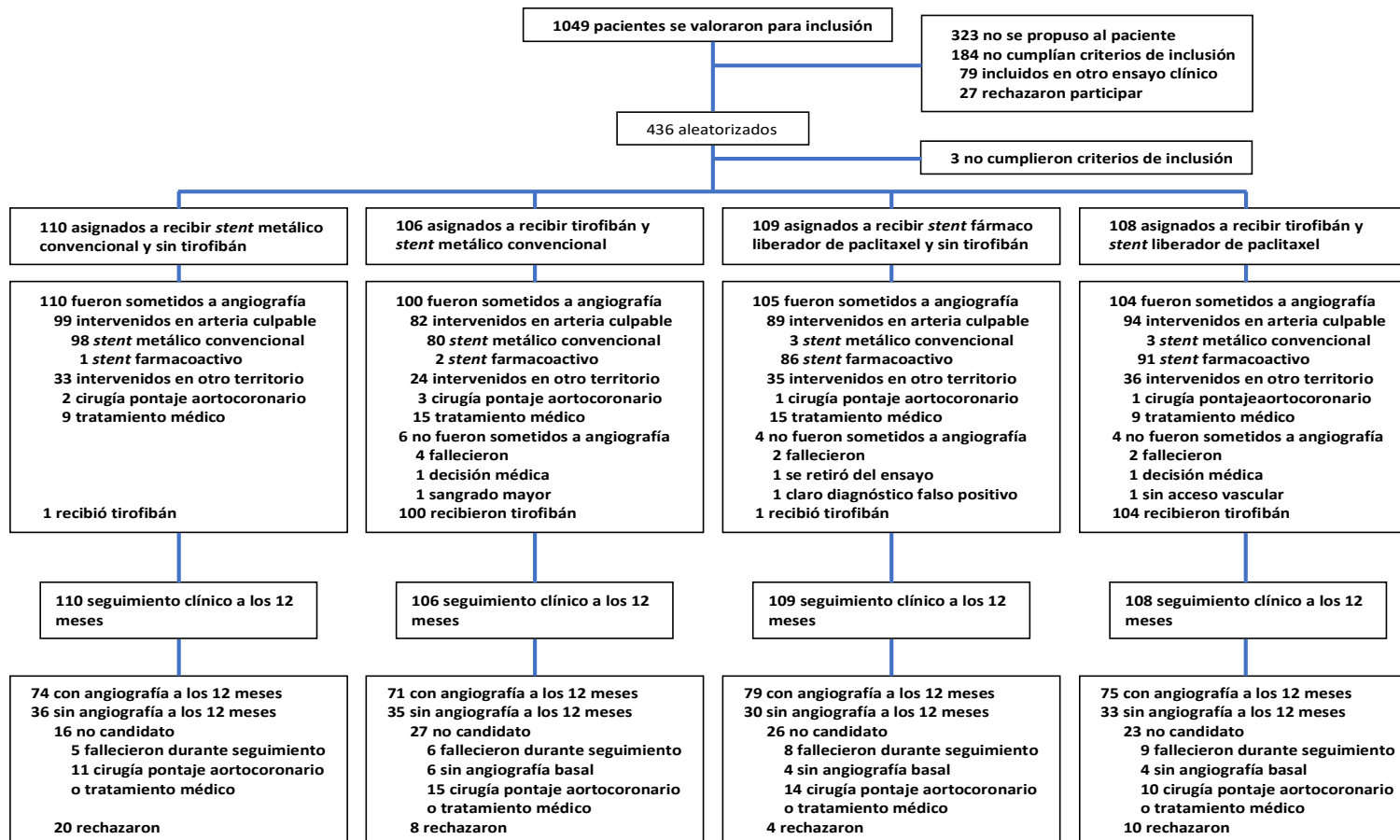


Figura 6. Flujo de pacientes ensayo GRACIA-3.

## 3.2. Software utilizado

Para todo el desarrollo empleado en el proyecto se ha utilizado como lenguaje de programación Python, en su versión 3.7, utilizando como entorno de ejecución Anaconda3 (soporte para Python 3.x).

Para la implementación de los distintos algoritmos utilizados y el cálculo de todos los resultados de esta tesis doctoral se ha empleado la biblioteca de código abierto scikit-learn<sup>22</sup>.

La librería scikit-learn<sup>22</sup> dispone de numerosas funciones entre las que destacan:

- Algoritmos de aprendizaje supervisados. La gran mayoría de algoritmos *machine learning* que entran en la clasificación de aprendizaje supervisado forman parte de scikit-learn, desde los modelos lineales generalizados, como la regresión lineal, la máquina de vectores de soporte, árboles de decisión, y métodos bayesianos.
- Validación cruzada. La librería cuenta con las técnicas necesarias para verificar la precisión de los modelos supervisados como la validación cruzada.
- Algoritmos de aprendizaje no supervisados. De la misma forma que para el aprendizaje supervisado, esta librería cuenta con una gran variedad de algoritmos disponibles para aprendizaje no supervisado, comenzando por la agrupación, el análisis factorial, el análisis de componentes principales y las redes neuronales no supervisadas.
- Varios conjuntos de datos o *dataset*. Esta librería pone a disposición varios *dataset* que son muy útiles para poder practicar con Python. Varios de estos *dataset* fueron utilizados a lo largo de esta tesis antes de implementar nuestro modelo.
- Extracción y selección de características. Esta librería es muy útil para identificar las variables más significativas a partir de las cuales crear modelos supervisados.
- Comunidad. Una de las principales razones detrás de uso de herramientas de código abierto es la gran comunidad que tiene.

También se han utilizado las librerías Python eli5<sup>23</sup>, Tkinter y Bokeh.

Eli5 se ha utilizado para ejecutar el algoritmo MDA y calcular la importancia de las variables en la predicción del modelo. Eli5 es una librería Python que ayuda a depurar clasificadores *machine learning* y explicar sus predicciones.

Tkinter se ha utilizado para el desarrollo de la calculadora que se ejecuta en PC o MAC. Tkinter es una librería Python orientada a diseñar interfaces gráficas de usuario (GUI) sencillas para aplicaciones de escritorio desarrolladas en Python. Se considera un estándar para el desarrollo de interfaces gráficas de usuario para Python.

Bokeh se ha utilizado para representar las curvas ROC y PR en las calculadoras y situar el paciente analizado en las mismas. Bokeh es una librería de visualización interactiva de Python pensada para ser utilizada en los navegadores web modernos. Su objetivo es proporcionar una construcción de gráficos sencillos, novedosos e interactivos.

### **3.3. Hardware utilizado**

Para la construcción del modelo machine learning, el proyecto ha sido desarrollado y ejecutado en la nube, en Amazon Web Services (AWS). Dentro de los recursos disponibles en AWS se ha utilizado una instancia EC2 (*hardware*). Debido a que el proyecto no maneja una base de datos grande, puede ser ejecutado en un PC o MAC con una capacidad media, instalando previamente el soporte de Python (Anaconda o similar) especificado en el apartado de software utilizado.

Para la construcción de la calculadora web, el proyecto ha sido desarrollado y ejecutado en la nube, en Google Colaboratory (Google Colab). Google Colab es la herramienta de Google en la nube para ejecutar código Python y crear modelos de machine learning a través de la nube de Google y con la posibilidad de hacer uso de sus GPUs (*graphics processing unit*).

# Capítulo 4

## 4. Metodología

La **figura 7** incluida en la siguiente página muestra las fases que hemos seguido para construir nuestro modelo *machine learning*:

- Preparación del modelo.
- Entrenamiento del modelo.
- Evaluación del modelo.

En la preparación del modelo: definimos la tarea u objetivo del mismo, seleccionamos los sujetos del estudio, discriminamos las variables de entrada (*features*) y la variable de salida o resultado (*label* o *output*), y procedemos a la recopilación y preparación de datos. Como hemos indicado en el capítulo de material, el conjunto de datos procesados proviene del estudio GRACIA-3<sup>21</sup> y se dividió en dos conjuntos. El primero, el conjunto de datos de entrenamiento, se utilizó para contruir el modelo; el segundo, el conjunto de *test*, se utilizó para evaluar el funcionamiento del modelo.

En la construcción y entrenamiento del modelo, el modelo se desarrolló incluyendo una selección de variables, el entretenimiento con diferentes clasificadores *machine learning*, el ajuste de hiperparámetros, y la aplicación de una metodología de validación cruzada (*cross-validation*)<sup>24,25</sup> para elegir la mejor selección de variables y los mejores hiperparámetros (validación cruzada interior).

Finalmente, al evaluar el modelo, el conjunto de datos de *test* se ejecutó con el modelo *machine learning* final para estimar el rendimiento del modelo desarrollado en un escenario clínico del mundo real. Para esta evaluación se utilizó un modelo de validación cruzada (validación cruzada exterior). Adicionalmente, se realizaron comparaciones del modelo *machine learning* frente a *scores* clásicos de predicción de riesgo de sufrir reestenosis.

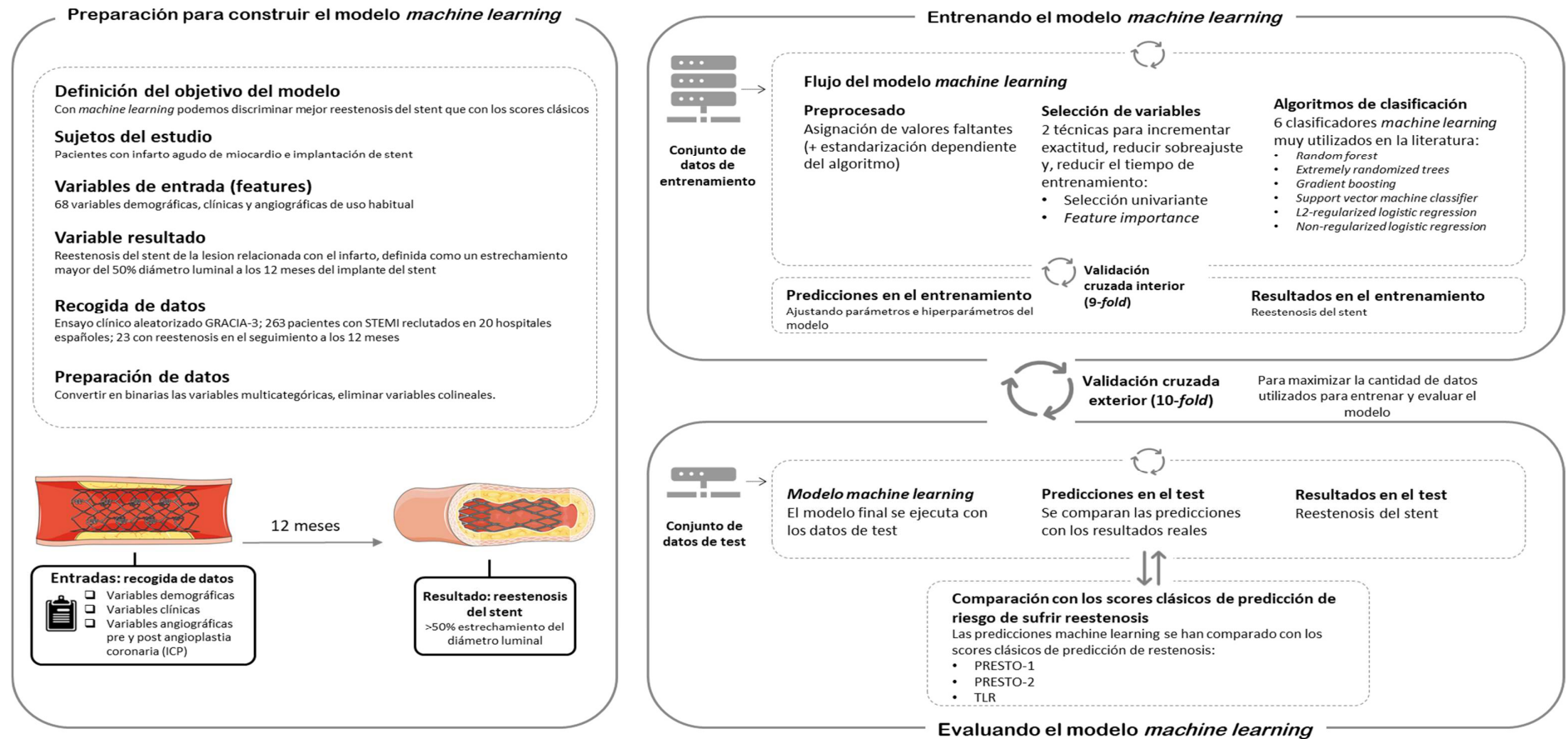


Figura 7. Resumen de las fases que se han seguido para construir el modelo *machine learning*.

## 4.1. Preparación para construir el modelo *machine learning*

### 4.1.1. Definición del objetivo del modelo

El objetivo de nuestro modelo es predecir la reestenosis del *stent* a los 12 meses en pacientes con infarto agudo de miocardio con elevación del segmento ST sometidos a intervención coronaria percutánea con *stent* (ICP). Para desarrollar un modelo *machine learning* de fácil aplicación, realizamos nuestra predicción sin variables adicionales a las obtenidas en la práctica clínica habitual.

### 4.1.2. Sujetos del estudio

El conjunto de sujetos utilizado para el desarrollo del modelo está compuesto por 263 pacientes del ensayo GRACIA-3<sup>21</sup> que finalmente se sometieron a seguimiento angiográfico 12 meses después del episodio de infarto y colocación de *stent*; 23 (8,7%) de los cuales presentaron reestenosis coronaria.

### 4.1.3. Variables de entrada

Comenzamos nuestro modelo *machine learning* supervisado especificando nuestras variables de entrada (*features*)<sup>10</sup>. Los datos de entrada consisten en las características demográficas y clínicas del paciente, examen físico, parámetros bioquímicos y el análisis coronario cuantitativo (QCA) para la evaluación de las dimensiones de la arteria coronaria antes y después de la implantación del *stent*. Todos ellos parámetros disponibles en cualquier procedimiento rutinario de intervencionismo coronario. En total, se utilizaron 68 variables procedentes del ensayo GRACIA-3.

### 4.1.4. Variable resultado

En cuanto al resultado correspondiente (*label*), se definió reestenosis angiográfica como un estrechamiento superior al 50% del diámetro de la luz en el segmento objetivo; definido como todas las porciones del vaso coronario que recibieron tratamiento dentro de la zona del *stent*, incluidos los márgnes proximal y distal de 5 mm<sup>26</sup>.

#### 4.1.5. Preparación de datos

Un paso importante en la preparación de cualquier modelo *machine learning* consiste en preprocesar los datos crudos (*raw data*) para convertirlos en un conjunto de variables que los algoritmos *machine learning* puedan utilizar.

Para este propósito, las variables con varias categorías se codificaron en variables binarias. La información relacionada con el segmento coronario donde se produjo la lesión se agrupó en 3 variables atendiendo al vaso donde se realizó la ICP (arteria descendente anterior izquierda, circunfleja izquierda o arteria coronaria derecha). Los valores faltantes se completaron con la mediana y la moda de cada variable continua y categórica, respectivamente.

## 4.2. Entrenando el modelo *machine learning*

Los modelos *machine learning* se basan en distintos conjuntos de datos secundarios (subconjuntos del conjunto de datos disponible): uno para el entrenamiento y otro conjunto de datos de *test* para evaluar el modelo. Los conjuntos de datos de entrenamiento son la base de aprendizaje de los modelos y el conjunto de *test* se utiliza con los modelos finales ajustados para evaluar qué modelo es mejor.

Se requieren diferentes enfoques en la preparación de datos cuando el conjunto de datos utilizado para entrenar el modelo no es grande y tiene resultados (*outputs*) desbalanceados. El objetivo principal es maximizar la cantidad de datos utilizados para entrenar el modelo y estimar su rendimiento futuro. Este fue nuestro caso con 23 de 263 pacientes que desarrollaron reestenosis en el seguimiento a los 12 meses. Para abordar este problema, realizamos una metodología k-fold<sup>24</sup> de validación cruzada, para dividir aleatoriamente el conjunto de datos en k partes del mismo tamaño: las partes k-1 formaban nuestro conjunto de datos de entrenamiento y el conjunto restante se utilizó como conjunto de datos de test para evaluar el modelo. En nuestro modelo, utilizamos un k-fold estratificado dividido en 10 partes (k=10), una configuración habitual que garantiza un número mínimo de casos de la clase minoritaria tanto en el entrenamiento como en la evaluación (**figura 8**).



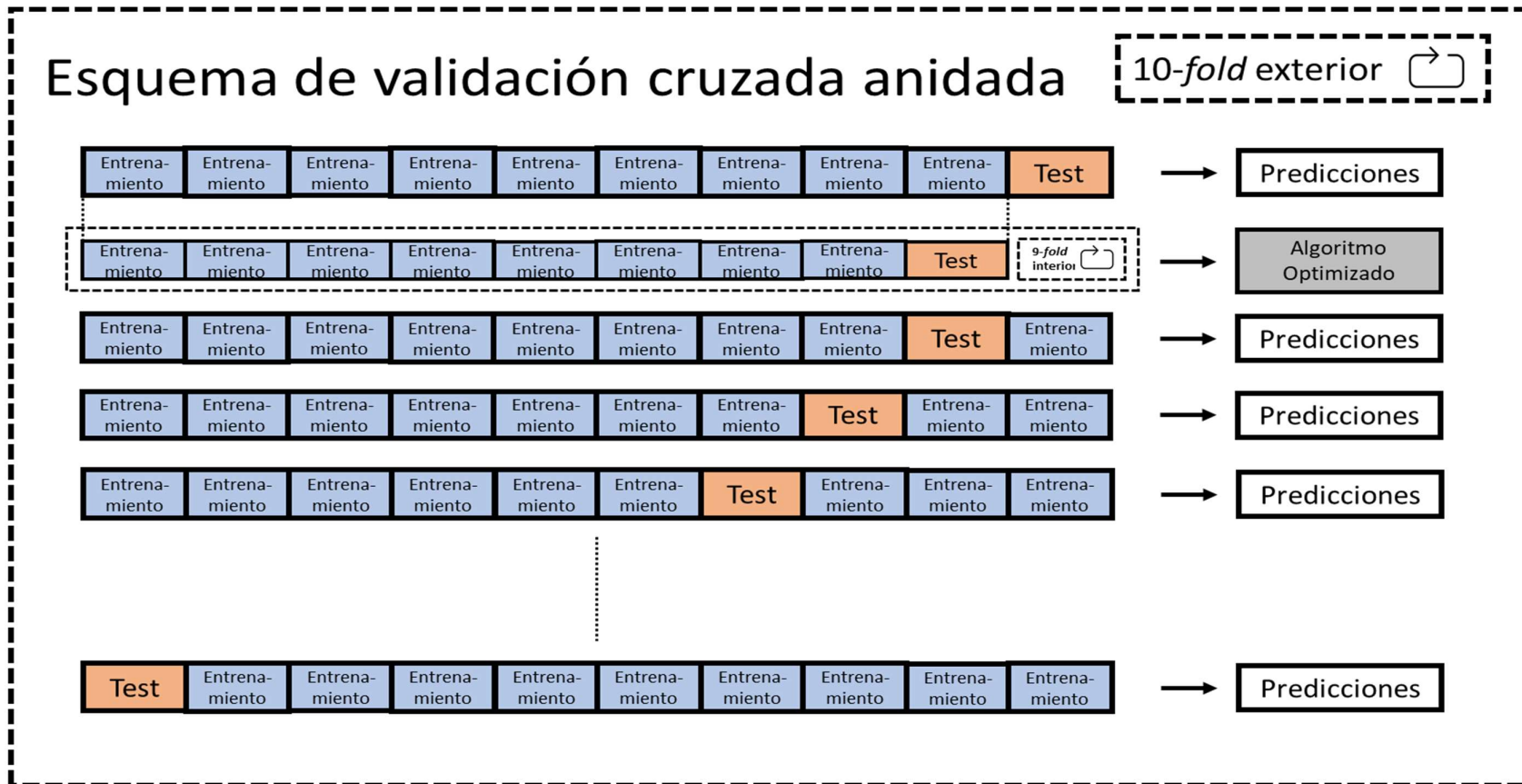


Figura 8. Esquema de validación cruzada anidada seguido para entrenar y validar el modelo.

### 4.2.1. Modelos *machine learning* y técnicas de selección de variables

Aplicamos y comparamos el rendimiento de seis clasificadores (algoritmos) *machine learning* que se usan ampliamente en la literatura<sup>10,27</sup>:

- *Random forest*
- *Extremely randomized trees*
- *Gradient boosting*
- *Support vector machine classifier*
- *L2-regularized logistic regression*
- *Non-regularized logistic regression*

Tanto el algoritmo *random forest* como el algoritmo *extremely randomized trees* son técnicas que utilizan una aproximación de perturbar y combinar especialmente diseñada para árboles de decisión<sup>28-30</sup>, entendiendo perturbar como la introducción de aleatoriedad en la construcción del clasificador. La predicción se obtiene utilizando un conjunto de clasificadores de tipo árboles de decisión, que se obtienen de forma ligeramente diferente en *random forest* y en *extremely randomized trees*.

En *random forest*<sup>28,31</sup>, cada árbol del conjunto se construye con un subconjunto de datos del conjunto de datos de entrenamiento obtenido por muestreo con reemplazo (*bootstrap*).

Cuando se divide cada nodo durante la construcción del árbol, la mejor división se busca bien utilizando todas las variables o un subconjunto aleatorio de tamaño, *max\_features*.

El propósito de utilizar estas dos fuentes de aleatoriedad es reducir la varianza del estimador final. De hecho, los árboles de decisión individuales muestran elevada varianza y tienden al sobreajuste. El modelo combinado con todos los árboles consigue reducir mucho la varianza.

En la implementación usada de *random forest* (scikit-learn<sup>22</sup>) no se sigue exactamente la publicación original sobre *random forest*<sup>32</sup>, combina los clasificadores promediando su predicción probabilística, en vez de dejar que cada clasificador vote por una clase concreta.

En los *extremely randomized trees*<sup>33</sup> la aleatorización va un paso más allá en la forma en la que se realizan las divisiones. Igual que en *random forest* se elige un subconjunto aleatorio de las variables, pero en vez de buscar los umbrales óptimos, los que más discriminan (impureza, entropía), en el subconjunto de variables elegidas, se eligen estos umbrales de forma aleatoria para cada variable y el mejor de ellos es el que se utiliza para dividir.

Con esta aproximación consigue reducir un poco más la varianza, con un ligero incremento en algunos casos del sesgo.

Los parámetros principales para ajustar *random forest* y *extremely randomized trees* son *n\_estimators* y *max\_features*<sup>34,35</sup>. El primero es el número de árboles utilizados en el modelo y el segundo es el tamaño de los subconjuntos aleatorios de variables para considerar a la hora de dividir un nodo.

En cuanto al número de árboles, cuantos más árboles mejor, pero a medida que se incrementa se necesita más tiempo de computación y a partir de un punto ya no se obtiene una mejoría significativa. En cuanto al número de variables utilizadas para dividir, cuantas menos mayor reducción de la varianza, pero también un incremento mayor del sesgo.

Otros parámetros que suelen fijarse cuando se utilizan estos algoritmos son:

*max\_depth=None* y *min\_samples\_split=2*, que son la profundidad del árbol y el número de casos en cada nodo para dividirlo, lo que no limita la construcción de los árboles, aunque se consuma mucha memoria.

La mejor aproximación para fijar los parámetros es hacer un esquema de validación cruzada, que es lo que hemos hecho en nuestro caso (*k-fold* interno).

En *random forest* suele fijarse *bootstrap* a true para que elija subconjuntos de datos para entrenamiento, pudiendo calcularse el error con los datos no utilizados en el entrenamiento. En *extremely randomized trees* suele fijarse *bootstrap* a false y se utiliza el conjunto completo de datos de entrenamiento para la construcción de los árboles, calculando posteriormente el error con el esquema de validación cruzada.

Para mejorar el rendimiento en la construcción de estos modelos las librerías permiten paralelización, que se regula con otro parámetro, de forma que se aprovechen al máximo el número de *cores* disponibles para esta tarea en la máquina.

En nuestro caso los parámetros fijados han buscado la mayor simplicidad, dando buenos resultados, poca profundidad, muchos árboles y una única variable para dividir los nodos. Los parámetros fijos utilizados se fijaron en pruebas previas, observando

aquellos parámetros que apenas variaban entre *folds*. Se dejó como hiperparámetro para iterar la profundidad de los árboles porque fue donde observamos mayor variación entre *folds*.

*Gradient boosting*<sup>36</sup> o *gradient tree boosting* es una generalización del *boosting*. Es un algoritmo que puede usarse tanto para clasificación como para regresión. Sus ventajas son que puede usarse con variables heterogéneas, tiene un poder predictivo alto y que es robusto frente a *outliers*. Su desventaja es la escalabilidad, como todo *boosting*, que debido a su naturaleza secuencial es difícilmente paralelizable.

En clasificación, el número de clasificadores débiles (*weak learners*, normalmente árboles de clasificación) se controla con el parámetro *n\_estimators*. El tamaño de cada árbol se puede controlar fijando la profundidad del árbol con el parámetro *max\_depth* o fijando el número de *leaf nodes* con el parámetro *max\_leaf\_nodes*. El parámetro *learning\_rate* es un hiperparámetro que se fija en el rango (0.0, 1.0), y controla el sobreajuste. Es un parámetro que controla cuánto contribuye cada árbol a la mejora del modelo, de forma que con valores bajos se evita contribuciones altas, aunque son necesarios más árboles para entrenar el modelo.

*Support vector machine classifier*<sup>37,38</sup> es un clasificador para obtener alta precisión. Este clasificador separa los datos de entrada (los pacientes) utilizando un hiperplano con el mayor margen posible entre los conjuntos de datos a separar. Su rasgo más relevante es lo que se conoce como el truco del *kernel* (*kernel trick*) que se utiliza para manejar espacios de entrada no lineales, no separables por un hiperplano. Este truco consiste en transformar el espacio de entrada en un espacio de mayor dimensión en el que los conjuntos sí sean separables por un hiperplano.

En nuestro caso el *kernel* utilizado ha sido RBF (*Radial Basis Function*) *kernel='rbf'* que es no lineal.

Otro de los parámetros que habitualmente se fijan con *support vector machine classifier* es el parámetro C. Este parámetro controla la regularización. La regularización es la penalización por los errores cometidos en la clasificación. Cuando C es pequeño, el clasificador admite más sujetos mal clasificados (alto sesgo, baja varianza). Cuando C es alto, el clasificador penaliza mucho los errores y los intenta evitar (bajo sesgo, alta varianza). Valores de C muy altos nos llevan a problemas de sobreajuste.

Hemos entrenado también 2 modelos basados en regresión logística, por ser nuestro problema un problema de clasificación binario, que en medicina se gestionaría en muchos casos con este tipo de algoritmos de decisión.

Los dos modelos de regresión logística se han entrenado uno de ellos con regularización y otro sin regularizar. El primero suele ser más habitual en *machine learning*, pero se ha incluido el no regularizado para comparar resultados. De las posibles regularizaciones se ha seleccionado la de Ridge (L2)<sup>39</sup>.

Cuando usamos regularización minimizamos la complejidad del modelo a la vez que minimizamos la función de coste. Esto resulta en modelos más simples que tienden a generalizar mejor. Los modelos que son excesivamente complejos tienden a sobreajustar, es decir, a encontrar una solución que funciona muy bien para los datos de entrenamiento, pero muy mal para datos nuevos. Nos interesan los modelos que además de aprender bien, tengan un buen rendimiento con datos nuevos.

En la regularización Lasso, también llamada L1<sup>40</sup>, la complejidad  $C$  se mide como la media del valor absoluto de los coeficientes del modelo. Lasso nos va a servir de ayuda cuando sospechemos que varias de las variables de entrada son irrelevantes. Al usar la regularización Lasso, estamos fomentando que la solución sea poco densa. Es decir, favorecemos que algunos de los coeficientes acaben valiendo 0. Esto puede ser útil para descubrir cuáles de los atributos de entrada son relevantes y, en general, para obtener un modelo que generalice mejor. Lasso funciona mejor cuando las variables de entrada no están muy correlacionadas entre ellas.

En la regularización Ridge, también llamada L2, la complejidad  $C$  se mide como la media del cuadrado de los coeficientes del modelo. Al igual que ocurría en Lasso, la regularización Ridge se puede aplicar a varias técnicas de aprendizaje automático. Ridge nos va a servir de ayuda cuando sospechemos que varias de las variables de entrada estén correlacionadas entre ellas. Ridge hace que los coeficientes acaben siendo más pequeños. Esta disminución de los coeficientes minimiza el efecto de la correlación entre las variables de entrada y hace que el modelo generalice mejor. Ridge funciona mejor cuando la mayoría de las variables pueden ser relevantes.

ElasticNet combina las regularizaciones L1 y L2. Con un parámetro podemos indicar qué importancia relativa tienen Lasso y Ridge, respectivamente. Usaremos ElasticNet cuando tengamos un gran número de variables de entrada, algunas de ellas serán irrelevantes y otras estarán correlacionadas entre sí.

Los seis clasificadores (*random forest*, *extremely randomized trees*, *gradient boosting*, *support vector machine classifier*, *L2-regularized logistic regression* y *non-regularized logistic regression*) fueron entrenados con 2 técnicas de selección de variables diferentes, haciendo un total de 12 modelos comparables diferentes.

El primer método de selección de variables se basó en la selección univariante donde se utilizaron pruebas ANOVA para seleccionar aquellas variables que tenían la relación más fuerte con la presencia de reestenosis a los 12 meses después de la implantación del *stent*. El segundo método de selección de variables se basó en la importancia de las variables (*feature importance*), donde usando un clasificador *extremely randomized trees* obtenemos una puntuación para cada variable del conjunto de datos; cuanto mayor sea la puntuación, más importante es la variable respecto a la variable a predecir.

El código desarrollado utilizado para entrenar los modelos se puede consultar en [https://github.com/PedroDorado/Restenosis\\_IA\\_Gracia3](https://github.com/PedroDorado/Restenosis_IA_Gracia3)

#### **4.2.2. Esquema de validación cruzada (*cross-validation*) y ajuste de hiperparámetros (*hyperparameter tuning*)**

Todos los modelos se entrenaron y evaluaron con el esquema de validación cruzada anteriormente mencionado de 10-folds y 20 repeticiones (**figura 8**).

También se llevó a cabo un ajuste de hiperparámetros de los modelos para mejorar el rendimiento de los clasificadores *machine learning*. Los hiperparámetros de los modelos son las propiedades que rigen todo el proceso de entrenamiento<sup>41</sup>; p.ej. el número de variables seleccionadas por la técnica de selección de variables o en los clasificadores basados en la combinación de árboles de decisión (*ensemble classifiers*), la profundidad de éstos. La configuración para la selección de hiperparámetros incluyó la mejor estrategia de selección de variables de acuerdo con el mejor resultado utilizando una validación cruzada estratificada (*g-fold* interior en **figura 8**) que garantiza un número mínimo de casos de la clase minoritaria y mide el resultado como el área bajo la curva ROC (AUC-ROC), que es la misma métrica que utilizamos para evaluar nuestros modelos finales. Este enfoque es computacionalmente costoso debido a su naturaleza altamente iterativa, pero es asumible cuando se trata de conjuntos de datos pequeños y permite el uso de todo el conjunto de datos sin separar una parte de los datos para calcular el error de generalización<sup>42,43</sup>. La optimización de los parámetros se realizó con un algoritmo de búsqueda aleatorio (*randomized searching algorithm*)<sup>44</sup>. Los valores fijos de los hiperparámetros no optimizados y los rangos de los optimizados para cada algoritmo de clasificación y selección de variables se muestran en la siguiente tabla (**tabla 2**).

Cada modelo *machine learning* está compuesto de un algoritmo de selección de variables y un clasificador. En la tabla se muestran los hiperparámetros fijos y los rangos de los hiperparámetros ajustados con el esquema de validación cruzada anteriormente descrito. Los hiperparámetros se han etiquetado de acuerdo con su implementación en la librería Python scikit-learn. Los hiperparámetros no incluidos en la tabla son los estándares inicializados por los métodos de la librería scikit-learn.

**Tabla 2.** Lista de los hiperparámetros ajustados en los modelos machine learning utilizados para construir nuestro modelo final.

ANOVA: selección de variables basada en ANOVA; FI: selección de variables basada en la importancia dada por un clasificador *random forest*; LR: regresión logística; RF: *Random Forest*; ERT: *extremely randomized trees*; GB: *gradient boosting classifier*; SVC: *support vector machine classifier with RBF kernel*; LR\_noreg: regresión logística sin regularización.

Algoritmos	Hiperparámetros fijos	Rango de hiperparámetros ajustados
ANOVA	Ninguno	k: [2,4,6,8,10,12,14,16,18,20]
FI	Ninguno	k: [2,4,6,8,10,12,14,16,18,20]
LR	penalty = 'l2'	C: [0.1,1,10,100,1000]
RF	n_estimators = 500, max_features = 1	max_depth: [2,3,4,5,10]
ERT	n_estimators = 500, max_features = 1	max_depth: [2,3,4,5,10]
GB	n_estimators = 500, max_features = 1	max_depth: [2,3,4,5,10]
SVC	kernel = 'rbf', probability = True	C: [0.1,1,10,100,1000]
LR_noreg	penalty = 'none' solver = 'newton-cg'	Ninguno

Modelo final: ANOVA + ERT.

Variables seleccionadas: diabetes; enfermedad en 2 o más vasos; trombos post-ICP; flujo TIMI post-ICP; colesterol total anormal.

Hiperparámetros ERT: n\_estimators = 500, max\_features = 1, max\_depth = 4



## 4.3. Evaluando el modelo *machine learning*

### 4.3.1. Esquema de evaluación

Una vez completado el entrenamiento de todos los modelos *machine learning* siguiendo lo anteriormente descrito, los modelos se evaluaron con el esquema de validación cruzada mencionado anteriormente y representado en la **figura 8** como *10-fold* exterior.

Este esquema se ejecutó de forma repetida utilizando 20 repeticiones. Para esta fase de *test* utilizamos el conjunto de datos que habíamos dejado separado del entrenamiento en el proceso de validación cruzada.

Cada vez que evaluamos una nueva instancia de nuestros modelos usando estos datos seguimos todos los pasos previamente descritos para el entrenamiento del modelo, incluyendo a selección de variables y el ajuste de hiperparámetros.

A continuación (**figura 9**) se muestra un ejemplo de uno de los *extremely randomized trees* (ERT) utilizado en nuestro modelo para predecir la reestenosis del *stent*.

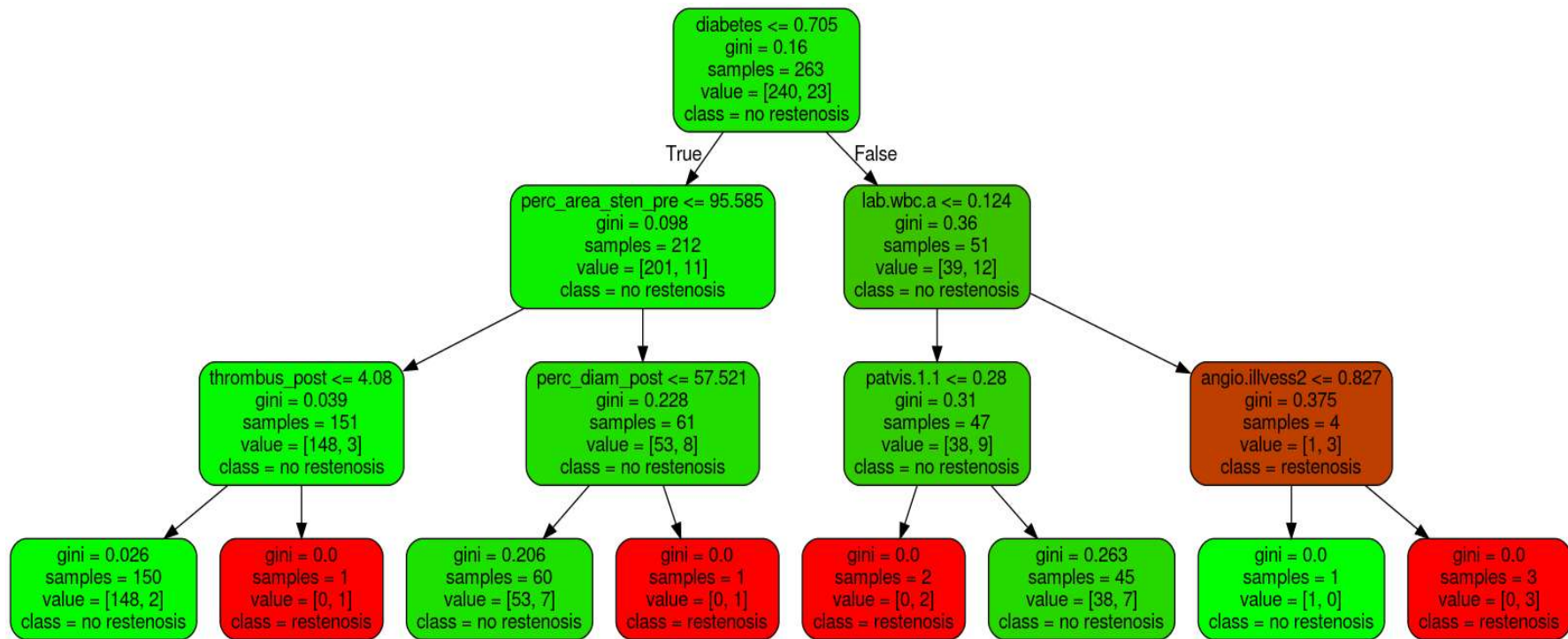


Figure 9. Ejemplo de uno de los extremely randomized trees (ERT) utilizado en nuestro modelo para predecir la reestenosis del *stent*.

Los nodos verde y rojo muestran divisiones con todos los casos en las clases sin reestenosis o reestenosis, respectivamente; el color mixto muestra nodos con una proporción intermedia. Los nodos en el último nivel determinan el resultado de clasificación del árbol.

#### 4.3.2. Comparación del modelo *machine learning* propuesto con los scores clásicos de predicción de riesgo de sufrir reestenosis

Hemos comparado el rendimiento del modelo *machine learning* desarrollado con los discriminadores predictivos inferenciales desarrollados con modelos estadísticos clásicos: PRESTO-1, PRESTO-2 and TLR.

De forma resumida, los scores PRESTO se desarrollaron a partir del ensayo PRESTO (*Prevention of Reestenosis with Tranilast and its Outcomes*) que incluyó 1.312 pacientes<sup>45</sup>. Los investigadores construyeron 2 scores de riesgo: PRESTO-1, que utilizó variables previas al implante de *stent* (sexo femenino, tamaño de los vasos, longitud de la lesión, diabetes, tabaquismo, lesión tipo C, cualquier ICP previa y angina inestable) que son fáciles de evaluar y con frecuencia se han informado como predictores robustos de reestenosis; PRESTO-2 utilizó variables clínicas y angiográficas relacionadas con reestenosis e identificadas a partir del conjunto de datos PRESTO (diabetes mellitus tratada, no fumador, tamaño del vaso, longitud de la lesión, lesión tipo C, ubicación ostial y ICP previa). Ambos scores alcanzaron un área bajo la curva ROC de 0,63 en la población PRESTO.

El score TLR (edad <60, ICP previa, angioplastia de tronco principal izquierdo sin protección, angioplastia de injerto de vena safena, diámetro mínimo del *stent* y longitud total del *stent*) utilizó datos del registro EVENT (*Evaluation of Drug-Eluting Stents and Ischemic Events*) con 8,829 pacientes<sup>46</sup>. TLR logró un área bajo la curva ROC de 0,68 en su cohorte de validación independiente.

Para la evaluación externa de los scores PRESTO y TLR que hemos llevado a cabo, utilizamos tanto los modelos de regresión logística generados por los coeficientes exactos que dichos estudios otorgan a cada variable predictora, así como los modelos generados por los valores enteros de dichos coeficientes, más típicamente empleados en la aplicación de los scores de riesgo a la práctica clínica diaria. La evaluación de modelos de regresión logística con ambos métodos es una de las recomendaciones establecidas en la declaración TRIPOD<sup>47</sup> para la evaluación externa e interna de scores diagnósticos y pronósticos. Por otra parte, es significativo también destacar que calculamos los scores de riesgo usando el mismo esquema de validación cruzada y divisiones descritas en el desarrollo del modelo *machine learning*, obteniendo 200 submuestras de validación aleatorias (10-*folds* con 20 repeticiones), con el objetivo de poder comparar posteriormente las métricas de evaluación de dichos scores con las de los modelos ML desarrollados en este estudio.

### 4.3.3. Métricas utilizadas en la evaluación

Las diferencias en las características demográficas, clínicas del paciente, examen físico, fracción de eyección, parámetros bioquímicos rutinarios, y análisis coronario cuantitativo para cada partición (reestenosis y no reestenosis) se compararon utilizando pruebas de  $\chi^2$  para variables categóricas y ANOVA para variables continuas.

En un problema de decisión binaria como nuestro estudio, nuestros clasificadores etiquetaron los ejemplos como positivos (reestenosis coronaria) o negativos (no reestenosis coronaria). Las decisiones de los clasificadores se representaron en tablas de contingencia con cuatro categorías (**Tabla 3**): los verdaderos positivos fueron ejemplos correctamente etiquetados como positivos, los falsos positivos se refieren a ejemplos negativos incorrectamente etiquetados como positivos; los verdaderos negativos corresponden a negativos etiquetados correctamente como negativos y; finalmente, los falsos negativos se refieren a ejemplos positivos incorrectamente etiquetados como negativos; donde la exhaustividad o valor predictivo positivo se define como el número de verdaderos positivos sobre todos los positivos predichos, la sensibilidad como el número de verdaderos positivos sobre el número real de positivos; especificidad como el número de verdaderos negativos sobre el número real de negativos; y el valor predictivo negativo como el número de verdaderos negativos sobre el número de negativos predichos.

El análisis de la curva ROC<sup>48</sup> se utilizó para evaluar la capacidad predictiva de cada modelo *machine learning* y los *scores* clínicos de riesgo clásicos de dos maneras diferentes. Primero, concatenamos las predicciones de todos los conjuntos de *test* y construimos una curva ROC. En segundo lugar, calculamos y guardamos el área bajo la curva ROC de cada conjunto de *test* del esquema de validación cruzada exterior. Utilizamos la primera estrategia para trazar las curvas ROC con fines ilustrativos y la segunda para calcular las puntuaciones medias, los intervalos de confianza y la significancia estadística en las pruebas t de Student entre los modelos *machine learning* y los *scores* clínicos de riesgo clásicos<sup>49</sup>. Calculamos estos intervalos de confianza y la significancia de la prueba t teniendo en cuenta que los resultados en cada *k-fold* no son completamente independientes<sup>50</sup>.

Además, calculamos las curvas exhaustividad sensibilidad o curvas PR (*Precision-Recall*). Las curvas PR también se usan en *machine learning* en problemas de decisión binarios, ya que pueden dar una idea más intuitiva del comportamiento del modelo cuando los datos utilizados están desbalanceados<sup>51,52</sup>. Se representaron las curvas PR

tanto referidas a reestenosis como para no reestenosis, para una mejor interpretación de los resultados de los modelos. Para no reestenosis, exhaustividad y sensibilidad son equivalentes al valor predictivo negativo y la especificidad.

También hemos evaluado el rendimiento del modelo *machine learning* desarrollado en determinados escenarios de uso concretos, seleccionados según las curvas ROC y PR, de acuerdo con su sensibilidad, especificidad, exhaustividad y valor predictivo negativo.

Finalmente, calculamos la importancia de las variables del mejor modelo midiendo cómo disminuye el área bajo la curva ROC cuando una variable no está disponible a través del método conocido como *Permutation Importance or Mean Decrease Accuracy (MDA)*<sup>28</sup>. El método consiste en reemplazar cada variable del conjunto de datos de *test* con una variable columna de ruido aleatorio y medir el rendimiento del modelo *machine learning*. El peso de las variables con impacto positivo en el modelo predictivo se escala a 1. El algoritmo también está disponible en la librería de código abierto de Python, eli5<sup>23</sup>. Para que el ruido tenga la misma distribución que la variable original, lo que se hace es permutar los valores de dicha variable aleatoriamente y de esta manera se puede utilizar el mismo modelo. Si elimináramos la variable estaríamos cambiando el número de variables que metemos al modelo y habría que reentrenarlo.

#### **4.4. Trasladando el modelo *machine learning* a la práctica clínica. Desarrollo de una calculadora de riesgo de reestenosis coronaria**

Utilizando el modelo *machine learning* desarrollado y entrenado, hemos construido una calculadora de riesgo de reestenosis del stent tras ICP.

Esta calculadora se ha desarrollado para 2 entornos:

- Aplicación web (Google Colab). En la nube, ejecutable con cualquier navegador.
- Aplicación para ejecutar en cualquier PC (Windows, MAC) con Python (Tkinter).

#### 4.4.1. Calculadora con Google Colab

Google Colab<sup>53</sup> (*Colaboratory*) es la herramienta de Google en la nube para ejecutar código Python y crear modelos *machine learning* a través de la nube de Google y con la posibilidad de hacer uso de sus GPUs (*graphics processing unit*).

Google Colaboratory es un entorno gratuito de Jupyter Notebook que no requiere configuración y que se ejecuta completamente en la nube.

Python es un lenguaje de programación que permite su utilización con diversos paradigmas de programación, como programación orientada a objetos o programación funcional, pero quizás la característica más importante en este contexto sea que también se usa como lenguaje de scripting y que es un lenguaje interpretado.

Jupyter es un entorno interactivo que permite desarrollar código Python de manera dinámica. Jupyter se ejecuta en local como una aplicación cliente-servidor y posibilita tanto la ejecución de código como la escritura de texto, favoreciendo así la interactividad del entorno y que se pueda entender el código como la lectura de un documento.

La principal ventaja que ofrece Google Colab es que libera a nuestra máquina de tener que llevar a cabo un trabajo demasiado costoso tanto en tiempo como en potencia o incluso nos permite realizar ese trabajo si nuestra máquina no cuenta con recursos suficientemente potentes.

Otro de los beneficios que tiene lo indica el propio nombre, *colaboratory*, es decir, colaborativo, nos permite realizar tareas en la nube y compartir nuestros cuadernos (desarrollos) si necesitamos trabajar en equipo.

Google Colab es gratuito y forma parte de la suite de aplicaciones de Google en la nube. Por ello, para utilizarlo basta con acceder a nuestra cuenta de Google y, o bien entrar directamente al enlace de [Google Colab](#) o ir a nuestro Google Drive, pulsar el botón de «Nuevo» y desplegar el menú de «Más» para seleccionar «Colaboratory», lo que creará un nuevo cuaderno (*notebook*).

Esta forma de trabajar es visualmente poco atractiva, pero nos permite compartir y ejecutar los modelos en entornos web colaborativos que no consumen apenas recursos.

#### 4.4.2. Calculadora con Tkinter

Esta solución se ha desarrollado completa en Python utilizando como añadido una librería para interfaces gráficas: Tkinter<sup>54</sup>. Tkinter es una librería Python orientada a diseñar interfaces gráficas de usuario (GUI) sencillas para aplicaciones de escritorio desarrolladas en Python. Se considera un estándar para el desarrollo de interfaces gráficas de usuario para Python.

Esta calculadora ya tiene un aspecto visual mejorado y es más sencilla para su manejo por un profesional médico.

#### 4.4.3. Representación del paciente en las curvas ROC y PR

En las calculadoras de riesgo se han representado las curvas ROC y PR (reestenosis y no reestenosis) del modelo desarrollado, ubicando en las mismas al paciente cuya probabilidad de reestenosis queremos estimar, de forma que pueda valorarse correctamente su riesgo de reestenosis. No es suficiente con utilizar la puntuación que devuelve el modelo *machine learning*, hay que poner esta puntuación en contexto respecto a la población y al problema concreto.

Para representar las curvas y situar el paciente analizado en las mismas se ha utilizado la librería Python Bokeh. Bokeh es una librería de visualización interactiva de Python pensada para ser utilizada en los navegadores web modernos. Su objetivo es proporcionar una construcción de gráficos sencillos, novedosos e interactivos.

**Tabla 3.** Tablas de contingencia con 4 categorías y métricas asociadas.

Estado respecto a la enfermedad		
Resultado de la prueba de estudio	Enfermo	Sano
Positivo	Verdadero positivo (VP)	Falso positivo (FP)
Negativo	Falso negativo (FN)	Verdadero negativo (VN)
$\text{Exhaustividad} = \frac{VP}{VP+FP} ; \text{Sensibilidad} = \frac{VP}{VP+FN}$ $\text{Especificidad} = \frac{VN}{VN+FP} ; \text{Valor predictivo negativo} = \frac{VN}{VN+FN}$		



## Capítulo 5

### 5. Resultados

#### 5.1. Características de la población de estudio

Las variables de datos continuas y categóricas de los pacientes utilizados para el desarrollo del modelo *machine learning* se muestran en las siguientes tablas (**tabla 4 y 5**, respectivamente). Como se reseñaba anteriormente, 23 pacientes (8,7%) presentaron reestenosis binaria angiográfica a los 12 meses de seguimiento del total de 236 pacientes incluidos en este análisis. En el análisis bivalente, los pacientes que experimentaron reestenosis tenían significativamente más probabilidades de tener las siguientes características: diabetes, colesterol anormal, enfermedad coronaria en 2 o más vasos, mayor número de stents implantados, menor diámetro luminal mínimo post-ICP, mayor porcentaje de diámetro de la estenosis post-ICP, mayor posibilidad de trombo post-ICP y menor flujo TIMI post-ICP.

Tabla 4. Diferencias en las variables de entrada continuas de la población

Variable (feature)	Toda la población	Reestenosis	No reestenosis	p-valor
N	263	23	240	
<i>Variables demográficas y clínicas</i>				
Edad, años	59,9±11,9	56,6±12,8	60,2±11,8	0,163
Peso, Kg	78,5±12,4	80,6±15,5	78,3±12,1	0,409
Altura, cm	168,2±8,3	167,3±8,9	168,2±8,3	0,628
Pulso, lpm	74,5±14,89	73,7±11,9	74,6±15,2	0,064
Presión arterial sistólica, mm Hg	135,6±22,5	138,6±18,7	135,3±22,8	0,496
Presión arterial diastólica, mm Hg	78,1±14,0	78,9±13,0	78,0±14,2	0,775
Número diario de cigarrillos fumados	11,6±16,4	6,4±9,7	12,2±16,8	0,110
Consumo diario del acohol, unidades	0,35±1,29	0,04±0,21	0,38±1,34	0,240
<i>Variables angiográficas</i>				
Diámetro luminal mínimo pre-ICP, mm	0,84±0,54	0,63±0,45	0,86±0,55	0,058
% diámetro de estenosis pre-ICP, %	71,92±16,44	78,04±14,66	71,35±16,51	0,068
% área de estenosis pre-ICP, %	89,42±11,34	93,13±6,73	89,07±11,62	0,109
Longitud de la lesión pre-ICP, mm	14,51±8,18	15,19±9,84	14,44±8,03	0,684
Diámetro luminal mínimo post-ICP, mm	2,36±0,59	2,05±0,67	2,38±0,58	0,025
% diámetro de estenosis post-ICP, %	23,00±10,27	29,31±20,67	22,53±8,95	0,008
% área de estenosis post-ICP, %	39,66±14,01	46,01±19,92	39,19±13,41	0,053
Ganancia inmediata post-ICP, mm	1,56±0,61	1,45±0,65	1,57±0,61	0,434
<i>Variables funcionamiento ventrículo izquierdo</i>				

Volumen telediastólico, mL	124,65±44,73	123,73±42,19	124,74±45,05	0,917
Volumen telesistólico, mL	56,38±28,02	51,32 ±8,09	56,87±28,78	0,365
Fracción de eyección, %	55,34±11,71	58,04±8,62	55,08±11,94	0,258
<b><i>Variables bioquímicas</i></b>				
Creatinina quinasa (CK), U/I	2281,9±1904,8	1771,1±1204,2	2325,4±1948,9	0,265
Creatinina quinasa MB (CK-MB), U/I	285,0±227,7	239,8±183,0	289,3±231,4	0,342
Colesterol total, mg/dL	182,7±37,8	190,57±45,74	181,93±37,02	0,320
Colesterol-LDL, mg/dL	106,5±31,6	110,3±38,9	106,1±30,8	0,573
Plaquetas, x10 <sup>3</sup> /μL	231,7±71,6	253,9±111,1	229,5±66,5	0,119
Leucocitos, x10 <sup>3</sup> /μL	11,33±8,81	10,77±3,59	11,39±9,16	0,748
Hemoglobina, g/dL	14,46±1,61	14,45±1,38	14,47±1,64	0,954
Hematocrito, %	42,56±4,53	42,38±3,81	42,58±4,6	0,840
Creatinina, mg/dL	0,98±0,27	0,9±0,14	0,99±0,27	0,115

ICP = intervencionismo coronario percutáneo

Tabla 5. Diferencias en las variables de entrada categóricas de la población

Variable (feature)	Toda la población	Reestenosis	No reestenosis	p-valor
N	263	23	240	
<b>Variables demográficas y clínicas</b>				
Sexo varón	219 (83,3)	19 (82,6)	200 (83,3)	0,837
Diabetes mellitus	51 (19,4)	12 (52,2)	39 (16,2)	<0,001
Diabetes mellitus tipo 1	12 (4,6)	2 (8,7)	10 (4,2)	0,655
Hipertensión arterial	108 (41,1)	11 (47,8)	97 (40,4)	0,641
Fumador				0,823
Fumador activo	134 (51)	12 (52,2)	122 (50,8)	
Exfumador	70 (26,6)	5 (21,7)	65 (27,1)	
Nunca ha fumado	59 (22,4)	6 (26,1)	53 (22,1)	
Dislipemia	126 (47,9)	11 (47,8)	115 (47,9)	0,833
H <sup>a</sup> familiar enfermedad cardiovascular	56 (21,3)	5 (21,7)	51 (21,3)	0,831
Algún consumo de alcohol	37 (14,2)	1 (4,3)	36 (15,0)	0,234
Angina previa	37 (14,2)	1 (4,3)	36 (15,0)	0,234
Infarto de miocardio previo	15 (5,7)	2 (8,7)	13 (5,4)	0,862
ICP previo	9 (3,4)	1 (4,3)	8 (3,3)	0,713
<b>Medicación</b>				
IECA	22 (8,4)	3 (13,0)	19 (7,9)	0,662
ARA II	6 (2,3)	0	6 (2,5)	0,971
Betabloqueantes	28 (10,6)	4 (17,4)	24 (10,0)	0,479
Calcioantagonistas	6 (2,3)	1 (4,3)	5 (2,1)	0,971
Nitroglicerina	231 (87,8)	21 (91,3)	210 (87,5)	0,840
Aspirina	243 (92,4)	19 (82,6)	224 (93,3)	0,194
Clopidogrel	20 (7,6)	3 (13,0)	17 (7,1)	0,557

<b>Variables angiográficas</b>				
Arteria coronaria enferma				0,429
Descendente anterior (LAD)	106 (40,9)	7 (31,8)	99 (41,8)	
Circunfleja (LCX)	40 (15,4)	2 (9,1)	38 (16,0)	
Coronaria derecha (RCA)	113 (43,7)	13 (59,1)	100 (42,2)	
Enfermedad en $\geq 2$ vasos	90 (34,2)	15 (65,2)	75 (31,3)	0,003
Enfermedad en $\geq 3$ vasos	21 (8,0)	3 (13,0)	18 (7,5)	0,609
Morfología de la lesión				0,139
Tipo A	19 (7,2)	0	19 (7,9)	
Tipo B1	124 (47,2)	9 (39,1)	115 (47,9)	
Tipo B2	99 (37,6)	11 (47,8)	88 (36,6)	
Tipo C	10 (3,8)	2 (8,6)	8 (3,3)	
No disponible	11 (4,2)	1 (4,3)	10 (4,2)	
Stent farmacoactivo	137 (52,1)	12 (52,2)	125 (52,1)	0,834
Número de stents implantados en el segmento infartado				0,018
0	1 (0,4)	1 (5,6)	0 (0)	
1	206 (84,4)	11 (61,1)	195 (86,3)	
2	33 (13,5)	5 (27,7)	28 (12,4)	
$\geq 3$	4 (1,7)	1 (5,6)	3 (1,3)	
Número de stents en segmentos no infartados de la arteria infartada				0,122
0	230 (87,4)	19 (82,6)	211 (87,9)	
1	28 (10,7)	3 (13,0)	25 (10,4)	
2	4 (1,5)	0	4 (1,7)	
$\geq 3$	1 (0,4)	1 (4,4)	0	
Número de stents en arterias no relacionadas con el infarto				0,523
0	211 (80,2)	19 (82,6)	192 (80,0)	

1	47 (17,9)	3 (13,0)	44 (18,4)	
2	2 (0,8)	0	2 (0,8)	
≥ 3	3 (1,1)	1 (4,4)	2 (0,8)	
Uso de tirofiban	131 (49,8)	11 (47,8)	120 (50,0)	0,985
Resultado satisfactorio en la angioplastia	231 (93,9)	14 (82,3)	217 (94,8)	0,182
Fenómeno de no reflujo	15 (6,1)	2 (11,8)	13 (5,7)	0,645
Trombo TIMI pre-ICP				0,862
0	56 (21,6)	3 (13,6)	53 (22,4)	
1	91 (35,1)	9 (40,9)	82 (34,6)	
2	31 (12,0)	2 (9,1)	29 (12,2)	
3	26 (10,0)	2 (9,1)	24 (10,1)	
4	24 (9,3)	2 (9,1)	22 (9,3)	
5	31 (12,0)	4 (18,2)	27 (11,4)	
Trombo TIMI post-ICP				0,016
0	212 (86,2)	13 (76,5)	199 (86,9)	
1	22 (9,0)	1 (5,9)	21 (9,1)	
2	2 (0,8)	0	2 (0,9)	
3	5 (2,0)	0	5 (2,2)	
4	3 (1,2)	1 (5,9)	2 (0,9)	
5	2 (0,8)	2 (11,7)	0	
Flujo epicárdico TIMI pre-ICP				0,721
0	28 (10,6)	4 (17,4)	24 (10,0)	
1	14 (5,3)	1 (4,4)	13 (5,4)	
2	28 (10,6)	3 (13,0)	25 (10,4)	
3	193 (73,4)	15 (65,2)	178 (74,2)	
Flujo epicárdico TIMI post-ICP				0,002
0	3 (1,2)	2 (10,5)	1 (0,4)	
1	1 (0,4)	1 (5,3)	0	

2	10 (4,0)	2 (10,5)	8 (3,5)	
3	236 (94,4)	14 (73,7)	222 (96,1)	
Flujo microvascular TMPG pre-ICP				0,991
0	68 (27,9)	5 (26,3)	63 (28,0)	
1	12 (4,9)	1 (5,3)	11 (4,9)	
2	30 (12,3)	2 (10,5)	28 (12,4)	
3	134 (54,9)	11 (57,9)	123 (54,7)	
Flujo microvascular TMPG post-ICP				0,179
0	24 (10,5)	4 (25,0)	20 (9,4)	
1	26 (11,3)	1 (6,3)	25 (11,7)	
2	36 (15,7)	4 (25,0)	32 (15,0)	
3	143 (62,4)	7 (43,7)	136 (63,9)	
<b>Variables bioquímicas</b>				
Colesterol total anormal	19 (7,2)	6 (26,1)	13 (5,4)	0,007
Colesterol-LDL anormal	11 (4,2)	3 (13,0)	8 (3,3)	0,150
Plaquetas anormales	1 (0,4)	1 (4,3)	0	0,284
Leucocitos anormales	19 (7,2)	3 (13,0)	16 (6,7)	0,505
Hemoglobina anormal	9 (3,4)	0	9 (3,8)	0,714
Hematocrito anormal	5 (1,9)	0	5 (2,1)	0,922
Creatinina anormal	7 (2,7)	0	7 (2,9)	0,876

ARA II = antagonistas de los receptores de la angiotensina II; IECA = inhibidores enzima convertidora de la angiotensina;  
ICP = intervencionismo coronario percutáneo;

## 5.2. Comparación de los modelos *machine learning* con los discriminadores predictivos inferenciales desarrollados con modelos estadísticos clásicos

En la **figura 10** se muestran las curvas ROC para los 6 modelos *machine learning* entrenados junto con la selección de variables utilizando ANOVA. La selección con ANOVA dio mejores resultados que la selección de variables usando la importancia de variables (*feature importance*) que se basa en *random forest*.

El mejor modelo *machine learning* final con un área bajo la curva ROC de 0,77 (95%: IC, 0,66-0,89) se obtuvo con el clasificador *extremely randomized trees*. No se obtuvieron diferencias significativas entre *extremely randomized trees* y el resto de los modelos *machine learning*, excepto con el clasificador *support vector machine* ( $P=0,003$ ).

La predicción del modelo desarrollado con *extremely randomized trees* logró una mejora estadísticamente significativa en comparación con los discriminadores predictivos inferenciales desarrollados con modelos estadísticos clásicos ( $p < 0,001$ ; **Figura 11**); PRESTO-1 (AUC-ROC 0,57 [0,43-0,71]), PRESTO-2 (AUC-ROC 0,58 [0,43-0,72]) y TLR (AUC-ROC 0,62 [0,50-0,74]).

Según lo ya descrito en la parte metodológica de este estudio, reportamos los resultados de las curvas ROC de los scores de riesgo PRESTO y TLR basados en sus coeficientes exactos y enteros (**Figura 11**). Como se puede observar, en términos de área bajo la curva ROC, ambos métodos muestran resultados similares. Sin embargo, y como era de esperar, los scores basados en coeficientes redondeados muestran una estratificación en un menor número de grupos, al tomar los scores de riesgo únicamente 6 posibles valores. Esto no es necesariamente negativo, ya que, dependiendo de la zona de la curva, la clasificación aleatoria de pacientes junto a una infra o sobrestimación de coeficientes puede ser más exacta que la generada por el modelo de coeficientes exactos. A la hora de comparar el resto de las métricas evaluadas en este estudio, se han considerado únicamente los scores de riesgo basados en coeficientes redondeados, por mostrar éstos, en términos de área bajo la curva ROC, un comportamiento ligeramente superior.



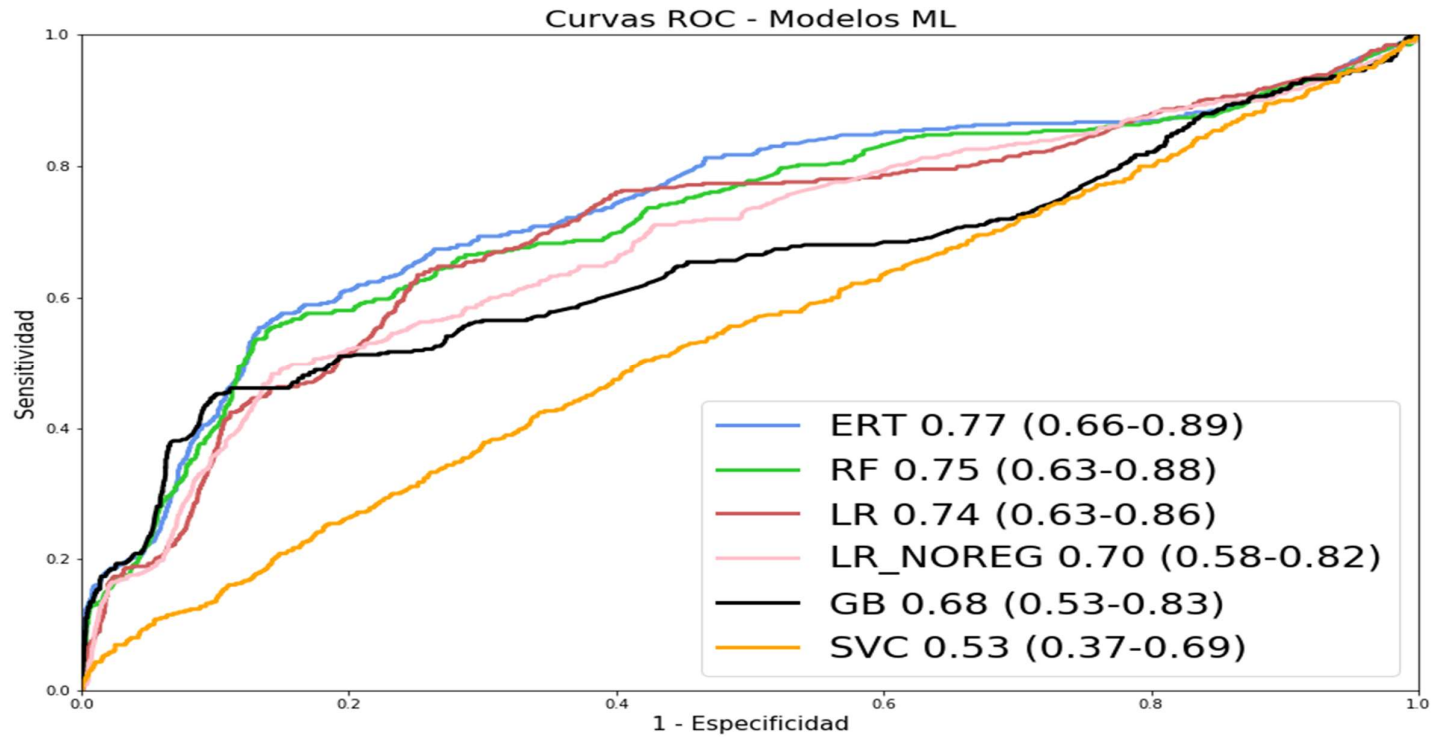
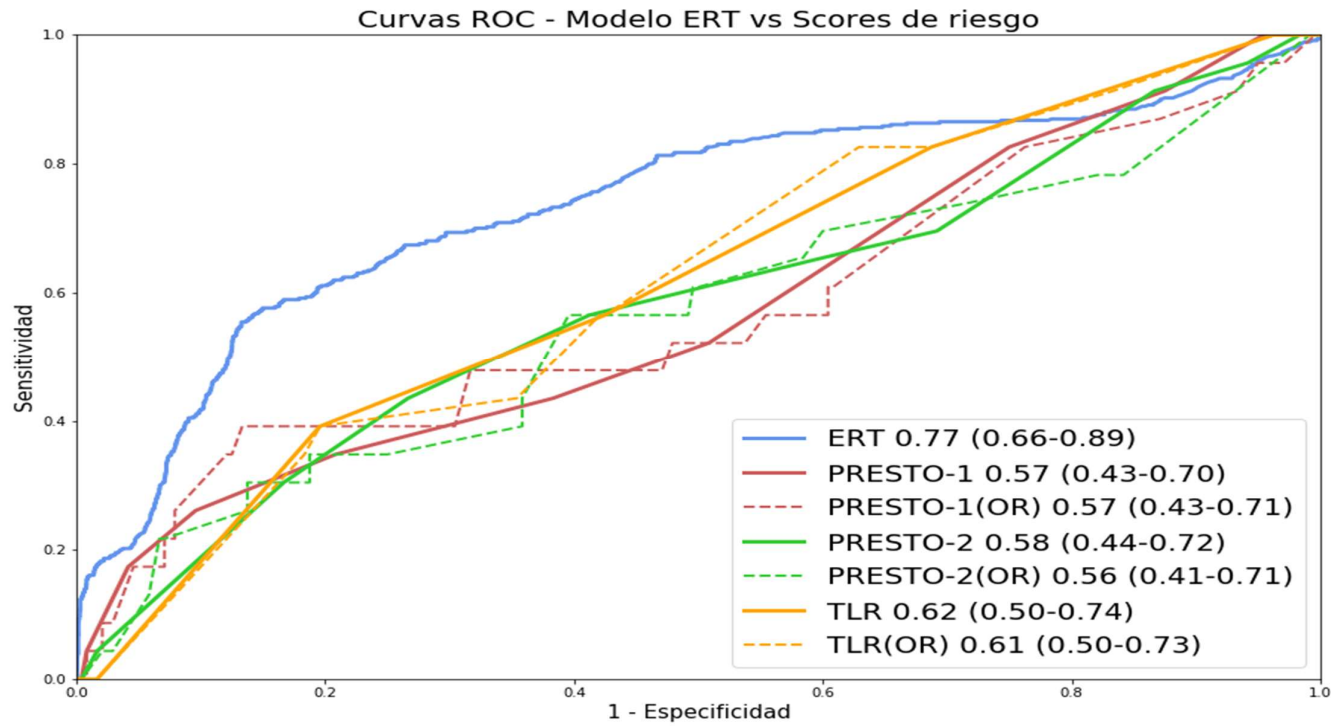


Figura 10. Curvas ROC para los modelos *machine learning* entrenados.

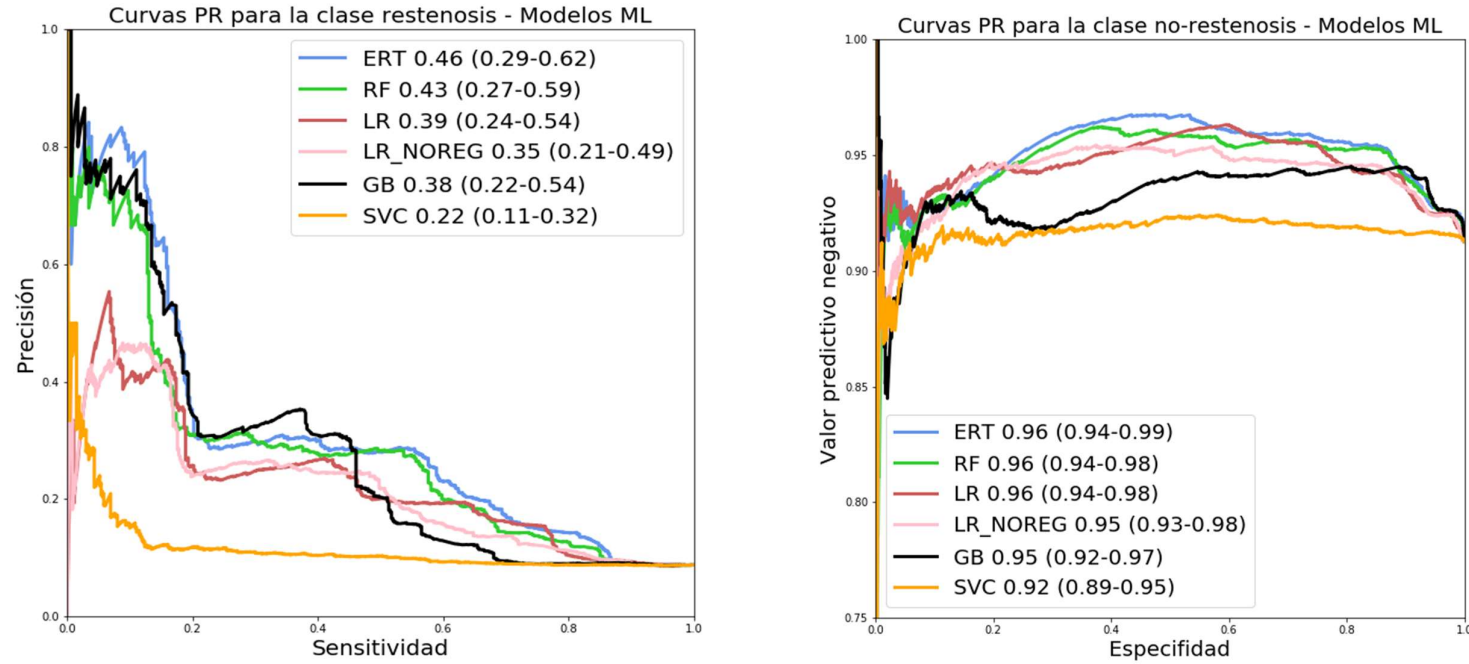
ERT: *extremely randomized trees*; RF: *random forest*; LR: *regresión logística con regularización L2*; LR\_NOREG: *regresión logística sin regularización*; GB: *gradient boosting*; SVC: *support vector machine classifier*.



**Figura 11.** Curvas ROC comparando el mejor modelo *machine learning* desarrollado con los *scores* de riesgo clásicos.

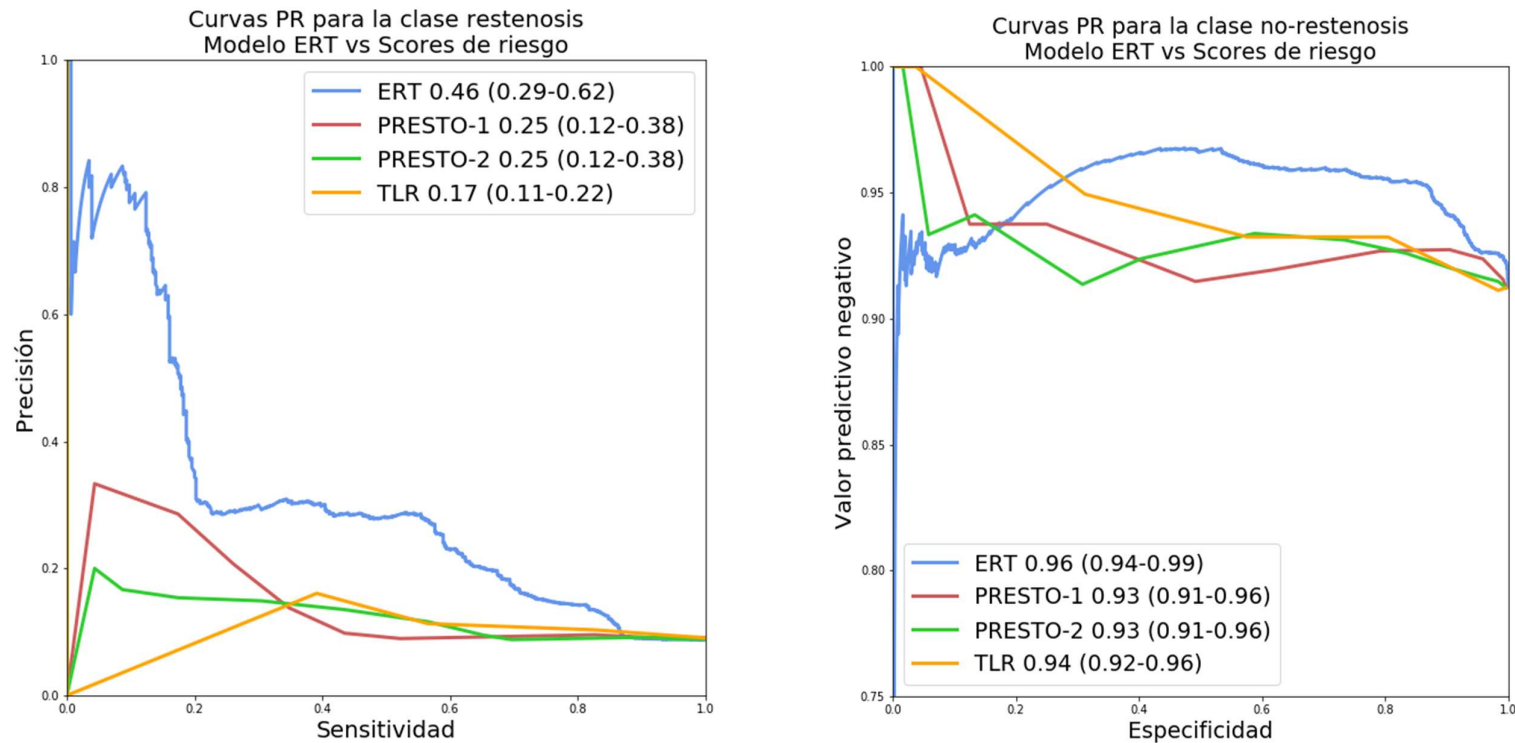
Los *scores* de riesgo con el sufijo (OR) se calcularon con los *odds-ratios* exactos de cada modelo. ERT: *extremely randomized trees*

Las curvas PR nos mostraron una imagen más precisa del rendimiento de los modelos. Estas curvas calculadas sobre los mismos modelos entrenados resaltan la diferencia visual entre los dos espacios de decisión binaria: reestenosis frente a no reestenosis. Las curvas PR para la predicción de reestenosis mostraron una exhaustividad y una sensibilidad por debajo de 0,5; sin embargo, para la predicción de no reestenosis presenta una elevada área bajo la curva PR, presentando tanto alta exhaustividad como alta sensibilidad, donde la alta exhaustividad se relaciona con una baja tasa de falsos positivos y una alta sensibilidad se relaciona con una baja tasa de falsos negativos (**Figura 12**). Esta diferencia existe porque en este espacio de decisión el número de casos negativos excede en gran medida el número de casos positivos; y proporciona a nuestro algoritmo ERT un excelente valor predictivo negativo. Asimismo el área bajo la curva PR para el modelo *extremely randomized trees* desarrollado fue 0,96 (0,96-0,99) con una ventaja significativa ( $P=0,003$ ) sobre los scores PRESTO-1, PRESTO-2 y TLR (**Figura 13**).



**Figura 12.** Curvas PR (*Precision-Recall*) comparando la predicción de reestenosis (*precision*) y no reestenosis (*valor predictivo negativo*) de los distintos modelos *machine learning* entrenados.

**ERT:** *extremely randomized trees*; **RF:** *random forest*; **LR:** regresión logística con regularización L2; **LR\_NOREG:** regresión logística sin regularización; **GB:** *gradient boosting*; **SVC:** *support vector machine classifier*.



**Figura 13.** Curvas PR (*Precision-Recall*) comparando la predicción de reestenosis (*precision*) y no reestenosis (*valor predictivo negativo*) del modelo *machine learning* y los scores de riesgo clásicos.

**ERT:** *extremely randomized trees*.

### 5.3. Interpretación de las curvas ROC y PR en diferentes escenarios de aplicación clínica

Para evaluar mejor la importancia clínica de nuestros resultados, comparamos el rendimiento de de nuestro modelo *extremely randomized trees* con el mejor score de riesgo clínico (TLR) en tres umbrales de aplicación concretos (**Tabla 6**), identificados en las curvas ROC y PR en las **Figuras 14 y 15**. Esto es un requisito crítico para un *score* predictivo, donde dependiendo del caso debemos minimizar el número de falsos negativos o falsos positivos.

Para un umbral considerado óptimo para alta exhaustividad (primer umbral evaluado), de 1.000 pacientes con infarto de miocardio sometidos a implante de *stent*, nuestro modelo *extremely randomized trees* podría predecir correctamente 181 (+18%) casos más en comparación con el *score* TLR. Así es capaz de lograr un 89,6% de verdaderos negativos (71,5% para el *score* TLR) y detecta un grupo de 32 pacientes con un 50% de riesgo de reestenosis que deben ser programados rutinariamente para un seguimiento más cercano.

Por el contrario, para un umbral óptimo de valor predictivo negativo (tercer umbral evaluado), nuestro modelo *machine learning* puede detectar un grupo de 504 pacientes con un 3% de riesgo de reestenosis evitando cargas de trabajo de seguimiento innecesarias y reduciendo los costes (TLR discrimina 318 pacientes y 5% de riesgo).

Estos tres puntos de corte (umbrales) nos definen diferentes zonas de riesgo. Con ellos podemos estratificar el riesgo de los pacientes. La zona de máxima exhaustividad, definida por el primer umbral y situada en la curva ROC a la izquierda de este umbral, es una zona donde el riesgo de sufrir reestenosis es muy alto en comparación con su prevalencia (>50% vs 9%). A cualquier paciente que se nos sitúe en esta zona, que serán pocos (3%), hay que realizarle un seguimiento más estrecho y de mayor coste porque su probabilidad de reestenosis es mucho más alta que en la población. El mejor *score* clásico es muy inferior en esta zona de máximo riesgo, como puede verse en la **Tabla 6** y en las **Figuras 14 y 15** (relación de falsos positivos frente a verdaderos positivos para TLR).

La zona de máximo valor predictivo negativo (desde el tercer umbral evaluado a la derecha de la curva ROC es una zona donde el riesgo de reestenosis se reduce a la tercera parte, baja de un 9% a <3%. Esta zona delimita un gran grupo de pacientes (más del 50%) donde podemos reducir el coste del seguimiento.

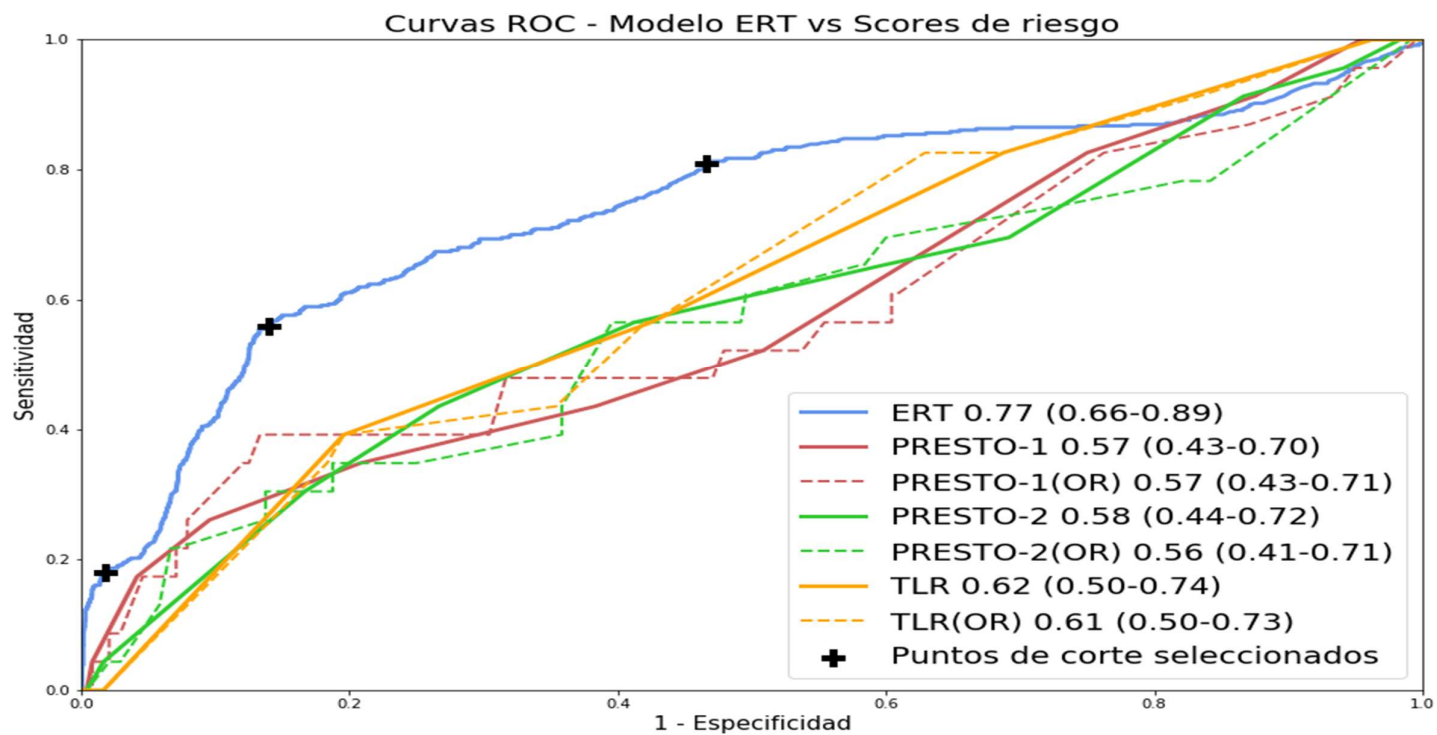
En las **Figuras 16 y 17** se representan las zonas de riesgo definidas por los umbrales. En la **Figura 16** se representa un paciente alto riesgo y en la **Figura 17** un paciente de bajo riesgo.

**Tabla 6.** Valores para ERT y TRL en tres umbrales diferentes de aplicación clínica

Valores de verdaderos positivos (VP), verdaderos negativos (VN), falsos positivos (FP) y falsos negativos (FN), sensibilidad, especificidad, exhaustividad y valor predictivo negativo (VPN) en tres umbrales diferentes para el modelo ERT y el score de riesgo clásico TLR. Para una mejor comprensión, los valores se han escalado para una población de 1.000 pacientes con la misma prevalencia de reestenosis del stent que la población del ensayo GRACIA-3.

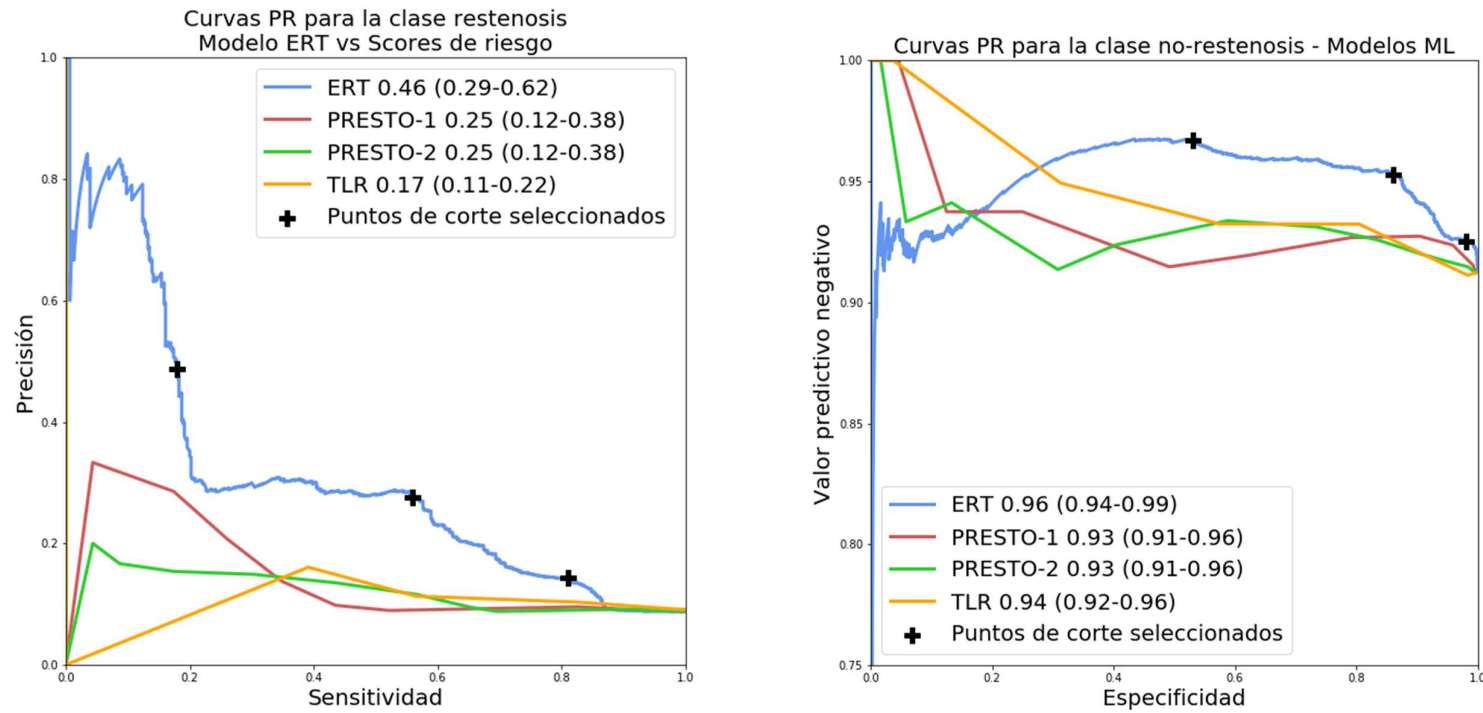
Modelo - umbral	Sensibilidad	Especificidad	Exhaustividad	VPN	VP	VN	FP	FN
ERT - 1	0,18	0,98	0,49	0,93	16	896	16	72
TLR - 1	0,18	0,78	0,07	0,91	16	715	197	72
ERT - 2	0,56	0,86	0,28	0,95	49	784	128	38
TLR - 2	0,56	0,58	0,11	0,93	49	533	379	39
ERT - 3	0,81	0,53	0,14	0,97	70	487	426	17
TLR - 3	0,81	0,33	0,10	0,95	70	301	611	17





**Figura 14.** Curvas ROC comparando el mejor modelo *machine learning* desarrollado con los scores de riesgo clásicos en 3 umbrales de uso clínico.

ERT: *extremely randomized trees*.



**Figura 15.** Curvas PR (*Precision-Recall*) comparando la predicción de reestenosis (*precision*) y no reestenosis (*valor predictivo negativo*) del modelo *machine learning* y los *scores* de riesgo clásicos en tres umbrales de uso clínico.

**ERT:** *extremely randomized trees*.

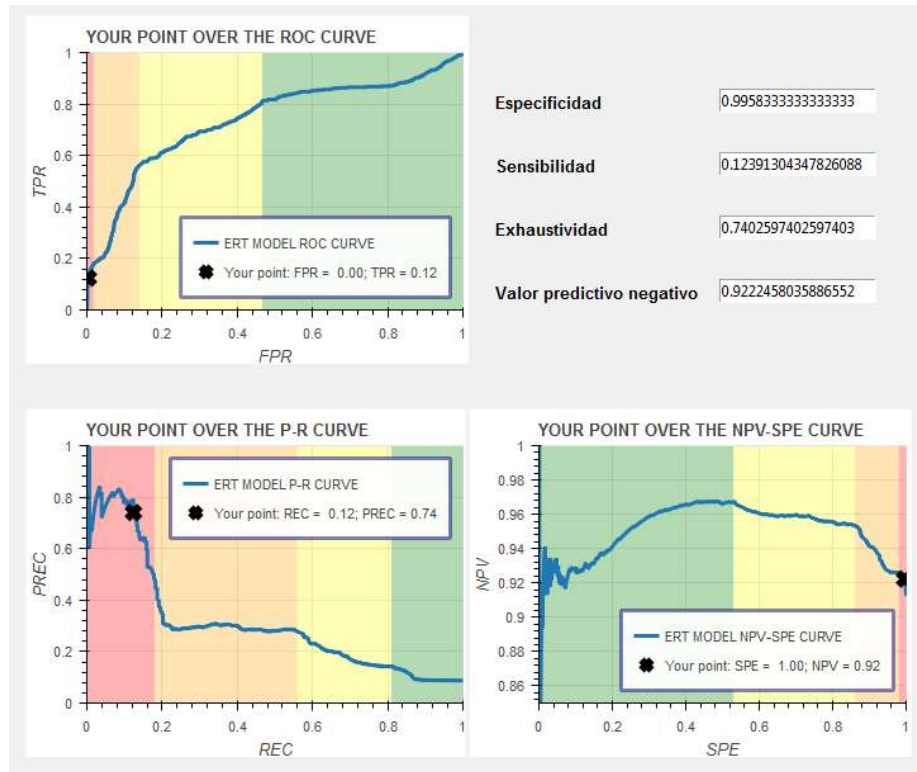


Figura 16. Zonas de riesgo definidas por los umbrales. Paciente de alto riesgo.

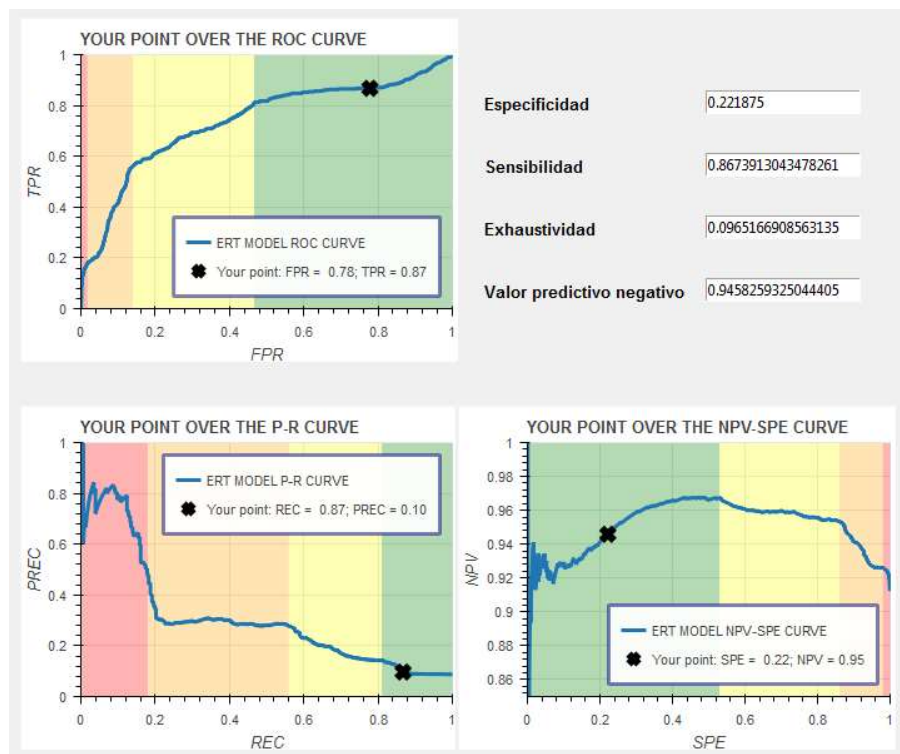


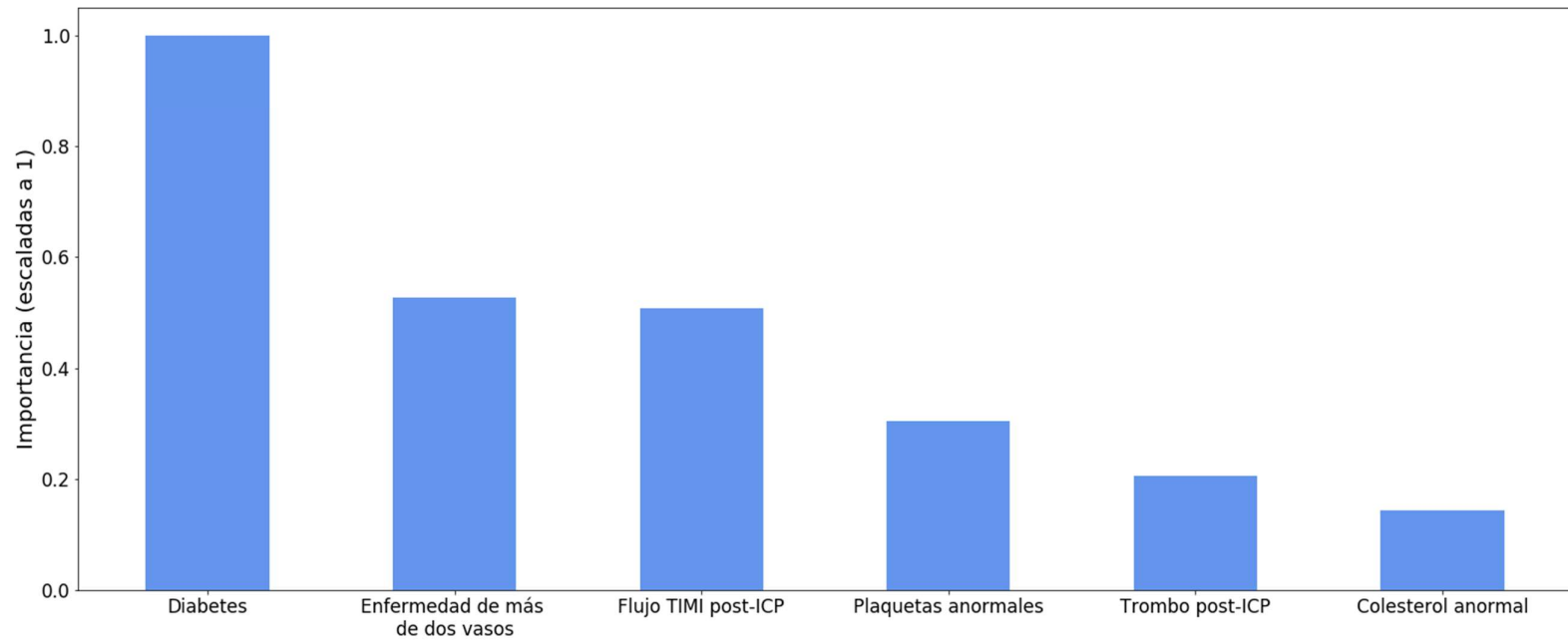
Figura 17. Zonas de riesgo definidas por los umbrales. Paciente de bajo riesgo.

## 5.4. Variables con mayor peso en la predicción

La **Figura 18** muestra la importancia de las primeras 6 variables predictoras en el modelo ERT: diabetes, enfermedad en dos o más vasos coronarios, flujo TIMI post-ICP, plaquetas anormales, trombo post-ICP y colesterol total anormal.

La diabetes, solo incluida en el score PRESTO-2, parece ser más importante que otros factores o hábitos de riesgo cardiovascular como fumar o consumir alcohol para predecir reestenosis. Entre las variables angiográficas, el flujo TIMI post-ICP fue la variable más importante de los principales predictores. Otras características que no se muestran en la figura presentan valores inferiores a 0,01 (escalados a la importancia de la variable principal) o valores negativos (variables que tienen una contribución negativa a la tarea de clasificación).

Esta selección de las variables más importantes en la predicción es diferente de la selección de variables utilizada durante la construcción del modelo. Durante la construcción del modelo, el número de variables, la selección de las mismas y el resto de los hiperparámetros del modelo se eligen exclusivamente con los datos del conjunto de entrenamiento en el esquema de validación cruzada anidado descrito. Si durante la construcción del modelo utilizamos todo el conjunto de datos para la selección de variables estaríamos filtrando información del conjunto de test al modelo, incurriendo en un sobreajuste y, por lo tanto, la evaluación del modelo sería demasiado optimista de forma errónea, con problemas de generalización posterior a otras poblaciones. Este tipo de error de cálculo no es infrecuente en la aplicación de la validación cruzada y, por ejemplo, en este estudio, el área bajo la curva ROC (AUC-ROC) para el modelo *extremely randomized trees* habría alcanzado un valor de 0,82 en vez del 0,77 indicado.



**Figura 18.** Gráfico de barras con la importancia de las variables. Importancia de cada una de las variables en la predicción de una reestenosis del *stent* en el seguimiento a 12 meses con un clasificador *extremely randomized trees classifier* (ERT).

Las variables más importantes fueron diabetes, enfermedad en 2 o más vasos coronarios, flujo TIMI post-ICP.

## 5.5. Aplicando el modelo a la práctica clínica. Propuesta de calculadora de riesgo de reestenosis del *stent* tras ICP

Utilizando el modelo *machine learning* desarrollado y entrenado, se ha construido una calculadora de riesgo de reestenosis del *stent* tras ICP.

Esta calculadora se ha desarrollado para 2 entornos:

- Aplicación web (Google Colab). En la nube, ejecutable con cualquier navegador.
- Aplicación para ejecutar en cualquier PC (Windows, MAC) con Python (Tkinter).

### 5.5.1. Calculadora con Google Colab

Se puede acceder a la calculadora desarrollada a través de la URL [https://colab.research.google.com/drive/1jwGpC1rhkVKZi3sWdc-yKm2VxVsgMFUM?authuser=1#scrollTo=Gw9KRvAON9\\_t](https://colab.research.google.com/drive/1jwGpC1rhkVKZi3sWdc-yKm2VxVsgMFUM?authuser=1#scrollTo=Gw9KRvAON9_t); o a través del siguiente enlace [Google Colab Calculadora Reestenosis](#).

La **Figura 19** muestra el aspecto de la calculadora tal y como aparece en la URL.

```

Notebook to run the model developed for the study:
'MACHINE LEARNING TO PREDICT STENT RESTENOSIS BASED ON DAILY DEMOGRAPHIC, CLINICAL AND ANGIOGRAPHIC CHARACTERISTICS'

In order to run the notebook:
• Write the inputs in the next cell. In case one of the features is not available, keep the values that are already written.
• Click on "Runtime -> Run all". Wait the notebook to run all cells. Check the results at the end of it.

[] # Data inputs
diabetes = 1
vessels_gtz = 0
thrombus_post_pci = 0
tmi_post_pci = 0
abnormal_cholesterol = 0
abnormal_platelets = 0

[]
import os
if os.path.exists('./finalized_model.sav') == False:
    !wget -nv https://raw.githubusercontent.com/IA-Cardiologia-Husa/Restenosis/master/models/finalized_model.sav

if os.path.exists('./probs.csv') == False:
    !wget -nv https://raw.githubusercontent.com/IA-Cardiologia-Husa/Restenosis/master/models/probs.csv

if os.path.exists('./Funcs_boksh.py') == False:
    !wget -nv https://raw.githubusercontent.com/IA-Cardiologia-Husa/Restenosis/master/calculator/Funcs_boksh.py

2019-10-31 08:44:38 URL: https://raw.githubusercontent.com/IA-Cardiologia-Husa/Restenosis/master/models/finalized_model.sav [348936/948636] -> "finalized_model.sav"
2019-10-31 08:44:31 URL: https://raw.githubusercontent.com/IA-Cardiologia-Husa/Restenosis/master/models/probs.csv [115588/115588] -> "probs.csv" []
2019-10-31 08:44:34 URL: https://raw.githubusercontent.com/IA-Cardiologia-Husa/Restenosis/master/calculator/Funcs_boksh.py [3469/3469] -> "Funcs_boksh.py" []

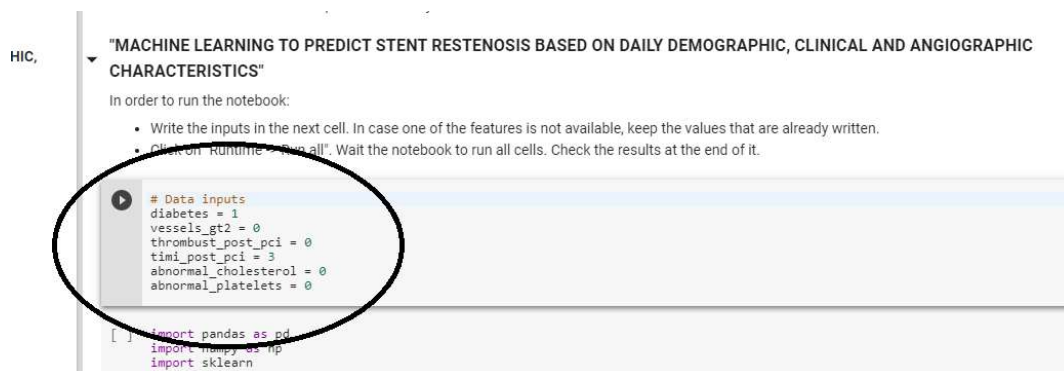
[]
from Funcs_boksh import *
model = pickle.load(open('finalized_model.sav', 'rb'))
model = model.named_steps['clf']
X = np.array([diabetes, vessels_gtz, thrombus_post_pci, tmi_post_pci,
              abnormal_cholesterol, abnormal_platelets]).reshape(1,-1)

```

**Figura 19.** Calculadora de riesgo de reestenosis tras angioplastia desarrollada en Google Colab (Google Colab Calculadora Reestenosis)

Esta forma de trabajar es visualmente poco atractiva, pero nos permite compartir y ejecutar los modelos en entornos web colaborativos que no consumen apenas recursos.

Lo primero que aparece es la entrada de variables. En esta zona reportaremos la información de las variables principales que utiliza el modelo (**Figura 20**).



```

# Data inputs
diabetes = 1
vessels_gt2 = 0
thrombus_post_pci = 0
timi_post_pci = 3
abnormal_cholesterol = 0
abnormal_platelets = 0

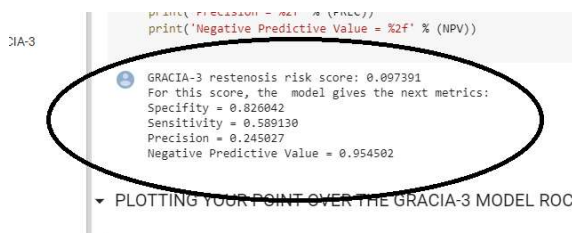
import pandas as pd
import numpy as np
import sklearn
import pickle

```

**Figura 20.** Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. Entrada de variables.

Las variables las introduciremos con “1” para Sí y con “0” para No. Todas son de este tipo salvo trombo\_post\_PCI, que puede tomar valores de 0, 1, 2, 3, 4 ó 5 y TIMI\_post\_PCI, que puede tomar valores 1, 2 ó 3.

Una vez introducidas las variables ejecutamos el modelo en el menú “Entorno de ejecución” seleccionando “Ejecutar todas”. El programa se descargará de GitHub el modelo desarrollado y lo ejecutará con las variables introducidas, devolviendo como resultado el *score* de reestenosis y las distintas métricas de especificidad (*specificity*), sensibilidad (*recall*), exhaustividad (*precision*) y valor predictivo negativo (*negative predictive value*) (**Figura 21**).



```

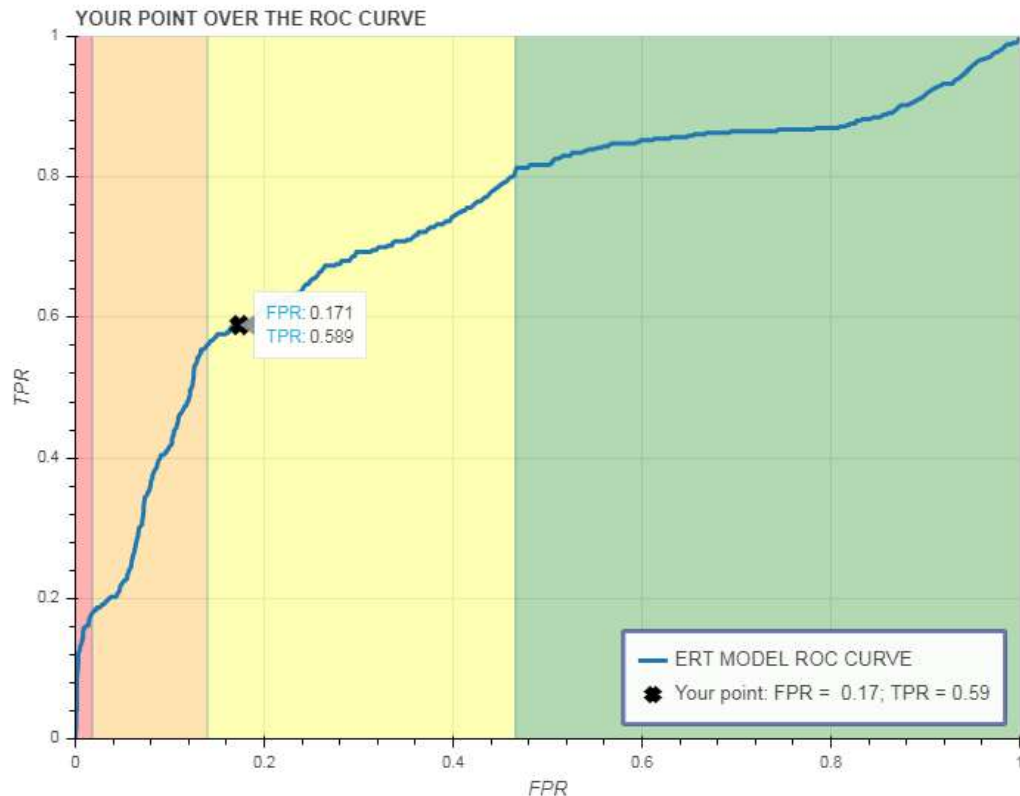
print('Negative Predictive Value = %2f' % (NPV))

GRACIA-3 restenosis risk score: 0.097391
For this score, the model gives the next metrics:
Specificity = 0.826042
Sensitivity = 0.589130
Precision = 0.245027
Negative Predictive Value = 0.954502

```

**Figura 21.** Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. *Score* y métricas calculados por el modelo.

El *score* que devuelve el modelo no es un resultado definitivo para decidir qué hacer con ese paciente. La probabilidad de reestenosis es baja para este tipo de pacientes y lo interesante es ver en qué punto se sitúa este paciente respecto al conjunto, ver en qué punto se nos sitúa el paciente en las curvas ROC o PR, y por tanto en qué zona de riesgo respecto a las que hemos definido (**Figura 22**).

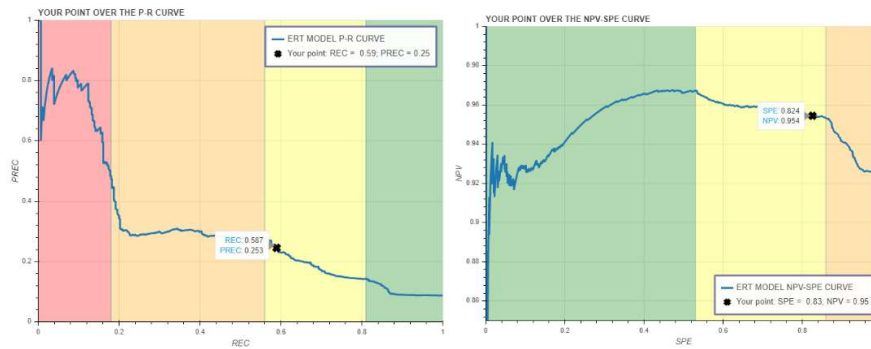


**Figura 22.** Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. Estratificación del riesgo del paciente basado en curvas ROC.

Con el paciente del ejemplo podemos ver que se sitúa en la zona de riesgo intermedio. Su *score* es bajo, porque la probabilidad de reestenosis lo es, pero es un paciente de riesgo intermedio. Se actuaría del mismo modo a como se viene actuando hasta la fecha, sin incrementar ni reducir su seguimiento. Los pacientes de riesgo alto (zonas naranja y roja) deben tener un seguimiento más cercano para evitar eventos no deseados. Los pacientes de riesgo bajo (zona verde) pueden tener un seguimiento mínimo para incrementar la eficiencia/coste en el seguimiento tras la intervención.

Podemos representar al paciente igualmente en las curvas PR (*precisión-recall*) tanto para reestenosis como para no-reestenosis (**Figura 23**).



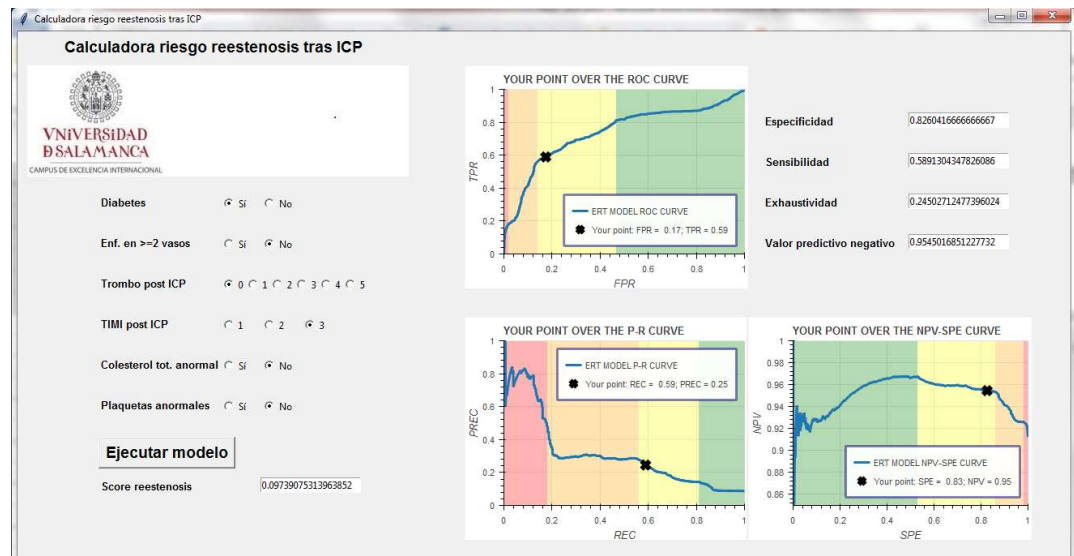


**Figura 23.** Calculadora de riesgo de reestenosis tras ICP desarrollada en Google Colab. Estratificación del riesgo del paciente basado en curvas PR para reestenosis y no reestenosis.

### 5.5.2. Calculadora con Tkinter

Esta solución se ha desarrollado completa en Python utilizando como añadido una librería para interfaces gráficas: Tkinter.

Esta calculadora ya tiene un aspecto visual mejorado y es más sencilla para su manejo por un profesional médico. Su aspecto es el siguiente (**Figura 24**).

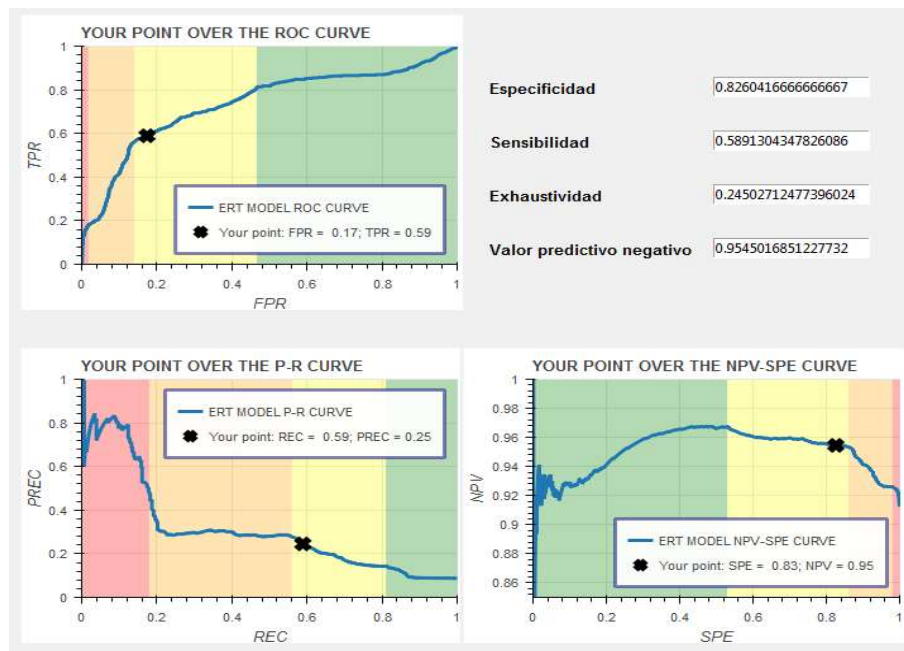


**Figura 24.** Calculadora de riesgo de reestenosis tras ICP desarrollada con Tkinter.

Para usarla es suficiente con seleccionar los valores de cada variable y pulsar el botón “Ejecutar el modelo”, lo cual devuelve el *score* de reestenosis, las gráficas ROC y

PR y las métricas de especificidad, sensibilidad, exhaustividad y valor predictivo negativo para ese paciente.

Al igual que con la calculadora desarrollada en Google Colab, el score no es suficiente para tomar una decisión. La probabilidad de reestenosis es baja para este tipo de pacientes y lo interesante es ver en qué punto se sitúa este paciente respecto al conjunto, ver en qué punto se nos sitúa el paciente en las curvas ROC o PR, y por tanto en qué zona de riesgo respecto a las que hemos definido. Se ha incluido en esta calculadora la misma funcionalidad que en la anterior, situando al paciente en la zona de riesgo para que el médico pueda en base a esta estratificación de riesgo tomar una decisión (**Figura 25**).



**Figura 25.** Calculadora de riesgo de reestenosis tras ICP desarrollada en Tkinter. Estratificación del riesgo del paciente basado en curvas ROC y curvas PR para reestenosis y no reestenosis.

Este desarrollo está también disponible en abierto en GitHub en la siguiente URL: [https://github.com/PedroDorado/Restenosis\\_IA\\_Gracia3/tree/master/calculator](https://github.com/PedroDorado/Restenosis_IA_Gracia3/tree/master/calculator); o a través del siguiente enlace [Tkinter Calculadora Reestenosis](#).

Para usarlo es suficiente con descargarse el programa y ejecutarlo en cualquier ordenador que tenga instalado el soporte para Python.

## Capítulo 6

### 6. Discusión

Usando *machine learning*, hemos desarrollado un modelo para predecir reestenosis coronaria tras ICP basado en variables obtenidas en la práctica clínica rutinaria a partir de un análisis retrospectivo del ensayo GRACIA-3. En comparación con scores clínicos clásicos de riesgo basados en regresión logística multivariante PRESTO-1, PRESTO-2 y TLR; nuestro modelo mejora la predicción en un 18% en términos de número de casos totales correctamente pronosticados.

#### 6.1. Reestenosis coronaria

Actualmente el ICP es el procedimiento más común para tratar las obstrucciones ateroscleróticas coronarias. El número total de ICPs realizadas en España en 2017 (últimos datos disponibles), con aproximadamente 45 millones de habitantes, fue de 70.928 procedimientos, con una proporción de 1.524 por millón de habitantes. Se implantaron 105.529 *stents* (el 90,3% farmacológicos). Además, 21.395 ICPs se llevaron a cabo durante la fase aguda del infarto de miocardio (30,2%)<sup>55</sup>

El tendón de aquiles del ICP es la reestenosis, una enfermedad iatrogénica, que conduce a un estrechamiento arterial recurrente en el lugar de la intervención. Los criterios que definen la reestenosis coronaria son angiográficos, existiendo un gran número de ellos; pero sin duda el más aceptado, es el estrechamiento superior al 50% del diámetro de la luz en el segmento objetivo; definido como todas las porciones del vaso coronario que recibieron tratamiento dentro de la zona del *stent*, incluidos los márgenes proximal y distal de 5 mm<sup>26</sup>.

La reestenosis es causada por una enfermedad inflamatoria crónica que produce la activación de las células inmunes y el crecimiento excesivo de las células vasculares del músculo liso dentro de la pared del vaso lesionado. Cuando se empezó a hacer ICP, a fines de la década de 1970, la ICP se limitaba a inflados de balón en la zona afectada.

Sin embargo, y aunque la arteria inicialmente se abría con éxito usando un balón, aproximadamente el 30% de todas las arterias coronarias comenzaban a cerrarse algunos meses después de la angioplastia realizada con balón debido a la reestenosis. A mediados de la década de 1980, los profesionales comenzaron a diseñar nuevos dispositivos para incrementar la durabilidad de los procedimientos de angioplastia, evitando la reestenosis. Uno de esos dispositivos fue el *stent*, un tubo de metal que se inserta después del inflado del balón, evitando así el remodelado negativo del vaso tratado, que era la principal causa de reestenosis después de la angioplastia convencional con balón. Pero, si bien los *stents* metálicos eliminaron muchos cierres arteriales abruptos, la reestenosis persistió en aproximadamente el 20% de los casos<sup>56</sup>. En la primera parte de este siglo, la solución para la reestenosis se alejó del *stent* metálico convencional hacia los *stents* recubiertos de productos farmacéuticos. Los profesionales comenzaron a probar una variedad de medicamentos que se sabe que interrumpen los procesos biológicos que causan la reestenosis. Los *stents* se recubrieron con fármacos inmunosupresores y antiproliferativos, comenzaron los ensayos clínicos y comenzó la era de los *stents* liberadores de fármacos, también conocidos como *stents* farmacológicos. Actualmente, estamos en la segunda generación de *stents* liberadores de fármacos y se ha logrado que la reestenosis se haya reducido al 10%<sup>57</sup> (Figura 26)<sup>57</sup>.

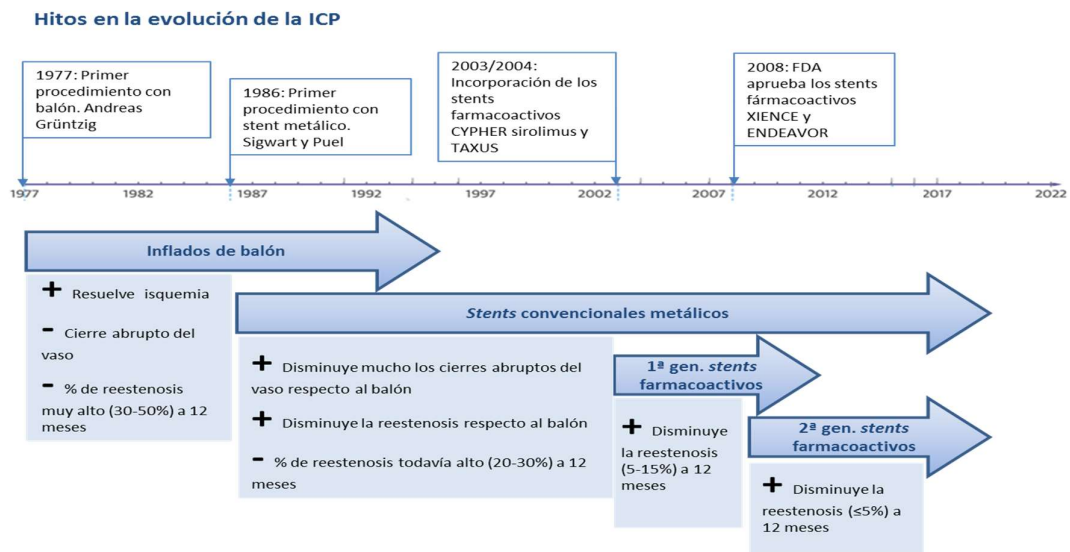


Figura 26. Hitos en la evolución de la ICP <sup>57</sup>.

Gráfico obtenido de Torrado J, Buckley L, Durán A, et al. Restenosis, Stent Thrombosis, and Bleeding Complications: Navigating Between Scylla and Charybdis. *J Am Coll Cardiol.* 2018;71(15):1676-1695

Sin embargo, y teniendo en cuenta el creciente número de ICPs realizadas en todo el mundo y el uso generalizado de *stents* farmacoactivos, la reestenosis sigue siendo un problema de salud de magnitud por motivos diferentes<sup>58,59</sup>. Primero, como es lógico, por el importante incremento que se ha producido en la utilización del *stent* intracoronario como forma de revascularización coronaria desplazando a la cirugía de pontaje aorto-coronario. En segundo lugar, y aunque la reestenosis puede estar asociada con una recurrencia de los síntomas de angina estable, se reconoce que hasta 1 de cada 3 pacientes presentan infarto de miocardio o angina inestable susceptible de repetir la intervención<sup>60,61</sup>. Por tanto, reestenosis se asocia en numerosas ocasiones con un cuadro coronario agudo, inestable, y de alto riesgo. Finalmente, aunque se han realizado diferentes estrategias para el tratamiento de la reestenosis coronaria basándose en mecanismos de actuación potencialmente diferentes, ninguna se ha mostrado superior. De hecho, la más utilizada es la angioplastia convencional con balón, que actúa principalmente por compresión del tejido neointimal. Se han utilizado también otras como la aterectomía (direccional y rotacional) y el láser con la finalidad de extraer, eliminar o ablacionar el máximo grado de proliferación neointimal; e incluso la irradiación intracoronaria, intentando inhibir el máximo grado de respuesta proliferativa. Finalmente, una última estrategia, cada vez más utilizada, es la implantación de un *stent* adicional en el interior del *stent* inicial. La falta de estudios aleatorizados comparativos para tratar la reestenosis coronaria, así como la falta de seguimientos angiográficos sistemáticos, hace que el tratamiento de la reestenosis sea un reto actual, especialmente en pacientes con reestenosis repetidas<sup>62</sup>.

## **6.2. Utilizando *machine learning* para predicción de reestenosis coronaria**

A pesar de conocerse muchos factores de riesgo independientes para desarrollar reestenosis coronaria (diabetes mellitus, vasos pequeños, bifurcaciones, lesiones largas, tipo y número de *stents*, etc), no se sabe con exactitud qué paciente, arteria o lesión sufrirá reestenosis.

Esta es sin duda la principal aportación de esta tesis; ser capaces de desarrollar un modelo *machine learning* para predecir el riesgo de reestenosis, con una mejor estratificación de riesgo para el paciente que los *scores* de riesgo clásico, y por tanto con relevancia clínica. De hecho, con recomendaciones personalizadas derivadas del modelo, a los pacientes con infarto agudo de miocardio sometidos a ICP con implante de un *stent* coronario, podrían evitarse seguimientos innecesarios y disminuir los

costes en aquellos pacientes con bajo riesgo de sufrir reestenosis coronaria y; por el contrario, mejorar los resultados en pacientes con alto riesgo, estableciendo pruebas de detección isquémica a los 12 meses o incluso realizando angiografía rutinarias y programadas, con el objeto de evitar sucesos clínicos adversos.

El ensayo GRACIA-3 es único en el desarrollo de este modelo predictivo *machine learning* por varias razones. En primer lugar, brinda la oportunidad de incorporar información (variables) con respecto al uso de *stents* metálicos convencionales frente *stents* farmacológicos en un contexto clínico, infarto agudo de miocardio con elevación del segmento ST, donde el beneficio de los nuevos *stents* farmacológicos aún es discutible y se sigue analizando<sup>63</sup>. Por tanto, nuestro modelo predictivo sería específicamente útil en este escenario específico. En segundo lugar, el número de estudios angiográficos en el seguimiento realizados en el ensayo GRACIA-3 (60,3%) es considerable y similar a los estudios contemporáneos en grandes cohortes de pacientes donde también se programaron cateterismos de seguimiento<sup>64</sup>. Finalmente, los datos provienen de un ensayo clínico aleatorizado, que involucró a 20 hospitales españoles, donde todas las angiografías se analizaron en un laboratorio central independiente. Por consiguiente, la metodología utilizada para definir la presencia de reestenosis fue de máximo rigor y calidad y representa una excelente oportunidad para desarrollar un modelo predictivo de reestenosis basado en *machine learning*.

Los algoritmos *machine learning* son una tendencia creciente en el análisis de datos biomédicos por su capacidad para encontrar patrones complejos, que generalmente aparecen en este campo<sup>10</sup>. Este enfoque se ha utilizado para estudiar varias patologías cardíacas<sup>27,65</sup>, como se detalla en el apartado final de esta discusión. En el caso de la reestenosis, hasta donde sabemos, el único estudio que ha utilizado un enfoque *machine learning* fue desarrollado por Cui et al<sup>66</sup>. Para este estudio se utilizaron dos cohortes de pacientes. Una primera cohorte de 400 pacientes con enfermedad coronaria sometidos a ICP en el hospital Anzhen de Pekin entre enero 2013 y noviembre 2013. A todos los pacientes se les implantó un stent farmacoactivo. Los pacientes fueron seguidos a los 1 y 6 meses. Al mes se realizaron muestras sanguíneas para determinar metabolitos plasmáticos de fosfolípidos y esfingolípidos relacionados con la presencia de reestenosis coronaria, cuya incidencia fue determinada en el estudio angiográfico realizado a los 6 meses. Los 36 pacientes en los que se identificó la presencia de reestenosis coronaria fueron comparados con los pacientes sin reestenosis (grupo control). Se identificó un conjunto de 6 metabolitos plasmáticos relacionados con un mayor riesgo de sufrir reestenosis. La validación del valor pronóstico de estos metabolitos se realizó en una segunda cohorte de 500 pacientes sometidos a ICP en el mismo hospital entre mayo 2014 y junio 2015. Utilizando dos clasificadores *machine*

*learning: random forest* y *support vector machine* los investigadores obtuvieron una excelente área bajo la curva ROC de 0,94, en la predicción de reestenosis coronaria. El clasificador *random forest* fue superior al *support vector machine*. Asimismo, los investigadores utilizaron un esquema de validación cruzada por 100 y un test de permutación en su metodología. Los resultados de este modelo son superiores al nuestro, área bajo la curva ROC de 0,77 en nuestro ejemplo. Sin embargo, los biomarcadores utilizados en el estudio de Cui et al. no suelen medirse en la práctica clínica habitual y su medición extrahospitalaria resulta costosa y complicada. En este sentido, desarrollar un modelo para predecir reestenosis basado en variables obtenidas en la práctica clínica como el nuestro, puede ser una herramienta valiosa para mejorar la estratificación de riesgo de pacientes rutinarios; y por tanto, ser más extrapolable a la rutina diaria (objetivo final de cualquier modelo *machine learning*).

### **6.3. Conjunto de datos pequeños y no balanceados para entrenar algoritmos *machine learning***

Posiblemente el paso más significativo en el desarrollo de nuestro modelo *machine learning* haya sido la construcción del conjunto de datos que hemos utilizado para entrenamiento, validación y *test* para determinar la técnica *machine learning* a utilizar en el clasificador. La clave es que los datos estén bien seleccionados, limpiados y transformados de forma que el algoritmo pueda descubrir fácilmente los patrones. Si los datos no son de calidad, incluso el mejor algoritmo nos dará unas predicciones de baja calidad. Esta fase de recolección y preparación de datos puede llevar entre un 80% y 90% del tiempo de desarrollo de un sistema *machine learning*.

Hay muchas herramientas para la selección, limpieza y transformación de datos: Trifacta<sup>67</sup>, OpenRefine<sup>68</sup>, Paxata<sup>69</sup>, Alteryx<sup>70</sup>, etc. La mayoría de ellas desarrolladas por compañías como IBM, Google, Accenture, Amazon, Microsoft, Tableau, etc. Sin embargo, es común en el ámbito científico el uso de bibliotecas libres de acceso y gratuitas como Pandas para Python<sup>71</sup>, Dplyr para R<sup>72</sup>, o Optimus para Apache Spark<sup>73</sup>.

Pandas fue la librería utilizada en esta tesis al ser de código abierto, proporcionar estructuras de datos y herramientas de análisis de datos de alto rendimiento y, ser fácil de utilizar para el lenguaje de programación Python. Es además uno de los formatos estándar para manejar datos estructurados en Python, y se integra perfectamente con los modelos de clasificación implementados en *scikit-learn*. Por lo tanto, Pandas lleva a cabo tareas importantes, como alinear datos para su comparación, fusionar conjuntos de datos, gestión de datos perdidos, etc. La estructura de datos que nosotros utilizamos

proveniente de Pandas se denomina *DataFrame*, que es una colección ordenada de columnas con nombres y tipos, parecido a una tabla de base de datos, donde una sola fila representa un único caso y las columnas representan atributos particulares. Los elementos de distintas columnas pueden ser de diferentes tipos.

La selección de los datos fue dificultosa y controvertida. Teniendo en cuenta que los clasificadores *machine learning* buscan miles de asociaciones, no siempre tenidas en cuenta en la lógica humana, determinamos incluir en el entrenamiento de los datos aquellas variables de entrada que los métodos de selección de características identificaron más rentables, independientemente de que el número de sucesos fuese bajo. Este es el ejemplo de la variable plaquetas anormales, que se mostró como la cuarta variable en importancia en el modelo predictivo, a pesar de que solamente un paciente de la muestra presentaba esta característica. Tal vez hubiese sido apropiado no utilizar esta variable de entrada, pero preferimos no influir en demasía en la selección de variables, priorizando la calidad de su recogida. No obstante, nos tranquilizó observar que la robustez del modelo *machine learning* desarrollado persistía aunque desechásemos esta variable alcanzando un área bajo la curva ROC de 0,75. Sin duda, seleccionar las variables que el clínico considere significativas en base a asociaciones previas documentadas científicamente es clave en el desarrollo de esta nueva tecnología, y el trabajo codo con codo entre ingenieros, estadísticos, físicos, médicos, etc, se antoja esencial.

Uno de los principales interrogantes del *machine learning* es su aplicación en el análisis de conjuntos de datos pequeños y no balanceados como en nuestro caso. Para estas ocasiones cada vez son más numerosos los ejemplos que utilizan métodos de remuestreo, *resampling methods*, basados en extraer muestras repetidamente a partir del *set* de datos de entrenamiento o del *test*. Se trata de métodos no paramétricos, que no requieren ninguna asunción sobre la distribución de la población. De los métodos utilizados habitualmente los más usados son la validación cruzada, *cross-validation*, y el *bootstrap*. Nosotros utilizamos la validación cruzada por su aplicabilidad en diferentes escenarios: estimar el error sobre el conjunto de *test* asociado a clasificadores tanto de regresión como clasificación, evaluar el rendimiento del modelo, pero también para optimizar el modelo seleccionando los hiperparámetros más adecuados.

De forma global existen tres tipos de validaciones cruzadas: validación cruzada aleatoria, validación cruzada dejando uno fuera y, validación cruzada de k iteraciones.

El método de validación cruzada aleatoria o simple se basa en dividir de manera aleatoria (y tamaño comparable) el set de observaciones disponible en dos partes, el set



de entrenamiento y el set de validación. El modelo es ajustado usando el set de entrenamiento, y este mismo modelo se usa para predecir las nuevas observaciones del set de validación, obteniendo así la estimación del error asociado. Los dos inconvenientes de este método son que la estimación del error asociado es muy variable, dependiendo que las observaciones hayan sido seleccionadas para formar parte del set de entrenamiento o validación y; sólo un subconjunto, la mitad aproximadamente, de observaciones se utiliza para entrenar el modelo, por lo que el error del set de validación puede sobreestimar el error asociado.

El método de validación cruzada dejando uno fuera, o *leave-one-out cross validation*, es muy parecido al método de validación simple previamente descrito. De nuevo, el set de datos se divide en dos partes; pero en este caso se utiliza una sola observación para el set de validación, dejando al resto para el set de entrenamiento. Es decir, el modelo se ajusta utilizando  $n-1$  observaciones de entrenamiento. Las ventajas respecto a la validación simple son: sufre menos errores de selección, reduce la variabilidad, reduce la aleatoriedad en los resultados. Sin embargo, como contrapartida sufre la desventaja del coste computacional al tener que ajustar  $n$  modelos, especialmente cuando  $n$  es elevado. Además, puede sufrir de *overfitting*, ya que se acaban empleando todos los datos para ajustar el modelo.

Finalmente, la validación cruzada de  $k$  interacciones, o *k-fold cross validation*, utilizado para esta tesis doctoral, supone una alternativa a los anteriores y se ajusta mejor a conjunto de datos pequeños y no balanceados. En este caso, las observaciones se dividen aleatoriamente en  $k$  grupos de igual tamaño. Uno de los  $k$  grupos se emplea como set de validación, mientras que el resto  $k-1$  grupos se emplean para entrenar el modelo. El error asociado se calcula sobre el grupo  $k$  excluido del modelo, proceso que se repite  $k$  veces (cada vez un grupo  $k$  diferente es usado como set de validación). Este proceso generalmente se repite un número de veces aleatorizando la división del dataset en  $k$  subpartes. Se obtienen, por tanto,  $k$  estimaciones del error asociado, por cada repetición, calculándose la estimación global promediando el total de los valores. Las ventajas de esta metodología sobre la anterior son que computacionalmente es menos costoso, ya que es posible elegir el número de entrenamientos-*test* que se pretenden realizar para evaluar la robustez del clasificado; la variabilidad de las estimaciones es menor, debido a como se dividen las observaciones en los  $k$  grupos, y por tanto las estimaciones son más precisas. El único inconveniente respecto a la validación cruzada dejando uno fuera es que tiene más errores de selección, aunque siempre menos que la validación cruzada simple.

Podemos decir que un modelo *machine learning* tiene dos tipos de parámetros: aquellas variables o datos que utilizamos para su entrenamiento, y sobre los que el algoritmo se ajusta y; los parámetros o valores que podemos ajustar al mismo algoritmo para realizar los cálculos de aprendizaje necesarios en su aprendizaje: este paso es conocido como el tuneado de los hiperparámetros, que no tiene otro objeto que obtener el mejor rendimiento de cada modelo.

A la hora de elegir el número de subdivisiones y repeticiones para la utilización de la validación cruzada externa (empleada para evaluar el error de generalización del modelo), se ha optado por el uso de 10-*fold*s con 20 repeticiones. Existe cierto debate sobre cuál es el mejor esquema de validación cruzada, siendo probablemente los esquemas más usados en la literatura los de 5-*fold*s con 2 repeticiones y 10-*fold*s con 10 repeticiones. Dado que el conjunto de datos empleado no contaba con un elevado número de casos, se ha optado por el uso de 10-*fold*s y 20 repeticiones, para obtener conjuntos de entrenamiento de mayor tamaño y obtener una aproximación más precisa del error de generalización. Por otro lado, en la *cross-validation* interior se han utilizado 9-*fold*s con una repetición, con el objetivo de mantener subconjuntos de datos del mismo tamaño de la validación exterior y no disparar el coste computacional del proceso. En el caso del algoritmo con mejor comportamiento basado en el clasificador *Extremely Randomized Trees*, el tiempo total de cómputo del esquema anidado de validación cruzada fue de 3 horas utilizando un procesador de 4 núcleos. Probablemente, la utilización de este esquema de validación cruzada anidada habría sido difícil de manejar con una muestra mayor. Por otro lado, el entrenamiento del modelo final ajustado con el total del conjunto de datos e hiperparámetros optimizados con una validación cruzada de 10-*fold*s y una repetición no tomó más de 15 minutos.

Significativa es la coincidencia con el estudio de Cui et al. que los algoritmos de decisión que obtuvieron mayor capacidad predictiva se basaron en el uso de árboles de decisión: *random forest* en el estudio de Cui et al. y *extremely randomized trees* en nuestro caso. Este tipo de clasificadores y la visualización de su efectividad son fácilmente comprensibles. En los árboles de decisión se realiza una división de clases atendiendo a aquel atributo que mejor distinga entre un grupo y otro, repitiéndose este proceso recursivamente hasta que todos los datos de un subconjunto pertenezcan a la misma clase. La efectividad del uso de árboles de decisión en machine learning en medicina se ha generalizado para detección de enfermedades desbalanceadas como el cáncer. Entre las ventajas de este tipo de algoritmos, se encuentra la gran flexibilidad que otorga la elección de sus hiperparámetros y su facilidad para manejar datos de cualquier tipo y con valores faltantes. Esto los convierte en una buena opción tanto como primera aproximación como para la elección de un modelo final, como ocurre en

el presente estudio. Como inconvenientes en este tipo de algoritmos podemos encontrar baja precisión si lo comparamos con otros clasificadores en caso de que los datos tengan una distribución uniforme sin grandes diferencias; distribución, como hemos señalado, no presente en nuestro caso.

Finalmente, es importante señalar que la construcción de este conjunto de datos, procedentes del estudio GRACIA-3, no debería limitarse al entrenamiento y validación de los modelos *machine learning* desarrollados durante esta tesis ni a la producción científica originada. Es importante poner a disposición de la comunidad científica estos conjuntos de datos con el objetivo de poder evaluar otras técnicas ahora no disponibles, e incluso mejorar nuestro modelo predictivo. Estos conjuntos de datos, procedentes de estudios multicéntricos y no fáciles de obtener, deben ser utilizados en retos tecnológicos como el *datathon*, basados en el fenómeno de la inteligencia artificial. El *datathon* es un concurso público y abierto, en el que personas interesadas en el tema, desde novatos a expertos, trabajan conjuntamente con bases de datos comunes con un objetivo concreto previamente establecido. Los participantes deben encontrar una solución al objetivo marcado en los conjuntos de datos proporcionados. Las soluciones más precisas y creativas son premiadas y de acceso público. El acceso a nuestro conjunto de datos para la comunidad científica estará disponible, si el Comité Científico del ensayo GRACIA-3 lo considera oportuno.

## **6.4. Los scores de riesgo desarrollados con modelos estadísticos clásicos**

La construcción de modelos o *scores* de riesgo de aparición de un suceso cardiovascular es de gran importancia en medicina, tanto para conocer qué variables influyen en su aparición, como para predecir su aparición. Al ser las enfermedades cardiovasculares la principal causa de muerte en nuestro entorno, es de gran interés el desarrollo de modelos de predicción de riesgo de padecer diferentes patologías cardiovasculares.

Para el cálculo de la probabilidad de aparición de un suceso binario como el nuestro (reestenosis, sí o no), el modelo matemático más habitual se ha basado en la utilización de métodos estadísticos clásicos como la regresión logística, que produce una ecuación en la que conocidos los valores de los diferentes factores asociados con la enfermedad se puede evaluar la presencia de dicha enfermedad.

Los pasos a seguir en el desarrollo de los scores de riesgo con modelos estadísticos clásicos comprenden: (i) un análisis bivalente inicial, de las variables explicativas para el suceso cardiovascular adverso utilizando pruebas estadísticas como Chi-cuadrado para variables categóricas o t-student para las numéricas. (ii) Análisis multivariante. Una vez que la adecuación de incluir cada variable en el modelo ha sido analizada por separado, es necesario analizar todas las variables en conjunto a través de un análisis multivariante, introduciendo variables categóricas como numéricas en un modelo de regresión logística con un método de selección por pasos (*stepwise*). La regresión logística es una de las metodologías más extendidas en la construcción de *scores*. La razón por la que se utiliza este tipo de regresión es que la variable dependiente es una probabilidad, por lo que debe estar acotada al intervalo (0,1). La variable dependiente representa la ocurrencia o no de un suceso. Otra ventaja añadida es que este tipo de regresión permite que las variables independientes sean cuantitativas, discretas o continuas, o categóricas; sin que haya ninguna restricción. Asimismo, no requiere el supuesto de normalidad de los errores. La capacidad predictiva se mide mediante la comparación entre el grupo de pertenencia observado y estimado por el modelo, que clasifica a los individuos en función de un punto de corte establecido, que por defecto será de 0,5. Es decir, si la probabilidad estimada es mayor a 0,5 el individuo será clasificado como 1, y viceversa. La capacidad predictiva se mide por el porcentaje de individuos que siendo 0 o 1, el modelo los clasifica como tal. (iii) Obtención del *score*. Una vez las variables han sido categorizadas, se asigna una puntuación entera (*integer score*) a cada variable proporcional al coeficiente continuo estimado del modelo de regresión logística, tal y como representa la **Tabla 7** en relación a los *scores* para riesgo de reestenosis coronaria clásicos, con los que nuestro modelo predictivo *machine learning* se comparó; PRESTO-1, PRESTO-2 y TLR. Los modelos estadísticos en los que se calcula la probabilidad de reestenosis según los factores predictivos independientes de reestenosis son los desarrollados tras los estudios PRESTO (*Prevention of Reestenosis with Tranilast and its Outcomes*)<sup>45</sup>, y registro EVENT (*Evaluation of Drug-Eluting Stents and Ischemic Events*)<sup>46</sup>.

**Tabla 7. Información utilizada en el Desarrollo de los scores clásicos de riesgo de reestenosis**

<b>Prevention of restenosis with tranilast and its outcomes (PRESTO) trial<sup>45</sup></b>				
<b>Variable</b>	<b>OR</b>	<b>IC 95%</b>	<b>P</b>	<b>Puntuación</b>
Longitud lesión >20 mm	2,07	1,34–3,21	0,001	4
Lesión tipo C	1,81	1,26–2,59	0,001	3
ICP previa	1,46	1,14–1,88	0,003	2
Diabetes mellitus	1,41	1,03–1,92	0,034	2
No fumador	1,39	1,05–1,83	0,022	2
Tamaño del vaso				
>4 mm	1 (referencia)			
3.5–4 mm	1,18	0,55–2,53	0,680	1
3–3.5 mm	1,44	0,71–2,94	0,317	2
≤3 mm	1,76	0,87–3,54	0,115	3
Angina inestable	1,19	0,94–1,51	0,147	1
Sexo femenino	1,15	0,87–1,52	0,317	1
<b>Evaluation of drug-eluting stents and ischemic events (EVENT) registry<sup>46</sup></b>				
<b>Variable</b>	<b>OR</b>	<b>IC 95%</b>	<b>P</b>	<b>Puntuación</b>
Edad <60 años	1,49	1,14–1,95	0,003	1
ICP previa	1,83	1,40–2,39	<0,001	2
ICP tronco coronario	3,14	1,30–7,57	0,010	3
ICP pontaje con safena	2,40	1,62–3,57	<0,001	2
Diámetro del stent ≤2,5 mm	1,54	1,17–2,02	0,001	1
Longitud total del stent ≥40 mm	1,78	1,35–2,35	<0,001	1

## 6.5. Declaración TRIPOD

Para la realización de este estudio, se han tenido en cuenta las recomendaciones de la declaración TRIPOD para la publicación y evaluación de modelos y scores de riesgo clínicos. En este sentido, y para facilitar la reproducibilidad del entrenamiento y aplicación de los modelos desarrollados en esta tesis, se han especificado los detalles de los modelos empleados, entre ellos, las librerías de código abierto que implementan los algoritmos de selección de variables y clasificación, sus hiperparámetros y los esquemas de validación cruzada. Así mismo, se ha puesto a disposición de otros investigadores el modelo de clasificación final entrenado con el total de la muestra y se proporcionan herramientas para la aplicación directa del mismo. Estos pasos se hacen necesarios a la hora de reportar cualquier modelo clínico basado en *machine learning*, los cuales pueden conllevar un número elevado de operaciones complejas irrealizables de forma manual y alejadas de los scores de riesgo clásicos basados en regresión logística. Por otro lado, en la evaluación interna y externa de modelos y scores de riesgo, se han aportado detalles de los resultados obtenidos en forma tanto de curvas ROC, típicamente empleadas en la literatura clínica, como de curvas PR, más ampliamente utilizadas en el ámbito de los clasificadores basados en *machine learning* aplicados a

clases desbalanceadas. Así mismo, se aporta una interpretación de dichas curvas con la elección de diferentes puntos de corte para la estratificación de grupos de riesgo y las métricas de clasificación asociadas a los mismos.

## 6.6. Curvas ROC y PR como medida de evaluación para un sistema clasificador binario

Los modelos *machine learning* funcionan de manera eficiente cuando se les proporciona una gran cantidad de datos y categorías balanceadas. Estas condiciones generalmente no se cumplen cuando se trata de datos que consisten en pacientes que desarrollarán o no una patología, como fue nuestro caso de la reestenosis coronaria. En este sentido, las curvas ROC y PR se usan comúnmente para presentar resultados para problemas de decisión binarios en *machine learning*. Cuando la distribución de clases está balanceada, las curvas ROC son adecuadas para valorar el modelo. Sin embargo, cuando se trata de un conjunto de datos muy desbalanceado como el nuestro, también se deben utilizar curvas PR. Una diferencia importante entre curvas ROC y PR es la representación visual de las curvas. El objetivo en el espacio ROC es estar en la esquina superior izquierda y cuando uno mira la curva ROC del modelo ERT desarrollado, parece estar bastante cerca del óptimo<sup>52</sup>. En el espacio PR, el objetivo es estar en la esquina superior derecha, y la curva PR muestra que el modelo ERT desarrollado es óptimo cuando se predice la no aparición de reestenosis, pero todavía hay un margen significativo para mejorar cuando se predice reestenosis. Por lo tanto, en nuestro ejemplo, las curvas PR ofrecen una imagen más precisa del rendimiento de nuestro algoritmo.

Las curvas PR pueden aportar claridad a la interpretación de los resultados y facilitar la elección de puntos de corte para la toma de decisiones basadas en el modelo. La particularidad de las curvas PR es su dependencia de la prevalencia, por lo que, en el caso de clases desbalanceadas, son mucho más sensibles al cambio en el rendimiento de los modelos de clasificación. Dicho de otra forma, cambios pequeños en las curvas ROC en problemas con clases desbalanceadas pueden ser significativos en las curvas PR, por lo que tanto la interpretación visual de la curva, como el valor numérico del área bajo la curva PR (o precisión media) puede ser más informativo dependiendo del uso que se quiera hacer del modelo, que puede ser el de buscar una alta sensibilidad (y predecir pacientes de bajo riesgo que no requieran seguimiento) o una alta especificidad (para predecir pacientes de alto riesgo que necesiten un seguimiento cercano). En general, estas conclusiones pueden ser extraídas igualmente del análisis

de la curva ROC, es decir, qué modelo domina en diferentes zonas de la curva, pero el uso de curvas PR supone un cambio de escala que tiene en cuenta el desbalanceo de clases y amplifica dicha dominancia. En el caso de la presente tesis, es reseñable el hecho de que el modelo final elegido tenía un mejor comportamiento que el resto en prácticamente todo el espectro de la curva ROC, y por tanto de la curva PR, por lo que su elección resultó trivial. Sin embargo, pueden tomarse como ejemplos las zonas de muy alta sensibilidad del modelo TLR y la zona de sensibilidad entre 0,3 y 0,5 del modelo basado en *gradient boosting*, en las cuales el modelo ERT era inferior, respectivamente. En este sentido, se ha incluido en las calculadoras para el uso del modelo el informe de los valores de sensibilidad, especificidad, valor predictivo negativo y precisión de cada predicción para dejar la libertad de elegir puntos de corte que se ajusten a cualquier otro requisito.

## **6.7. Variables predictoras en modelos *machine learning***

Como se ha señalado previamente, la reestenosis coronaria es una enfermedad iatrogénica y la identificación de pacientes, arterias, lesiones, etc., susceptibles de sufrir reestenosis coronaria ha sido y es una tarea importante<sup>74</sup>. Diferentes estudios, entre los que se incluyen los que desarrollaron los scores clásicos PRESTO-1, PRESTO-2, y TLR, han detectado una serie de factores relacionados con la aparición de reestenosis coronaria, tanto clínicos y angiográficos como de procedimiento; los más significativos se exponen en la **Tabla 8**.

**Tabla 8. Factores relacionados con la aparición de reestenosis coronaria en diferentes estudios**

Estudio	Pacientes	Seguimiento angiográfico	Factores predictivos
Kasaoka et al. <sup>75</sup> (1998)	1706	70%	Edad Diabetes Cirugía coronaria previa Longitud stent Diámetro referencia Diámetro luminal mínimo post-ICP
Bauters et al. <sup>76</sup> (1998)	463	81	Múltiples stents Stent tipo Palmaz-Schatz Longitud lesión Diámetro referencia del vaso
Kastrati et al. <sup>77</sup> (1997)	1349	80	Múltiples stents Diabetes mellitus Diámetro luminal mínimo post-ICP ICP previa Oclusión crónica Implantación en descendente anterior
Hoffman et al. <sup>78</sup> (1998)	291	100	Localización ostial Área seccional trasversal pre-ICP Diámetro luminal mínimo post-ICP
Singh et al. <sup>45</sup> (2004)	1312	100	Mujer Tamaño del vaso Longitud lesión Diabetes mellitus Tabaquismo Lesión tipo C ICP previa Angina inestable Lesión ostial
Stolker et al. <sup>46</sup> (2010)	8829	3,6%	Edad <60 ICP previa ICP de tronco principal izquierdo ICP de injerto de vena safena Diámetro mínimo del stent Longitud total del stent

Dentro de los factores clínicos, el más consistente es la presencia de diabetes. Otros factores clínicos relacionados han sido una edad avanzada o antecedentes de revascularización previa (tanto percutánea como quirúrgica)<sup>45,46,75,77</sup>. En relación con los factores angiográficos predictivos de reestenosis, los más frecuentemente identificados han sido la longitud de la lesión, la implantación en vasos de pequeño calibre, y la localización de la lesión a tratar en la arteria descendente anterior<sup>45,46,76,77</sup>. Otros factores menos comunes han sido la localización ostial de la lesión o la severidad de la estenosis preprocedimiento<sup>45,78</sup>. Los factores de procedimiento relacionados han sido un menor diámetro luminal mínimo al finalizar el procedimiento y la implantación



de múltiples stents en una misma arteria<sup>75-78</sup>. En nuestro modelo machine learning identificamos como variable de referencia una variable clínica, diabetes mellitus; tres variables angiográficas, enfermedad multivaso, flujo TIMI post-ICP y, presencia de trombo post-ICP y; dos variables bioquímicas, colesterol y plaquetas anormales. Variables clínicas y angiográficas se han repetido constantemente como asociadas a riesgo de reestenosis en estudios observacionales previos con análisis estadístico clásico basado en regresión logística múltiple.

## **6.8. Integrando los modelos *machine learning* en la práctica clínica diaria**

Las plataformas de *machine learning* está acercando sustancialmente el análisis predictivo, frente al análisis descriptivo, al ámbito médico. El proceso por el que estamos pasando nos recuerda a la evolución que tuvieron las bases de datos en los años 80 y 90 del siglo pasado: lo que en un principio era difícil de explicar, como funcionan y para qué sirven, ahora está de tal forma integrado en todos los sistemas que es complicado encontrarse con uno que no tenga una base de datos en su ordenador.

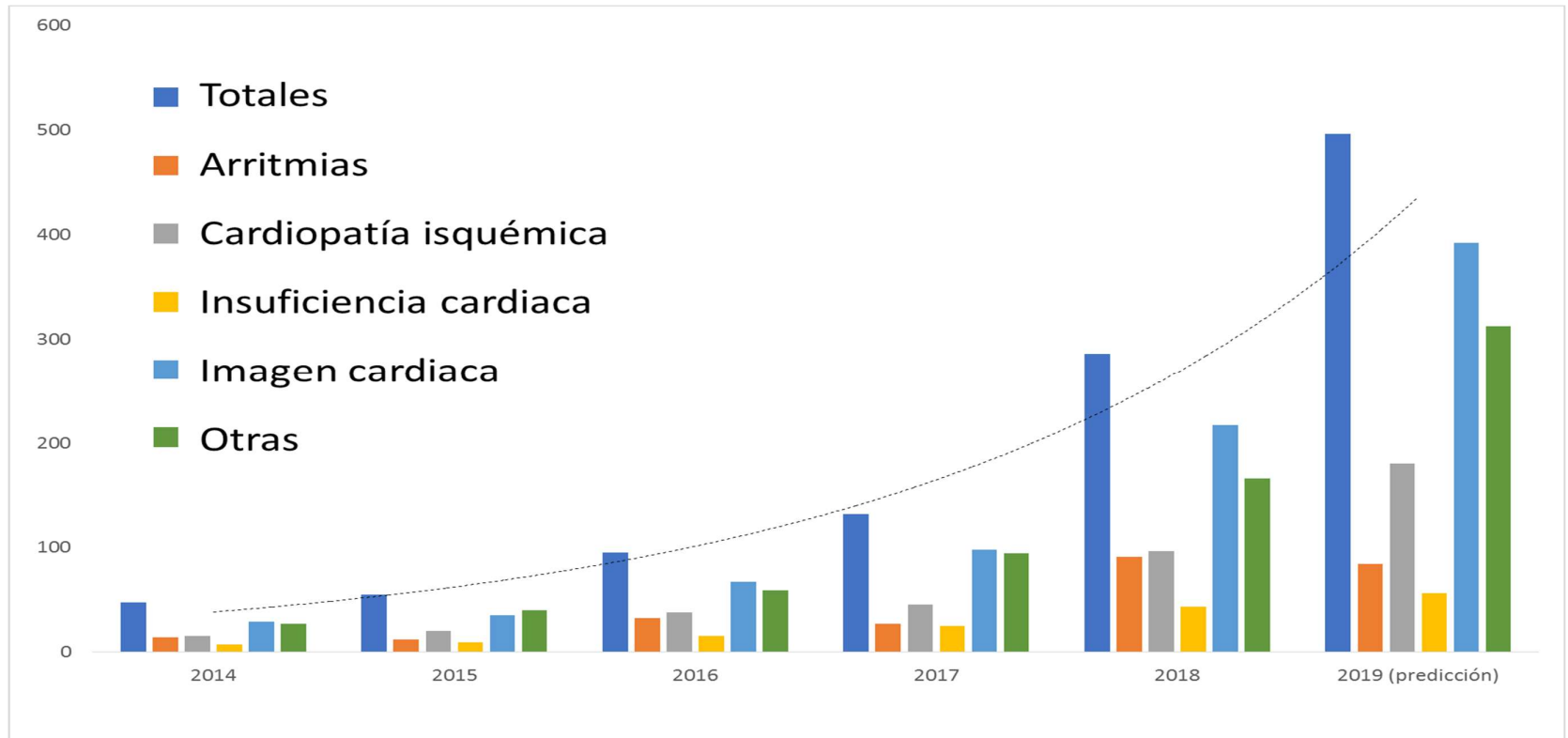
Es relevante la aportación que esta tesis presenta en relación con una demostración práctica de uso de un producto *machine learning*. Como mejora, existiría la posibilidad de dotar al sistema de la capacidad de trabajar con nuevos datos suministrados por la comunidad al hacer uso del modelo. De esta manera, podrían realizarse una validación en otras poblaciones y detectar nuevos patrones predictivos de la restenosis. Al contrario de los que sucede con los scores de riesgo tradicionales, nuestro sistema podría ajustarse a un entorno cambiante, adaptándose a las nuevas situaciones según se le vaya alimentando con nuevos datos.

Nuestra aportación proporciona un modelo *machine learning* que mejora de forma significativa el rendimiento de los *scores* clásicos de reestenosis. Además, proporciona las herramientas para que pueda llevarse a la práctica clínica diaria. Las calculadoras de riesgo desarrolladas son de uso muy sencillo por parte de personal médico y no requieren grandes recursos de ningún tipo. Las calculadoras de riesgo desarrolladas identifican la zona de riesgo en la que se sitúa el paciente, permitiendo estratificar a los mismos y tomar decisiones basadas en esta estratificación.

## 6.9. Inteligencia artificial en Cardiología

Aunque a menudo identificamos la inteligencia artificial como un concepto futurista y lejano, lo cierto es que es una tecnología que ya se emplea a día de hoy en todo tipo de áreas, entre ellas, la cardiología<sup>79</sup>. Gracias a la digitalización de grandes cantidades de datos, el desarrollo de algoritmos *machine learning* y la mejora de la capacidad de cálculo de los ordenadores de las últimas décadas, la inteligencia artificial puede abrir excelentes oportunidades para la automatización de tareas, la aplicación de la medicina de precisión o el avance en la investigación a través de la detección de patrones complejos en bases de datos médicas<sup>80</sup>. Un caso particular es el del análisis de imagen médica, en el que las técnicas *deep learning* han supuesto una auténtica revolución y su aplicación en el área de la cardiología está ya obteniendo excelentes resultados. Sin embargo, existe aún un largo camino que recorrer para aplicar de forma generalizada estas técnicas en la práctica clínica, para lo cual será fundamental la generación de grandes bases de datos con información de calidad y la evaluación e integración de la inteligencia artificial en contextos clínicos realistas.

Los ejemplos de inteligencia artificial, utilizando *machine learning* o *deep learning*, comienzan a ser numerosos en cardiología<sup>79</sup>. En esta sección presentamos una evolución de las contribuciones de inteligencia artificial a las diferentes áreas de aplicación en cardiología (**Figura 27**).



**Figura 27.** Evolución en las publicaciones referenciadas en PubMed sobre inteligencia artificial, *machine learning* y *deep learning* según el área de interés en cardiología.

En el material suplementario de esta tesis (anexos) se describen los detalles de las publicaciones por áreas de interés y la metodología de búsqueda empleada.

Se ha realizado una búsqueda detallada en PubMed de la evolución en el último lustro de las contribuciones de la inteligencia artificial a las diferentes áreas de aplicación en cardiología, y se han identificado un total de 673 artículos originales. Este material se ha incluido como material suplementario de esta tesis, incluyéndolo en sus anexos. Sin embargo, en este apartado hemos querido describir en detalle 19 ejemplos de diferentes áreas de la cardiología que utilizando inteligencia artificial han mostrado mejoras diagnósticas y terapéuticas, y que facilitan la comprensión de la metodología *machine learning* y *deep learning* aplicada a la cardiología. Esta información se ha resumido en la **tabla 9**<sup>81-99</sup>. Es importante reseñar que muchas de las publicaciones revisadas son resultado de iniciativas habituales en inteligencia artificial como retos promovidos en conferencias<sup>83,100</sup> y *datathons*. El *datathon* es un concurso público y abierto, en el que personas interesadas en el tema, desde novatos a expertos, trabajan conjuntamente con bases de datos comunes con un objetivo concreto previamente establecido. Los participantes deben encontrar una solución al objetivo marcado en los conjuntos de datos proporcionados. Las soluciones más precisas y creativas son premiadas y de acceso público.

### 6.9.1. Arritmias cardiacas

Una de las aplicaciones más extendidas del *machine learning* en cardiología es la predicción de arritmias cardiacas. Un gran número de estudios abordan modelos predictivos de aparición de fibrilación auricular, incluso paroxística, por su significancia e implicaciones clínicas; utilizando aprendizaje supervisado con sistemas predictivos *machine learning* compuestos de diferentes subprocesos: preprocesado de señal, extracción de variables significativas y algoritmos de clasificación<sup>81,101</sup>. Modelos similares *machine learning* se han desarrollado en la mejora de la gestión de alarmas de telemonitorización<sup>83</sup>, para predecir la aparición de arritmias ventriculares<sup>102,103</sup>, en la predicción a la respuesta de procedimientos de ablación invasivos como la crioablación<sup>82</sup>, incluso para predecir la mortalidad tras parada cardiaca resucitada<sup>84</sup>. Se han empleado técnicas *deep learning* con gran éxito en la detección de diferentes tipos de arritmias a partir del análisis directo de imágenes o señales electrocardiográficas<sup>85</sup>. Otra aplicación ha sido la búsqueda de fenotipos para clasificar miocardiopatías hipertróficas con diferente riesgo arrítmico utilizando aprendizaje no supervisado<sup>104</sup>.

### 6.9.2. Cardiopatía isquémica

La inteligencia artificial se ha aplicado a los datos rutinarios de la historia electrónica proveniente de atención primaria para predecir el riesgo de enfermedad cardiovascular en la población general; mostrándose superior a las escalas de riesgo tradicionalmente utilizadas<sup>86</sup>. Asimismo, se han aplicado técnicas de aprendizaje supervisado en la predicción pronóstica de la cardiopatía isquémica estable<sup>105</sup>, en el síndrome coronario<sup>87</sup>; o la predicción de mortalidad de enfermos con infarto de miocardio analizando resultados de hospitales individuales<sup>88</sup>, o grandes registros como el SWEDEHEART<sup>89</sup>. Los resultados son dispares debido al tamaño de las muestras, demostrando que las técnicas *machine learning* obtienen mejores resultados con tamaños de muestra grandes.

### 6.9.3. Insuficiencia cardiaca

Los sistemas *machine learning* podrían permitir optimizar los ingresos hospitalarios evitables por insuficiencia cardiaca identificando de forma más precisa qué pacientes son susceptibles de descompensación cardiaca tras el alta hospitalaria que las escalas de riesgo clásico<sup>90,106</sup>; este es un ejemplo de resultados contradictorios de la experiencia inicial en este campo tecnológico<sup>107</sup> que apunta a la importancia en el ajuste en la metodología de los sistemas<sup>108</sup>. Asimismo, hay estudios iniciales que abordan la utilidad de la inteligencia artificial como sistema de gestión en la telemonitorización remota de pacientes con insuficiencia cardiaca. Los estudios realizados hasta el momento han mostrado que la utilización de *machine learning* es factible<sup>109</sup>, y que es posible que pueda mejorar la evolución clínica de estos pacientes<sup>91</sup>. Otra área importante es el trasplante cardiaco donde sistemas *machine learning* se han aplicado para predecir la probabilidad de muerte o trasplante de pacientes incluidos en lista de espera o la predicción de éxito del paciente trasplantado<sup>92</sup>. Es reseñable que podemos predecir la respuesta clínica a la resincronización cardiaca utilizando sistemas *machine learning*<sup>110</sup>.

Un excelente ejemplo de combinación de metodología *machine learning* y *deep learning* en este campo es la probabilidad de clasificación de la disfunción diastólica analizando medidas ecocardiográficas<sup>111</sup>. Esta investigación, junto con otros ejemplos que han utilizado aprendizaje no supervisado<sup>93,112</sup>, ha demostrado que estos procedimientos pueden facilitar la estandarización e interpretación de patología cardiacas complejas, como el diagnóstico de insuficiencia cardiaca con función sistólica preservada, ayudando a la toma de decisiones.

Otro estudio realizado con una gran cantidad de datos del *Swedish Heart Failure Registry* ha sido capaz de identificar 4 tipos de fenotipos con diferentes evoluciones y respuestas terapéuticas a partir de técnicas no supervisadas<sup>94</sup>.

#### 6.9.4. Imagen cardíaca

Los ejemplos de inteligencia artificial utilizando datos provenientes de imagen comienzan a marcar tendencia y serán responsables de una nueva revolución en el mundo de la imagen cardíaca. Se han empleado con éxito técnicas de *machine learning* que utilizan los datos generados de la cuantificación de imágenes cardíacas para, por ejemplo, predecir la mortalidad cardiovascular a partir de bases de datos extensas ecocardiográficas<sup>95</sup>. Se han desarrollado modelos similares *machine learning* para diferenciar los patrones ecocardiográficos de hipertrofia ventricular fisiológica propia de los atletas, de la patológica propia de la miocardiopatía hipertrófica familiar<sup>96</sup>.

Aparte de los métodos *machine learning*, la metodología *deep learning* se ha utilizado para el análisis directo sobre imágenes en varios dominios de aplicación diferentes. La mayor parte de los esfuerzos se han enfocado en tareas de segmentación del tejido y estructuras anatómicas cardíacas (por ejemplo el endocardio), que por lo general es el paso previo para realizar cualquier otro tipo de estudio, como la detección de lesiones o la clasificación por patologías. El primer estudio puramente realizado con *deep learning* data de 2013, y en él se aplicó a la segmentación del ventrículo izquierdo a partir de ecocardiografía<sup>113</sup>. Los concursos tipo *datathon* se han utilizado sistemáticamente desde 2009 como herramienta de explotación de imágenes cardíacas, principalmente con objetivos de segmentación del ventrículo izquierdo y derecho, a partir de diferentes imágenes, fundamentalmente resonancia magnética y tomografía<sup>14,100,114</sup>. Durante los últimos años, no se ha cesado en la publicación de arquitecturas de redes neuronales que han ido mejorando el estado del arte en la segmentación del corazón; tanto en 2 dimensiones<sup>115,116</sup> como en 3 dimensiones<sup>100,117,118</sup>. Incluso se han realizado estudios científicos comparativos de la determinación automática del endocardio y cálculo de la fracción eyección ventricular izquierda a partir de algoritmos de *deep learning* aplicados directamente a imágenes DICOM frente a la determinación manual, con excelente similitud y ahorro consecuente de tiempos<sup>97</sup>.

En el campo de la imagen ecocardiográfica, se han realizado ya avances para implementar una interpretación totalmente automática, a través de la identificación de

vistas, la segmentación de la imagen, la cuantificación de estructuras y funciones y la detección de patologías<sup>119</sup>.

Entre otros resultados obtenidos en los últimos años relacionados con tareas no de segmentación, podemos citar como ejemplos el cálculo de la reserva fraccional de flujo a partir de imágenes de tomografía computarizada coronaria<sup>98</sup>, la medición de calcio en arterias coronarias<sup>120</sup>, la cuantificación y caracterización de tejidos de arterias coronarias y carótida<sup>121,122</sup>, la mejora o generación de imágenes cardiacas<sup>123</sup>, la detección de estenosis y aterosclerosis<sup>124</sup> o la diferenciación de la pericarditis constrictiva de la miocardiopatía restrictiva<sup>99</sup>.

#### **6.9.5. Otras aplicaciones**

La inteligencia artificial se está utilizando ampliamente en otros dominios de aplicación diferente; hecho que refleja su amplio espectro de uso. La búsqueda de patrones abarca multitud de funcionalidades, desde la predicción del riesgo cardiovascular a partir del análisis de fondo de ojo<sup>4</sup>, la predicción del fallo renal agudo tras cirugía cardiaca<sup>125</sup>, etc. Estos patrones pueden identificar mejor los grupos de pacientes con riesgo no similar, y su incorporación a la práctica clínica puede ayudar a eliminar indeterminaciones y mejorar el resultado clínico.

**Tabla 9. Contribuciones relevantes de la inteligencia artificial en las diferentes áreas de aplicación de la cardiología**

Referencia (año)	Área	Aplicación	Técnica	Método	Resultados
Ebrahimzadeh et al. <sup>81</sup> (2018)	Arritmias	Predicción de FA paroxística a partir de la variabilidad de frecuencia cardiaca	Aprendizaje supervisado	Datos: 106 señales de 53 pares de electrocardiogramas para el entrenamiento Algoritmos: KNN, SVM, RN	Sensibilidad (100%), especificidad (95%), exactitud (98%)
Budzianowski et al. <sup>82</sup> (2018)	Arritmias	Predicción de recurrencia de FA tras crioablación de venas pulmonares	Aprendizaje supervisado	Datos: 118 pacientes con 56 señales clínicas, laboratorio y del procedimiento de cada paciente Algoritmos: GB, SVM, sobremuestreo	Identificación de 7 predictores en concordancia con análisis estadístico univariante
Eerikainen et al. <sup>83</sup> (2016)	Arritmias	Clasificación de alarmas por arritmias cardiacas en telemetría	Aprendizaje supervisado	Datos: PhysioNet/Computing in Cardiology Challenge 2015 Algoritmos: RF	VP 95%, VN 83%
Nanayakkara et al. <sup>84</sup> (2018)	Arritmias	Predicción de mortalidad hospitalaria en pacientes con parada cardiaca resucitada a partir de un registro	Aprendizaje supervisado	Datos: Registro ANZICS, 39.566 pacientes Algoritmos: RL, GB, SVM, RN, RF, en conjunto (RF, SVM, GM)	Área bajo la curva del mejor algoritmo: 0,87 (frente al 0,80 de la escala APACHE III y 0,81 ANZROD)
Yildirim et al. <sup>85</sup> (2018)	Arritmias	Detección de hasta 17 tipos de arritmias a partir del ECG	Aprendizaje supervisado	Datos: 1.000 señales de ECG de la base de datos MIT-BIH Arrhythmia database Algoritmo: RN convolucional	Exactitud (91%)



Weng et al. <sup>86</sup> (2017)	Riesgo cardiovascular	Predicción de eventos cardiovasculares a 10 años a partir de historia electrónica	Aprendizaje supervisado	Datos: 378.256 individuos; datos demográficos, historiales médicos, prescripciones médicas y tests biológicos Algoritmos: RF, RL, GB, RN	Área bajo la curva del mejor algoritmo: 0,76 (frente al 0,72 de las escalas de predicción de riesgo ACC/AHA)
Huang et al. <sup>87</sup> (2017)	Cardiopatía isquémica	Predicción de eventos cardiacos adversos mayores en pacientes con síndrome coronario agudo a partir de historia electrónica	Aprendizaje supervisado	Datos: 2.930 pacientes y 268 variables Algoritmo: SVM, RF junto a técnicas de submuestreo y sobremuestreo	Área bajo la curva del mejor algoritmo: 0,672 (mejorando significativamente a la escala GRACE, +4,8%)
Shouval et al. <sup>88</sup> (2017)	Cardiopatía isquémica	Predicción de mortalidad a 30 días tras infarto de miocardio a partir de un registro	Aprendizaje supervisado	Datos: 2.782 pacientes con 54 variables Algoritmos: Naive Bayes, RL, RF, PART, ADABOOST, ADtree	Área bajo la curva del mejor algoritmo: 0,91 (frente al 0,87 de la escala GRACE y 0,82 de la escala TIMI). Identificación 15 variables predictoras
Wallert et al. <sup>89</sup> (2017)	Cardiopatía isquémica	Predicción de supervivencia a 2 años tras infarto de miocardio a partir de datos poblacionales, clínicos y de ecocardiograma	Aprendizaje supervisado	Datos: Registro SWEDEHEART, 51.943 pacientes y 54 variables Algoritmos: SVM, LR, RF y BoostedC5.0.	Área bajo la curva del mejor algoritmo (0,84), valor predictivo negativo (97%)
Shameer et al. <sup>90</sup> (2017)	Insuficiencia cardiaca	Predicción de reingresos hospitalarios por insuficiencia cardiaca a partir de la historia electrónica	Aprendizaje supervisado	Datos: MOUNT SINAI HEART FAILURE COHORT: 1.068 pacientes hospitalizados y 4.205 variables Algoritmo Naive Bayes	Área bajo la curva 0,78 y exactitud 84% modelos previos con áreas bajo la curva de 0,6-0,7)

Inan et al. <sup>91</sup> (2018)	Insuficiencia cardiaca	Seguimiento remoto de pacientes con insuficiencia cardiaca para el ajuste de tratamiento y evitar hospitalizaciones a partir de señales electrocardiográficas de dispositivos electrónicos inteligentes	Aprendizaje no supervisado	Datos: 45 pacientes (13 descompensados) Algoritmo: Graph Similarity Score	Diferencias significativas entre grupos (compensados frente descompensados)
Medved et al. <sup>92</sup> (2017)	Insuficiencia cardiaca	Predicción de diagnóstico en pacientes en lista de espera de trasplante cardiaco	Aprendizaje supervisado	Datos: Registro UNOS (United Network for Organ Sharing), 27,444 pacientes Algoritmos: RN	F1: 0,674, 0,680, y 0,680 para la predicción a 180, 365 y 730 días, respectivamente
Sanchez-Martinez et al. <sup>93</sup> (2018)	Insuficiencia cardiaca	Caracterización del fallo cardiaco con función de eyección conservada a partir de variables ecocardiográficas	Aprendizaje no supervisado	Datos: 156 pacientes y 22 variables (20 correspondientes a trazados de velocidad por ecocardiograma) Algoritmo: <i>clustering</i>	Verdaderos positivos (71%)
Ahmad et al. <sup>94</sup> (2018)	Insuficiencia cardiaca	Predicción de evolución y fenotipado de pacientes de insuficiencia cardiaca	Aprendizaje supervisado y no supervisado	Datos: 44.886 pacientes con datos demográficos, clínicos y farmacológicos. Algoritmos: Random Forest, cluster analysis	Identificación de 4 fenotipos con diferencia significativa en evolución y respuesta terapéutica

Samad et al. <sup>95</sup> (2018)	Imagen cardiaca	Predicción de mortalidad a partir de variables ecocardiográficas e historiales electrónicos	Aprendizaje supervisado	Datos: variables cuantificadas ecocardiográficas (331,317 ecos) y variables de la historia electrónica (171,510 pacientes) Algoritmos: RF	Área bajo la curva del mejor algoritmo: 0,89 (frente al 0,61 de la escala Framingham, 0,74 escala ACC/AHA, 0,79 escala Charlson)
Narula et al. <sup>96</sup> (2016)	Imagen cardiaca	Identificación de hipertrofia ventricular fisiológica frente a patológica a partir de variables ecocardiográficas	Aprendizaje supervisado	Datos: 77 atletas y 62 miocardiopatías hipertróficas familiares Algoritmos: SVM, RF, RN	Sensibilidad (96%), especificidad (77%) superiores a las de E/A, e', <i>strain</i> longitudinal
Knackstedt et al. <sup>97</sup> (2015)	Imagen cardiaca	Estudio comparativo entre la cuantificación automática frente a manual de la función ventricular izquierda ecocardiográfica	Aprendizaje supervisado	Datos: 255 pacientes Algoritmo: cuantificación automática mediante algoritmo de aprendizaje profundo para imágenes DICOM	Factibilidad medida automática (98%). Concordancia 0,83
Tesche et al. <sup>98</sup> (2018)	Imagen cardiaca	Cálculo de la reserva fraccional de flujo a partir de imágenes de tomografía coronaria	Aprendizaje supervisado	Datos: TC coronario 85 pacientes Algoritmo: SVM	Sensibilidad (79%), especificidad (90%), especificidad (94%)
Sengupta et al. <sup>99</sup> (2016)	Otras aplicaciones	Diferenciación entre pericarditis constrictiva y miocardiopatía restrictiva a partir de variables clínicas y ecocardiográficas	Aprendizaje supervisado	Datos: clínicos y ecocardiográficos de 50 pacientes con pericarditis constrictiva y 44 con miocardiopatía restrictiva Algoritmos: RF, KNN, SVM	Área bajo la curva del algoritmo 0,96 (frente al 0,82 de la velocidad anular mitral y, 0,64 del <i>strain</i> longitudinal).

ACC/AHA: American College of Cardiology/American Heart Association; DAI: desfibrilador automático implantable; ECG: electrocardiograma; FA: fibrilación auricular; GB: *gradient boosting*; KNN: K vecinos más próximos; RF: *random forest*; RL: regresión logística; RN: red neuronal; SVM: máquinas de vectores de soporte.

## Capítulo 7

### 7. Limitaciones

Como hemos dicho anteriormente, el procesamiento de conjuntos de datos pequeños y no balanceados es un problema recurrente en la metodología *machine learning*. Diferentes pautas sugieren utilizar cohortes mucho más grandes para desarrollar nuevos modelos y validar los existentes. Sin embargo, esto no siempre es posible cuando los sujetos para un estudio prospectivo implican procedimientos invasivos, como la angiografía coronaria posterior a una ICP, que actualmente no se recomienda en la práctica clínica. Aunque utilizamos diferentes enfoques metodológicos para abordar estos inconvenientes, nuestros resultados muestran algunas deficiencias inevitables.

Al comparar nuestro algoritmo *machine learning* con los scores de riesgo clínicos existentes (PRESTO-1, PRESTO-2 y TLR), estamos comparando el rendimiento externo de estos modelos clínicos con el rendimiento interno de nuestro algoritmo *machine learning*. Por lo tanto, para evitar esta ventaja en la validación interna de nuestro algoritmo, utilizamos para las puntuaciones de los scores de riesgo clínicos los *odds ratios* calculados en sus respectivos estudios de cohorte originales y las puntuaciones enteras informadas por sus autores, siguiendo la declaración TRIPOD (*TRIPOD statement*)<sup>47</sup>.

Otro problema inherente compartido por la estadística clásica y los nuevos modelos *machine learning* es la posible presencia de sesgo en el conjunto de datos utilizado para ajustar el algoritmo. El ensayo GRACIA-3 es una buena fuente de datos, debido al uso de registros de 20 hospitales españoles diferentes. Sin embargo, los datos se recopilaron hace 10 años y están restringidos a la población de España. Los datos actuales de ICP y cohortes más amplias pueden llevar a conclusiones adicionales y una mejor generalización de las predicciones. Para este propósito, proporcionamos nuestros

algoritmos *machine learning* desarrollados como software libre para otros investigadores.

Finalmente, una limitación conocida de los algoritmos *machine learning* es la falta de una interpretación directa del modelo, además de reportar la importancia de la contribución de las variables utilizadas para entrenar el modelo. Dependiendo del algoritmo *machine learning*, la construcción de una puntuación de riesgo evaluada manualmente puede no ser posible y, por lo tanto, la implementación de estas herramientas en la práctica clínica diaria puede requerir el uso de ordenadores y aplicaciones de software.

## Capítulo 8

### 8. Conclusiones

Las conclusiones finales de este trabajo son las siguientes:

1. Desarrollar un modelo *machine learning* para predecir la aparición de reestenosis coronaria en pacientes con infarto agudo de miocardio sometidos a intervencionismo coronario percutáneo es factible; y nos permite estimar el riesgo individual de forma sencilla a partir de parámetros clínicos, bioquímicos y angiográficos disponibles en cualquier procedimiento rutinario de intervencionismo coronario.
2. Se ha construido un conjunto de datos para entrenamiento y *test* donde evaluamos 6 clasificadores *machine learning* diferentes y 3 *scores* clásicos de riesgo de sufrir reestenosis. Este conjunto de datos podría utilizarse en el futuro para valorar el rendimiento de cualquier otro clasificador de riesgo. El acceso a este conjunto de datos para la comunidad científica está disponible, si el Comité Científico del ensayo GRACIA-3 lo considera oportuno.
3. Utilizar la selección de hiperparámetros basado en un esquema de validación cruzada, que hemos denominado validación cruzada interior, nos ha permitido elegir un modelo único como mejor modelo *machine learning* y poder compararlo con otros. El esquema completo utilizado es por tanto un esquema de validación cruzada anidada, que incluye la selección de variables como un hiperparámetro más. Este esquema aunque muy costoso computacionalmente, nos ha permitido contar con un conjunto de datos mayor para entrenar, aspecto clave en ejemplos con pocos datos y desbalanceados como nuestro caso.

4. El mejor modelo *machine learning* final, con un área bajo la curva ROC de 0,77 para predecir reestenosis coronaria 12 meses tras el implante de un stent coronario, se obtuvo con un clasificador *extremely randomized trees*. Este clasificador lleva la aleatoriedad del *random forest* un paso más allá al generar valores de corte aleatorios para cada variable, escogiendo el mejor de ellos.
5. La predicción del modelo *machine learning* desarrollado con *extremely randomized trees* logró una mejora estadísticamente significativa en comparación con los discriminadores predictivos inferenciales desarrollados con modelos estadísticos clásicos: PRESTO-1 (+20%), PRESTO-2 (+19%) y TLR (+15%).
6. La aplicación de la declaración TRIPOD, en la validación externa de los *scores* de riesgo clásicos para reestenosis (PRESTO-1, PRESTO-2, y TLR), utilizando los cocientes de probabilidades calculados en sus respectivos estudios de cohortes originales, no mostró diferencia alguna respecto a los cálculos utilizando valores enteros.
7. Las curvas PR nos mostraron una imagen más precisa del rendimiento del modelo *machine learning* desarrollado. Las curvas PR para la predicción de reestenosis mostraron una exhaustividad y una sensibilidad por debajo de 0,5; sin embargo, para la predicción de no reestenosis presentaron un área bajo la curva PR de 0,96, presentando tanto alta exhaustividad como alta sensibilidad, con una ventaja significativa sobre los *scores* clásicos.
8. Para un umbral óptimo de alta exhaustividad, nuestro modelo *machine learning* permite predecir correctamente 181 (+18%) casos en comparación con el mejor *score* de riesgo clásico, TLR. Además, detecta un grupo de 32 pacientes con un riesgo de reestenosis mayor o igual al 50% que deben ser programados rutinariamente para un seguimiento más cercano. Para un umbral óptimo de valor predictivo negativo, nuestro modelo *machine learning* puede detectar un grupo de 504 pacientes con un 3% de riesgo de reestenosis, pudiendo evitar seguimientos innecesarios y reducir costes (TLR discrimina 318 pacientes y 5% de riesgo).
9. Las 6 variables con mayor impacto en el modelo predictivo *machine learning* desarrollado, según la metodología MDA, fueron: diabetes, enfermedad en dos

o más vasos coronarios, flujo TIMI post-ICP, plaquetas anormales, trombo post-ICP y colesterol total anormal.

10. Se han desarrollado calculadoras de fácil uso en entorno web y como aplicación para ejecutar en Windows o Mac. Son calculadoras que permiten calcular un *score* de riesgo de reestenosis de cualquier paciente con la información de estas 6 variables habituales de la práctica clínica diaria. Estas calculadoras además representan una estratificación del riesgo del paciente, sirviendo de apoyo al médico en la toma de decisiones.



## Bibliografía

1. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep* 2014;16:441.
2. Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development* 1959;3:210-29.
3. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521:436-44.
4. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering* 2018;2:158-64.
5. Medrano IH, Guijarro JT, Belda C, et al. Savana: Re-using Electronic Health Records with Artificial Intelligence. *International Journal of Interactive Multimedia and Artificial Intelligence* 2018;4:8-12.
6. González-Mansilla A, Martínez-Legazpi P, Prieto A, et al. Valve area and the risk of overestimating aortic stenosis. *Heart* 2019;105:911-9.
7. Yoo JS, Kim JB, Jung SH, et al. Mitral durability after robotic mitral valve repair: analysis of 200 consecutive mitral regurgitation repairs. *J Thorac Cardiovasc Surg* 2014;148:2773-9.
8. Miller RA, Pople HE, Jr., Myers JD. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med* 1982;307:468-76.
9. Melero-Alegria JI, Cascon M, Romero A, et al. SALMANTICOR study. Rationale and design of a population-based study to identify structural heart disease abnormalities: a spatial and machine learning analysis. *BMJ Open* 2019;9:e024605.
10. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med* 2019;380:1347-58.
11. Murphy KP. Machine learning: a probabilistic perspective. MIT press; 2012.
12. Barreiro-Pérez M, Tundidor-Sanz E, Martín-García A, et al. First Magnetic Resonance Managed by a Cardiology Department in the Spanish Public Healthcare System. Experience and Difficulties of an Innovative Model. *Rev Esp Cardiol (Engl Ed)* 2018;71:365-72.
13. Baumgartner CF KL, Pollefeys M, Konukoglu E. An exploration of 2 D and 3 D deep learning techniques for cardiac MR image segmentation. *CoRR* 2017;abs/1709.04496.
14. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal* 2017;42:60-88.
15. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care - Addressing Ethical Challenges. *N Engl J Med* 2018;378:981-3.

16. Gijssberts CM, Groenewegen KA, Hoefler IE, et al. Race/Ethnic Differences in the Associations of the Framingham Risk Factors with Carotid IMT and Cardiovascular Events. *PLoS One* 2015;10:e0132321.
17. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: Addressing ethical challenges. *PLoS Med* 2018;15:e1002689.
18. Sturm I, Lapuschkin S, Samek W, Muller KR. Interpretable deep neural networks for single-trial EEG classification. *J Neurosci Methods* 2016;274:141-5.
19. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Boletín Oficial del Estado (BOE), num. 294.2018.
20. Mesko B, Hetenyi G, Gyorffy Z. Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Serv Res* 2018;18:545.
21. Sanchez PL, Gimeno F, Ancillo P, et al. Role of the paclitaxel-eluting stent and tirofiban in patients with ST-elevation myocardial infarction undergoing postfibrinolysis angioplasty: the GRACIA-3 randomized clinical trial. *Circ Cardiovasc Interv* 2010;3:297-307.
22. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* 2011;12:2825-30.
23. Mikhail K, Konstantin L. Welcome to ELI5's documentation! <https://eli5.readthedocs.io/en/latest/index.html> .2016-2017.
24. Rodriguez JD, Perez A, Lozano JA. Sensitivity analysis of kappa-fold cross validation in prediction error estimation. *IEEE Trans Pattern Anal Mach Intell* 2010;32:569-75.
25. Hastie RT, Friedman J. *The Elements of Statistical Learning*. Springer; 2009.
26. Kuntz RE, Baim DS. Defining coronary restenosis - Newer clinical and angiographic paradigms. *Circulation* 1993;88:1310-23.
27. Dorado-Díaz PI, Sampedro-Gómez J, Vicente-Palacios V, Sánchez PL. Applications of artificial intelligence in cardiology. The future is already here. *Rev Esp Cardiol (Engl Ed)* 2019;03008932.
28. Breiman L. Random forests. *Machine Learning* 2001;45:5-32.
29. Safavian SR, Landgrebe D. A survey of Decision Tree Classifier methodology. *Ieee Transactions on Systems Man and Cybernetics* 1991;21:660-74.
30. Breiman L, Friedman JH, Olshen RA, Stone CJ. *Classification and regression trees*: Wadsworth & Brooks/Cole Advanced Books & Software; 1984.
31. Louppe G. Understanding random forests: From theory to practice. *arXiv preprint arXiv:1407.7502* .2014.
32. Breiman L. Arcing Classifiers. *Annals of Statistics* 1998;26:801-49.
33. Geurts P, Ernst D, Wehenkel L. Extremely randomized trees. *Machine Learning* 2006;63:3-42.
34. Bernard S, Heutte L, Adam S. Influence of Hyperparameters on Random Forest Accuracy. *Multiple Classifier Systems, Proceedings* 2009;5519:171-80.
35. Probst P, Wright MN, Boulesteix A-L. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews-Data Mining and Knowledge Discovery* 2019;9.
36. Friedman JH. Greedy function approximation: A gradient boosting machine. *Annals of Statistics* 2001;29:1189-232.

37. Hsu C-W, Chang C-C, Lin C-J. A Practical Guide to Support Vector Classification: Department of Computer Science, National Taiwan University; 2003.
38. Bishop CM. Pattern recognition and machine learning: Springer; 2006.
39. Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* 1970;12:55-67.
40. Tibshirani RJ, 58(1), 267-288. Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 1996;58:267-88.
41. van Rijn JN, Hutter F, Acm. Hyperparameter Importance Across Datasets. Kdd'18: Proceedings of the 24th Acm Sigkdd International Conference on Knowledge Discovery & Data Mining 2018:2367-76.
42. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics* 2006;7:91.
43. Krstajic D, Buturovic LJ, Leahy DE, Thomas S. Cross-validation pitfalls when selecting and assessing regression and classification models. *J Cheminform* 2014;6:10.
44. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *J Mach Learn Res* 2012;13:281-305.
45. Singh M, Gersh BJ, McClelland RL, et al. Clinical and angiographic predictors of restenosis after percutaneous coronary intervention: insights from the Prevention of Restenosis With Tranilast and Its Outcomes (PRESTO) trial. *Circulation* 2004;109:2727-31.
46. Stolker JM, Kennedy KF, Lindsey JB, et al. Predicting restenosis of drug-eluting stents placed in real-world clinical practice: derivation and validation of a risk model from the EVENT registry. *Circ Cardiovasc Interv* 2010;3:327-34.
47. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med* 2015;162:W1-73.
48. Fawcett T. An introduction to ROC analysis. *Pattern Recognition Letters* 2006;27:861-74.
49. Airola A, Pahikkala T, Waegeman W, De Baets B, Salakoski T. An experimental comparison of cross-validation techniques for estimating the area under the ROC curve. *Computational Statistics & Data Analysis* 2011;55:1828-44.
50. Bouckaert RR. Choosing between two learning algorithms based on calibrated tests. Proceedings of the 20th International Conference on Machine Learning (ICML-03) 2003:51-8.
51. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One* 2015;10:e0118432.
52. Davis J, Goadrich M. The relationship between precision-recall and ROC curves. Proceedings of the 23rd International Conference on Machine Learning (ICML-06) 2006:233-40.
53. Google Colab [Internet]. 2019, at <https://colab.research.google.com/notebooks/welcome.ipynb>.)
54. TkInter - Python Wiki [Internet]. 2019, at <https://wiki.python.org/moin/TkInter>.)
55. SEC-Working-Group-for-the-2018-ESC/EACTS-Guidelines-for-Myocardial-Revascularization, Expert-Reviewers-for-the-2018-ESC/EACTS-Guidelines-for-

Myocardial-Revascularization, SEC-Guidelines-Committee. Comments on the 2018 ESC/EACTS Guidelines for Myocardial Revascularization. *Rev Esp Cardiol (Engl Ed)* 2019;72:16-20.

56. Laham RJ, Carrozza JP, Berger C, Cohen DJ, Kuntz RE, Baim DS. Long-term (4- to 6-year) outcome of Palmaz-Schatz stenting: paucity of late clinical stent-related problems. *J Am Coll Cardiol* 1996;28:820-6.

57. Torrado J, Buckley L, Durán A, et al. Restenosis, Stent Thrombosis, and Bleeding Complications: Navigating Between Scylla and Charybdis. *J Am Coll Cardiol* 2018;71:1676-95.

58. Califf RM, Fortin DF, Frid DJ, et al. Restenosis after coronary angioplasty: an overview. *J Am Coll Cardiol* 1991;17:2B-13B.

59. Byrne RA, Joner M, Kastrati A. Stent thrombosis and restenosis: what have we learned and where are we going? The Andreas Gruntzig Lecture ESC 2014. *Eur Heart J* 2015;36:3320-31.

60. Chen MS, John JM, Chew DP, Lee DS, Ellis SG, Bhatt DL. Bare metal stent restenosis is not a benign clinical entity. *Am Heart J* 2006;151:1260-4.

61. Dangas GD, Claessen BE, Caixeta A, Sanidas EA, Mintz GS, Mehran R. In-stent restenosis in the drug-eluting stent era. *J Am Coll Cardiol* 2010;56:1897-907.

62. Alfonso F, Cuesta J. The Therapeutic Dilemma of Recurrent In-Stent Restenosis. *Circ Cardiovasc Interv* 2018;11:e007109.

63. Feinberg J, Nielsen EE, Greenhalgh J, et al. Drug-eluting stents versus bare-metal stents for acute coronary syndrome. *Cochrane Database Syst Rev* 2017;8:CD012481.

64. Mindrescu C, Brener SJ, Guerchicoff A, et al. Impact of scheduled angiographic follow-up in patients treated with primary percutaneous coronary intervention for ST-segment elevation myocardial infarction. *J Interv Cardiol* 2013;26:319-24.

65. Westcott RJ, Tcheng JE. Artificial Intelligence and Machine Learning in Cardiology. *JACC Cardiovasc Interv* 2019;12:1312-4.

66. Cui S, Li K, Ang L, et al. Plasma Phospholipids and Sphingolipids Identify Stent Restenosis After Percutaneous Coronary Intervention. *JACC Cardiovasc Interv* 2017;10:1307-16.

67. Trifacta. Data Wrangling Tools & Software [Internet]. (Accessed 2019, at <https://www.trifacta.com/>.)

68. OpenRefine [Internet]. Google News Initiative. (Accessed 2019, at <http://openrefine.org/>.)

69. Paxata. Self-Service Data Integration and Management [Internet]. <https://www.paxata.com/>. 2019,

70. Alteryx. Self-Service Data Analytics Platform [Internet]. 2019, at <https://www.alteryx.com/>.)

71. Pandas: Python Data Analysis Library [Internet]. 2019, at <https://pandas.pydata.org/>.)

72. Dplyr. A Grammar of Data Manipulation [Internet]. 2019, at <https://dplyr.tidyverse.org/>.)

73. Optimus. Agile Data Science Workflows made easy with Pyspark [Internet]. 2019, at <https://hi-optimus.com/>.)

74. Kokkinidis DG, Waldo SW, Armstrong EJ. Treatment of coronary artery in-stent restenosis. *Expert Rev Cardiovasc Ther* 2017;15:191-202.
75. Kasaoka S, Tobis JM, Akiyama T, et al. Angiographic and intravascular ultrasound predictors of in-stent restenosis. *J Am Coll Cardiol* 1998;32:1630-5.
76. Bauters C, Hubert E, Prat A, et al. Predictors of restenosis after coronary stent implantation. *J Am Coll Cardiol* 1998;31:1291-8.
77. Kastrati A, Schömig A, Elezi S, et al. Predictive factors of restenosis after coronary stent placement. *J Am Coll Cardiol* 1997;30:1428-36.
78. Hoffmann R, Mintz GS, Mehran R, et al. Intravascular ultrasound predictors of angiographic restenosis in lesions treated with Palmaz-Schatz stents. *J Am Coll Cardiol* 1998;31:43-9.
79. Al'Aref S, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J* 2019;40:1975-86.
80. de la Torre Hernández JM, Edelman ER. From Nonclinical Research to Clinical Trials and Patient-registries: Challenges and Opportunities in Biomedical Research. *Rev Esp Cardiol (Engl Ed)* 2017;70:1121-33.
81. Ebrahimzadeh E, Kalantari M, Joulani M, Shahraki RS, Fayaz F, Ahmadi F. Prediction of paroxysmal Atrial Fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on HRV signal. *Comput Methods Programs Biomed* 2018;165:53-67.
82. Budzianowski J, Hiczkiewicz J, Burchardt P, et al. Predictors of atrial fibrillation early recurrence following cryoballoon ablation of pulmonary veins using statistical assessment and machine learning algorithms. *Heart Vessels* 2019;34:352-9.
83. Eerikainen LM, Vanschoren J, Rooijackers MJ, Vullings R, Aarts RM. Reduction of false arrhythmia alarms using signal selection and machine learning. *Physiol Meas* 2016;37:1204-16.
84. Nanayakkara S, Fogarty S, Tremeer M, et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. *PLoS Med* 2018;15:e1002709.
85. Yildirim O, Plawiak P, Tan RS, Acharya UR. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Comput Biol Med* 2018;102:411-20.
86. Weng SF, Reys J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One* 2017;12:e0174944.
87. Huang Z, Chan TM, Dong W. MACE prediction of acute coronary syndrome via boosted resampling classification using electronic medical records. *J Biomed Inform* 2017;66:161-70.
88. Shouval R, Hadanny A, Shlomo N, et al. Machine learning for prediction of 30-day mortality after ST elevation myocardial infarction: An Acute Coronary Syndrome Israeli Survey data mining study. *Int J Cardiol* 2017;246:7-13.
89. Wallert J, Tomasoni M, Madison G, Held C. Predicting two-year survival versus non-survival after first myocardial infarction using machine learning and Swedish national register data. *BMC Med Inform Decis Mak* 2017;17:99.
90. Shameer K, Johnson KW, Yahi A, et al. Predictive Modeling of Hospital Readmission Rates Using Electronic Medical Record-Wide Machine Learning: A Case-Study Using Mount Sinai Heart Failure Cohort. *Pac Symp Biocomput* 2017;22:276-87.

91. Inan OT, Baran Pouyan M, Javaid AQ, et al. Novel Wearable Seismocardiography and Machine Learning Algorithms Can Assess Clinical Status of Heart Failure Patients. *Circ Heart Fail* 2018;11:e004313.
92. Medved D, Ohlsson M, Hoglund P, Andersson B, Nuges P, Nilsson J. Improving prediction of heart transplantation outcome using deep learning techniques. *Sci Rep* 2018;8:3613.
93. Sanchez-Martinez S, Duchateau N, Erdei T, et al. Machine Learning Analysis of Left Ventricular Function to Characterize Heart Failure With Preserved Ejection Fraction. *Circ Cardiovasc Imaging* 2018;11:e007138.
94. Ahmad T, Lund LH, Rao P, et al. Machine Learning Methods Improve Prognostication, Identify Clinically Distinct Phenotypes, and Detect Heterogeneity in Response to Therapy in a Large Cohort of Heart Failure Patients. *J Am Heart Assoc* 2018;7:e008081.
95. Samad MD, Ulloa A, Wehner GJ, et al. Predicting Survival From Large Echocardiography and Electronic Health Record Datasets: Optimization With Machine Learning. *JACC Cardiovasc Imaging* 2019;12:681-9.
96. Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography. *J Am Coll Cardiol* 2016;68:2287-95.
97. Knackstedt C, Bekkers SC, Schummers G, et al. Fully Automated Versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: The FAST-EFs Multicenter Study. *J Am Coll Cardiol* 2015;66:1456-66.
98. Tesche C, De Cecco CN, Baumann S, et al. Coronary CT Angiography-derived Fractional Flow Reserve: Machine Learning Algorithm versus Computational Fluid Dynamics Modeling. *Radiology* 2018;288:64-72.
99. Sengupta PP, Huang YM, Bansal M, et al. Cognitive Machine-Learning Algorithm for Cardiac Imaging: A Pilot Study for Differentiating Constrictive Pericarditis From Restrictive Cardiomyopathy. *Circ Cardiovasc Imaging* 2016;9:e004330.
100. Bernard O, Lalande A, Zotti C, et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? *IEEE Trans Med Imaging* 2018;37:2514-25.
101. Xiong Z, Liu T, Tse G, et al. A Machine Learning Aided Systematic Review and Meta-Analysis of the Relative Risk of Atrial Fibrillation in Patients With Diabetes Mellitus. *Front Physiol* 2018;9:835.
102. Li Q, Rajagopalan C, Clifford GD. Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Trans Biomed Eng* 2014;61:1607-13.
103. Figuera C, Irusta U, Morgado E, et al. Machine Learning Techniques for the Detection of Shockable Rhythms in Automated External Defibrillators. *PLoS One* 2016;11:e0159654.
104. Lyon A, Ariga R, Mincholé A, et al. Distinct ECG Phenotypes Identified in Hypertrophic Cardiomyopathy Using Machine Learning Associate With Arrhythmic Risk Markers. *Front Physiol* 2018;9:213.
105. Betancur J, Otaki Y, Motwani M, et al. Prognostic Value of Combined Clinical and Myocardial Perfusion Imaging Data Using Machine Learning. *JACC Cardiovasc Imaging* 2018;11:1000-9.

106. Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of Machine Learning Techniques for Heart Failure Readmissions. *Circ Cardiovasc Qual Outcomes* 2016;9:629-40.
107. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-Day All-Cause Readmissions in Patients Hospitalized for Heart Failure: Comparison of Machine Learning and Other Statistical Approaches. *JAMA Cardiol* 2017;2:204-9.
108. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial Intelligence in Cardiology. *J Am Coll Cardiol* 2018;71:2668-79.
109. Fallahzadeh R, Pedram M, Ghasemzadeh H. SmartSock: a wearable platform for context-aware assessment of ankle edema. *Conf Proc IEEE Eng Med Biol Soc* 2016;2016:6302-6.
110. Kalscheur MM, Kipp RT, Tattersall MC, et al. Machine Learning Algorithm Predicts Cardiac Resynchronization Therapy Outcomes: Lessons From the COMPANION Trial. *Circ Arrhythm Electrophysiol* 2018;11:e005499.
111. Salem Omar AM, Shameer K, Narula S, et al. Artificial Intelligence-Based Assessment of Left Ventricular Filling Pressures From 2-Dimensional Cardiac Ultrasound Images. *JACC Cardiovasc Imaging* 2018;11:509-10.
112. Lancaster MC, Salem Omar AM, Narula S, Kulkarni H, Narula J, Sengupta PP. Phenotypic Clustering of Left Ventricular Diastolic Function Parameters: Patterns and Prognostic Relevance. *JACC Cardiovasc Imaging* 2018.
113. Carneiro G, Nascimento JC. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. *IEEE Trans Pattern Anal Mach Intell* 2013;35:2592-607.
114. Slomka PJ, Dey D, Sitek A, Motwani M, Berman DS, Germano G. Cardiac imaging: working towards fully-automated machine analysis & interpretation. *Expert Rev Med Devices* 2017;14:197-212.
115. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal* 2016;30:108-19.
116. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal* 2017;35:159-71.
117. Dou Q, Yu L, Chen H, et al. 3D deeply supervised network for automated segmentation of volumetric medical images. *Med Image Anal* 2017;41:40-54.
118. Zheng Q, Delingette H, Duchateau N, Ayache N. 3-D Consistent and Robust Segmentation of Cardiac Images by Deep Learning With Spatial Propagation. *IEEE Trans Med Imaging* 2018;37:2137-48.
119. Zhang J, Gajjala S, Agrawal P, et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. *Circulation* 2018;138:1623-35.
120. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Isgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Med Image Anal* 2016;34:123-36.
121. Abdolmanafi A, Duong L, Dahdah N, Cheriet F. Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography. *Biomed Opt Express* 2017;8:1203-20.
122. Lekadir K, Galimzianova A, Betriu A, et al. A Convolutional Neural Network for Automatic Characterization of Plaque Composition in Carotid Ultrasound. *IEEE J Biomed Health Inform* 2017;21:48-55.

123. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. *IEEE Trans Med Imaging* 2018;37:491-503.
124. Zreik M, Lessmann N, van Hamersvelt RW, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal* 2018;44:72-85.
125. Lee HC, Yoon HK, Nam K, et al. Derivation and Validation of Machine Learning Approaches to Predict Acute Kidney Injury after Cardiac Surgery. *J Clin Med* 2018;7.



# **Anexos**

**Anexo 1. Artículo publicado. “Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí”**

**Anexo 2. Artículo enviado. "Machine Learning to Predict Stent Reestenosis Based on Daily Demographic, Clinical and Angiographic Characteristics"**

**Anexo 3. Código fuente desarrollado (Python) para la construcción del modelo *machine learning* de predicción de reestenosis**

## **Anexo 4. Código fuente calculadoras desarrolladas en Python para predecir reestenosis del stent tras ICP**

### **Anexo 4.1 Calculadora Google Colab**

### **Anexo 4.2 Calculadora Tkinter**

**Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)**

## **Anexo 6. Aprobación CEIm proyecto ANGIO**

# **Anexos**



**Anexo 1. Artículo publicado. “Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí”**

Anexo 1. Artículo publicado. PDF 11 caras.

Review article

# Applications of artificial intelligence in cardiology. The future is already here

P. Ignacio Dorado-Díaz,<sup>a,b,◇</sup> Jesús Sampedro-Gómez,<sup>a,b,◇</sup> Víctor Vicente-Palacios,<sup>a,c</sup> and Pedro L. Sánchez<sup>a,b,\*</sup>

<sup>a</sup> Servicio de Cardiología, Hospital Universitario de Salamanca-Instituto de Investigación Biomédica de Salamanca (IBSAL), Universidad de Salamanca, Salamanca, Spain

<sup>b</sup> Centro de Investigación Biomédica en Red de Enfermedades Cardiovasculares (CIBERCV), Instituto de Salud Carlos III, Madrid, Spain

<sup>c</sup> Philips Healthcare, Madrid, Spain

**Keywords:**  
Artificial intelligence  
Machine learning  
Deep learning  
Cardiology

**Palabras clave:**  
Inteligencia artificial  
Machine learning  
Deep learning  
Cardiología

## ABSTRACT

There is currently no other hot topic like the ability of current technology to develop capabilities similar to those of human beings, even in medicine. This ability to simulate the processes of human intelligence with computer systems is known as artificial intelligence (AI). This article aims to clarify the various terms that still sound foreign to us, such as AI, machine learning (ML), deep learning (DL), and big data. It also provides an in-depth description of the concept of AI and its types; the learning techniques and technology used by ML; cardiac imaging analysis with DL; and the contribution of this technological revolution to classical statistics, as well as its current limitations, legal aspects, and initial applications in cardiology. To do this, we conducted a detailed PubMed search on the evolution of original contributions on AI to the various areas of application in cardiology in the last 5 years and identified 673 research articles. We provide 19 detailed examples from distinct areas of cardiology that, by using AI, have shown diagnostic and therapeutic improvements, and which will aid understanding of ML and DL methodology.

© 2019 Sociedad Española de Cardiología. Published by Elsevier España, S.L.U. All rights reserved.

## Aplicaciones de la inteligencia artificial en cardiología: el futuro ya está aquí

### RESUMEN

Existen pocos temas de actualidad equiparables a la posibilidad de la tecnología actual para desarrollar las mismas capacidades que el ser humano, incluso en medicina. Esta capacidad de simular los procesos de inteligencia humana por parte de máquinas o sistemas informáticos es lo que conocemos hoy en día como inteligencia artificial (IA). Este artículo pretende aclarar diferentes términos que todavía nos resultan lejanos como IA, *machine learning* (aprendizaje automático, AA), *deep learning* (aprendizaje profundo, AP), *data science* o *big data*; describir en profundidad el concepto de IA y sus tipos, las técnicas de aprendizaje y la metodología que se utiliza en el AA, el análisis en imagen cardiaca con AP, la aportación de esta revolución tecnológica a la estadística clásica, sus limitaciones actuales, sus aspectos legales y, fundamentalmente, sus aplicaciones iniciales en cardiología. En este sentido se ha realizado una búsqueda detallada en PubMed de la evolución en el último lustro de las contribuciones de la IA a las diferentes áreas de aplicación en cardiología, y se ha identificado un total de 673 artículos originales. Se describen en detalle 19 ejemplos de diferentes áreas de la cardiología que utilizando IA han mostrado mejoras diagnósticas y terapéuticas, y que facilitarán la comprensión de la metodología AA y AP.

© 2019 Sociedad Española de Cardiología. Publicado por Elsevier España, S.L.U. Todos los derechos reservados.

## INTRODUCTION

Few topics are as consequential, even in medicine, as the potential ability of modern technology to develop capabilities similar to those of human beings. The ability of machines or computer systems to mimic human intelligence processes is called artificial intelligence (AI). AI is becoming highly advanced in other professional disciplines and our current challenge is to transfer all of this development to the medical field and specifically to cardiology.

This article aims to clarify for the reader various terms that may still seem foreign, not just AI, but also machine learning (ML), deep learning (DL), data science, and big data, and to describe in detail the concept of AI and its types, the learning techniques and methodology used by ML, cardiac imaging analysis with DL, the contribution of this technological revolution to classic statistics, and its current limitations, legal aspects, and, critically, initial applications in cardiology.

## DATA SCIENCE

The terms data science, big data, AI, and ML have different meanings, despite belonging to the same discipline: the analysis

\* Corresponding author: Servicio de Cardiología, Hospital Universitario de Salamanca-IBSAL-CIBERCV, P.º San Vicente 58-187, 38007 Salamanca, Spain.

E-mail address: [pedrolsanchez@secardiologia.es](mailto:pedrolsanchez@secardiologia.es) (P.L. Sánchez).

◇ Both authors contributed equally to this work as first authors.

## Abbreviations

AI: artificial intelligence  
DL: deep learning  
ML: machine learning  
NN: neural network

and extraction of information from data. However, these concepts are sometimes used interchangeably and in ambiguous ways.

The term big data was first coined in 2005 by R. Magoulas, who described it as a massive volume of data that, due to size, overwhelms traditional storage and processing software. In turn, the world of big data revolves around 7 Vs: data volume, velocity, variety, veracity, validity, volatility, and value. That is, big data generate information at high speed, controlling the integrity of the data and exhibiting huge variety in terms of nature and type (eg, text, images, videos, different systems, and different providers). In cardiology, big data would be viable within a framework of national, European, or global collaboration, involving data homogenization and sharing by hospital entities to create large-volume repositories. These large databanks can support our daily practice by helping to establish protocols, promote standardized and early disease diagnosis, aid prognosis of disease progression, and support treatment planning for our patients.<sup>1</sup>

This type of information processing would not be possible without recent technological advances in AI. AI uses tools able to identify existing patterns in data. It was defined for the first time in the 1950s and encompasses multiple subdisciplines, from expert systems and robotics to ML. It is precisely this last term that we will focus on in this review due to the major advances in the field in recent years.

The term machine learning was first coined in 1959 by AL Samuel.<sup>2</sup> Its objective is to develop algorithms enabling computer systems to make decisions and learn from the results; these systems would thus be able to learn to do something without having been explicitly provided with the programming required.

Among the many ML techniques, there is more and more interest in DL models.<sup>3</sup> These base their predictive power on artificial neural networks (NNs) and are characterized by having multiple layers of information processing (transformations) that allow them to analyze datasets with more complex patterns. These algorithms are enabling pivotal advances in speech and image recognition; the latter is particularly important in cardiology.

The data science field encompasses all aspects related to the derivation of the existing information contained in data. It involves scientific methods, processes, and systems to extract knowledge or develop a better understanding of data, normally applying novel processing techniques such as AI. Although AI and ML are often used as synonyms, they are different. AI covers a broader scope of this technological revolution and includes both ML and DL (figure 1A). The right-hand panel of figure 2 illustrates the interest triggered by these concepts in the last 3 years, with increasing interest in ML, AI, data science, and DL. *Big data* is the most common term in Spain, southern Europe, and South America, *machine learning* in the United States, Canada, the United Kingdom, northern Europe, and Australia and *deep learning* in China and Japan.

## ARTIFICIAL INTELLIGENCE

Nowadays, it is difficult to find a universal definition of what is known as AI. The term itself is often applied to the field of

computer science that tries to mimic human cognitive processes, learning capacity, and knowledge storage. Other definitions are broader and encompass the understanding and construction of intelligent entities, generally understood as computer software. In terms of types of task, AI can refer to automated systems capable of, for example, translating a document, recognizing people by their facial features, or driving a car. However, AI is not limited to imitating human tasks: in some cases, it is able to beat the best expert in a field by making decisions with lower error rates than humanly possible or by identifying patterns imperceptible to the human eye.<sup>4</sup> Therefore, AI allows information to be analyzed with a distinct approach from the traditional one. Thus, when responding to an event, we are no longer limited to describing the available information. AI allows other questions to be answered: what happened? (diagnosis), what will happen? (prediction), and what should I do? (prescription).

The most important application areas of AI in the health care field include the following: automatic speech recognition and natural language processing; prediction, recommendation, and diagnostic algorithms; computer vision and image analysis; robotics; and AI and expert systems.

### Automatic speech recognition and natural language processing

The aim of these disciplines is to develop mechanisms for communication between people and machines using natural language. In the medical field, automatic speech recognition is already being used to record patients' clinical information.<sup>5</sup> In addition, natural language processing is allowing, as already seen in several examples in Spain in the field of cardiology, disease classification and the selection of the most appropriate cohort for a clinical study by analyzing medical record registries.<sup>6</sup>

### Prediction, recommendation, and diagnostic algorithms

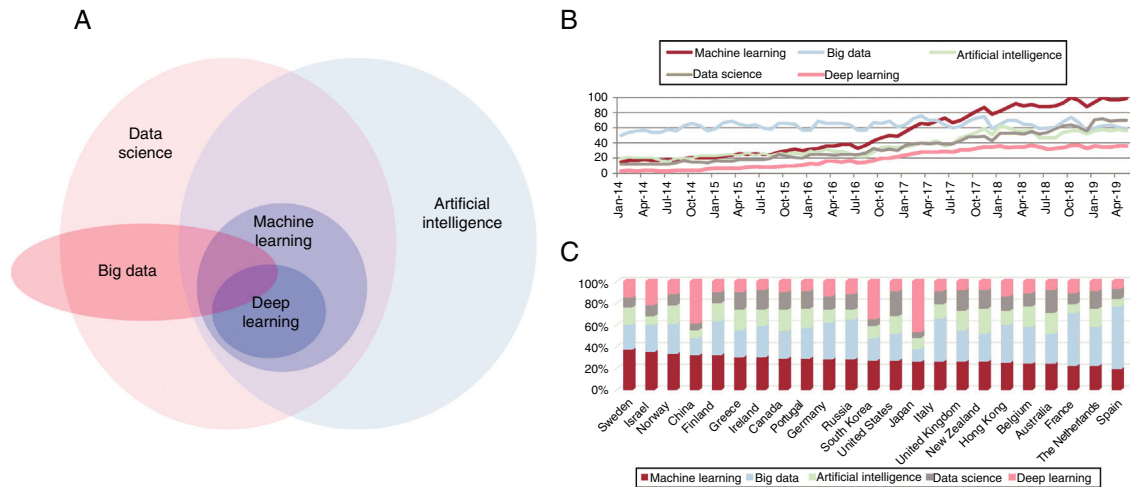
This is probably the most mature area in the worlds of medicine and cardiology and of ML and DL. Interest primarily lies both in the automation of repetitive tasks, such as the evaluation of diagnostic tests, and in knowledge generation through clinical data analysis. A major part of this article focuses on describing the methodology to be followed in this type of study and introducing use cases in cardiology.

### Computer vision and image analysis

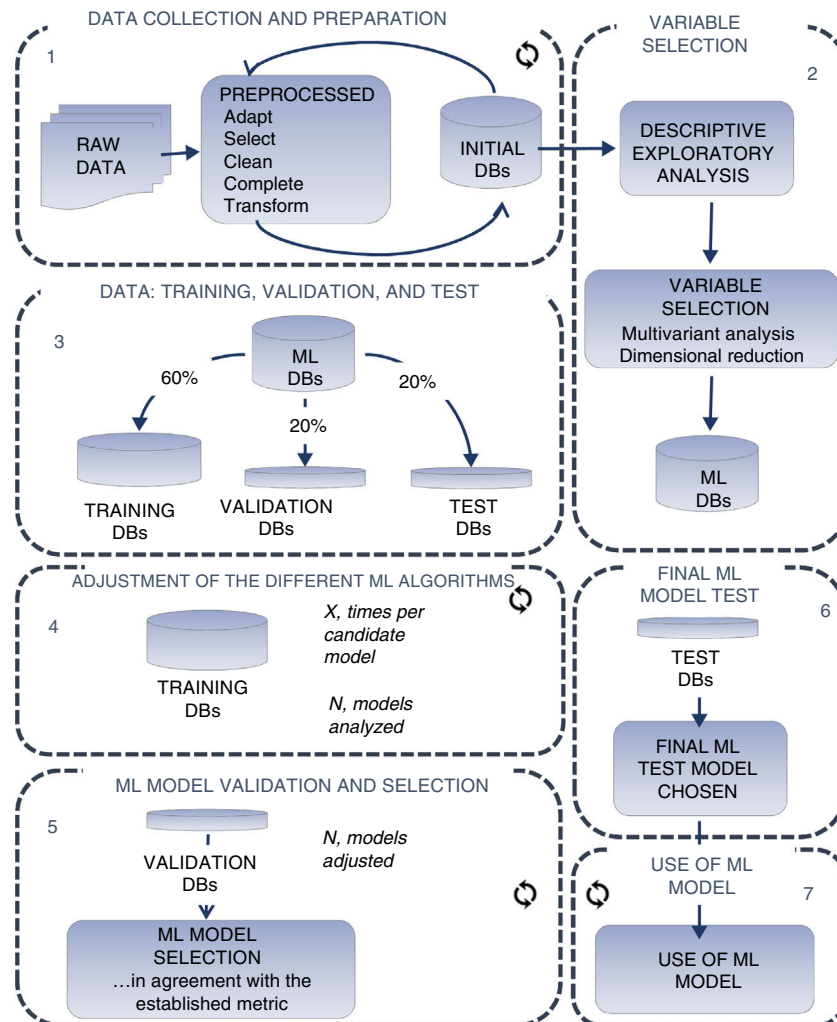
This scientific discipline encompasses methods to acquire, process, analyze, and understand real-world images in order to produce numerical or symbolic information that can be processed by a computer. These techniques have undergone a major revolution in recent years due to the application of DL algorithms and it is one of the disciplines making the greatest contribution to medicine today, and therefore to cardiology, as detailed below.

### Robotics and artificial intelligence

In robotics and AI, the objective is to build physical systems with intelligent behavior. This field has undergone many years of development but is experiencing a boom in the various areas it includes (eg, automatic speech recognition, computer vision). In the cardiovascular area, its development has been ongoing for years in the field of surgery, with a clear example being the Da Vinci surgical system.<sup>7</sup>



**Figure 1.** A: Venn diagram of the most commonly used terms in the data science disciplines. B: total searches (source: Google Trends) in the last 5 full years of terms related to artificial intelligence and data science; the vertical axis of the diagram represents the proportion of a topic with respect to the total number of searches on the topics. C: the most searched term in each country in the same period.



**Figure 2.** Application methodology of a machine learning model. DB, database; ML, machine learning.

## Expert systems

In AI, an expert system is software that emulates the decision-making ability of a human expert. These systems are based on rules or even clinical cases. For some time, expert medical systems have been available that try to simulate the reasoning of specialists and provide the probable diagnosis and optimal patient management, even in cardiology.<sup>8</sup>

## MACHINE LEARNING: TECHNIQUES

ML is an area of data science that forms part of what is known today as AI. ML involves the creation of systems that learn automatically; they understand by autonomously “learning” the ability to recognize complex patterns, without the need for human intervention and in datasets of any kind, including numerical, visual, and textual. As the experience of these systems increases, that is, as they are provided with new data, their performance improves to a point that may even exceed the human capacity in that task.

Although there are several ML techniques, they are usually grouped into 2 types: supervised learning and unsupervised learning.

Supervised learning techniques are undoubtedly the most widely used methods in ML and those with the best results. These procedures rely on a dataset from which the response variable to be predicted (eg, diagnosis, parameter, segmentation) is determined through the correct labeling of examples. Depending on the type of prediction, classification algorithms or regressive algorithms are used. In the first case, the aim is to identify 2 or more classes using a series of variables. In contrast, regressive algorithms seek to approximate a continuous value as much as possible.

In unsupervised learning techniques, there is no information on the variable to be predicted. These techniques must learn from the relationships among the elements of a dataset and classify them without relying on labels or categories. To do this, they look for structures, patterns, or characteristics in the source data that can be reproduced in new datasets. For this task, clustering methods are the most commonly used approaches.

## MACHINE LEARNING: METHODOLOGY

The construction of an ML model does not simply involve the application of a learning algorithm to a database, but is a whole process,<sup>9,10</sup> which usually includes the steps shown in figure 2.

The first stages are common to conventional statistics; raw data are converted into information with structured data (preprocessed), and an initial database is constructed (step 1 of figure 2). From this database, a descriptive and exploratory analysis is performed to identify and select the most significant variables; these variables will be directly applied to the ML algorithms (step 2 of figure 2).

**Table 1**  
Most used algorithms in machine learning

Algorithm	Type	Description
Random forest	Ensemble method	Combination of uncorrelated decision trees
Gradient boosting	Ensemble method	Combination of stepwise decision trees
Logistic regression	Regression	Regression analysis used to predict the result from a categorical variable
Support vector machines	Supervised classifier	Classification through construction of separating hyperplanes
k-nearest neighbors	Supervised classifier	Estimation of the density function of predictor variables according to class
Linear discriminant analysis	Linear discriminant	Fisher's linear discriminant generalization
Naive Bayesian classifier	Supervised probabilistic classifier	Probabilistic classifier based on Bayes theorem

The next stage, now specific to ML techniques, is the division of the dataset into 3 subsets—training, validation, and testing—, typically 60%, 20%, and 20%, respectively (step 3 of figure 2).

The training set is the dataset used to adjust the various ML algorithms selected (step 4 of figure 2). A wide range of classification and regression algorithms is available<sup>11</sup> (table 1), from more classic linear techniques such as logistic regression or linear discriminant analysis to more modern ones such as aggregation-type or bagging (random forest) algorithms and packaging or boosting (xgboost) algorithms. In addition, subsampling and oversampling techniques can be used when, as often occurs in health data, one category has many more cases than another. In theory, no algorithm is better than another; its ability to make a good adjustment will depend on the characteristics of the data (eg, number of variables, linearity, normality, missing values, and continuous or categorical variables).

Once the adjustment has been made, the validation data subset is used to evaluate the quality of the model (step 5 of figure 2). To do this, the aim is to maximize the metric of greatest interest in our particular case, such as area under the ROC curve, precision, sensitivity, and accuracy (figure 3). It is common for this training-validation process to be repeated a number of times while randomizing both subsets, which is known as k-fold cross-validation. The objective is to optimize the internal parameters of the algorithm used, evaluate the robustness of the model, and determine whether the model is subadjusting or overadjusting the data, trying to find a balance between the 2 scenarios.

Once the final model has been constructed, the test data subset is used to verify that the final ML model behaves as expected with data that has not been used for its construction or validation (step 6 of figure 2). If this result differs from that obtained in the validation set, the dataset used for training is probably insufficient and should be expanded if a reliable estimator is needed before its generalized use (step 7 of figure 2).

For the implementation of ML, the most commonly used open source programming languages are currently Python and R. Both platforms can avail of libraries such as scikit-learn (Python) and caret (R) with implementations of the most widely used techniques and algorithms.

## DEEP LEARNING AND CARDIAC IMAGING

Image analysis is the field with the most rapid advances in AI and is therefore highly relevant in cardiology. The daily analysis of various cardiological images can be tedious and time consuming. However, our day-to-day work can already be improved by the various tools, based on NNs, available for automatic image processing.

NNs<sup>3</sup> are a type of AI algorithm that is analogous to the learning process occurring in the neurons of the brain. Since its development as a computational model, training algorithms and network architectures have emerged that have considerably improved

learning accuracy and efficiency, functioning with ever smaller amounts of training data. When these architectures consist of numerous layers of neurons, the term DL is used. Currently, the most commonly applied NNs include convolutional NNs, recursive NNs, generative adversarial networks, and U-nets, each with different uses and architectural subtypes. NNs are also very flexible and can be used in supervised, unsupervised, and reinforcement learning contexts. For their implementation, various open source platforms are available, such as TensorFlow, Pytorch, Keras, and Caffe. The disadvantages of DL techniques include their high computational cost. In addition, a high degree of expertise is required for their correct adjustment and, in the case of supervised learning, a set of manually annotated images is required, which can sometimes be extremely expensive.

Within the field of cardiac imaging, AI techniques have several objectives. These include segmentation and identification of the different structures of the heart (figure 4),<sup>12,13</sup> classification of images cataloged with different conditions, lesion detection and segmentation, image registration, and classification of tissues from histological images. Other tasks are related to the generation of artificial images that are as realistic as possible.

Research applying DL techniques to cardiological imaging has burgeoned in the last 3 years. However, the use of DL techniques in medical imaging has made more of a mark in other disciplines, such as neurology and pneumology.<sup>14</sup> Among the challenges likely to be explored in the coming years are combined analyses of different imaging sources, the incorporation of clinical data and medical reports, and the study of temporal changes in cardiac images.

### LIMITATIONS OF MACHINE LEARNING AND DEEP LEARNING

The main characteristic of AI in terms of ML or DL models is that learning is based on the identification of patterns in datasets. This philosophy is simultaneously a strong point, because computers are extremely efficient and precise at finding such patterns when they exist, and a limitation, for several reasons. First, because the amount of data required to obtain an accurate model can be substantial. This can be a problem in medicine, where the implementation of automatic data collection systems is just beginning, such systems must meet established legal and ethical criteria, and some rare diseases inevitably have few studies. In addition, even with algorithms with good results for a dataset, ML and DL models suffer from an inability to correctly detect and classify cases that they have not previously seen. Along these lines, the reliability and quality of the data source are essential for an

algorithm to be realistic and correct. If the dataset used to adjust the model is itself biased, the model may not be well generalized to other populations.<sup>15</sup> This problem is nothing new in our daily work because we accept risk scales based on data from other populations, such as the Framingham score for cardiovascular risk, resulting in erroneous results that overestimate or underestimate the risk upon application to other populations.<sup>16</sup> Initiatives are underway to standardize the requirements that must be met by an AI algorithm for its possible practical application.<sup>17</sup> For all of these reasons, it seems clear that particular effort will be needed for the integration of datasets from different populations.

Another important limitation is the opacity and interpretability of the ML models, particularly the DL models. These techniques are used as “black boxes”, which are fed inputs to obtain an output, namely a prediction. Thus, ML offers us answers in the form of predictions, but not a biological explanation. Although there are various methodologies to interpret the results of a model and verify its correct functioning through analysis of the specific weights of parameters or variables and to highlight the most discriminating parts or the individual explanation of each prediction,<sup>18</sup> it is not currently possible to know exactly why the more complex ML and DL models make a certain decision. This limitation makes it difficult to generate knowledge and to apply AI to an area as critical as medicine.

### MACHINE LEARNING AND CLASSIC STATISTICS: SIMILARITIES AND DIFFERENCES

Although these 2 specialties have different origins, they also have common features, despite their distinct purposes.

The cornerstone of classic statistics is inference. In inference, the observations in a sample are extended to the entire population, generally creating a mathematical model that defines the relationship among variables. The strength of ML lies in a prediction based on the available information, without the need to know the mechanisms relating the variables to each other. These characteristics are not exclusive because, to a greater or lesser extent, both disciplines use inferential techniques to improve their results (ML) or predictive algorithms to corroborate inferential conclusions (statistics).

Compared with statistical techniques, those of ML do not require prior suppositions about the variables, have the ability to handle cases with missing data, and improve their reliability when large volumes of data are available by detecting complex relationships among variables.

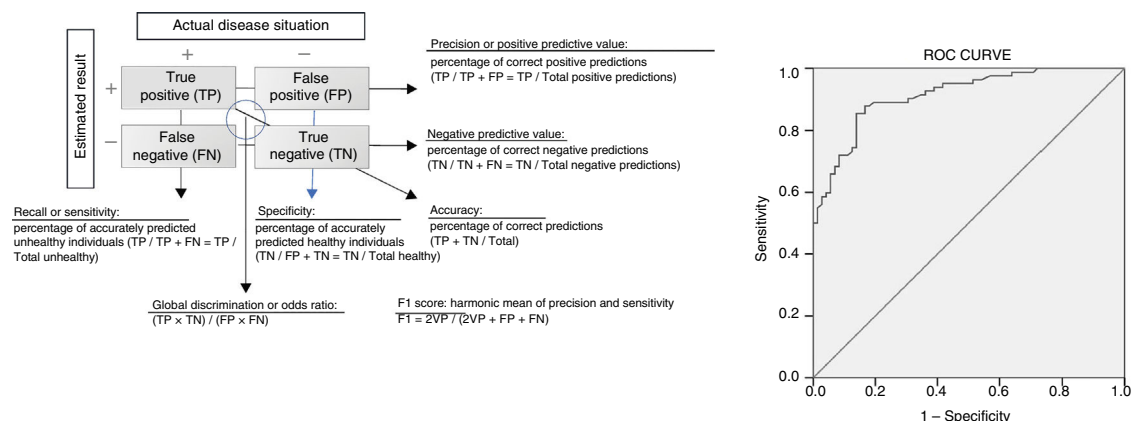
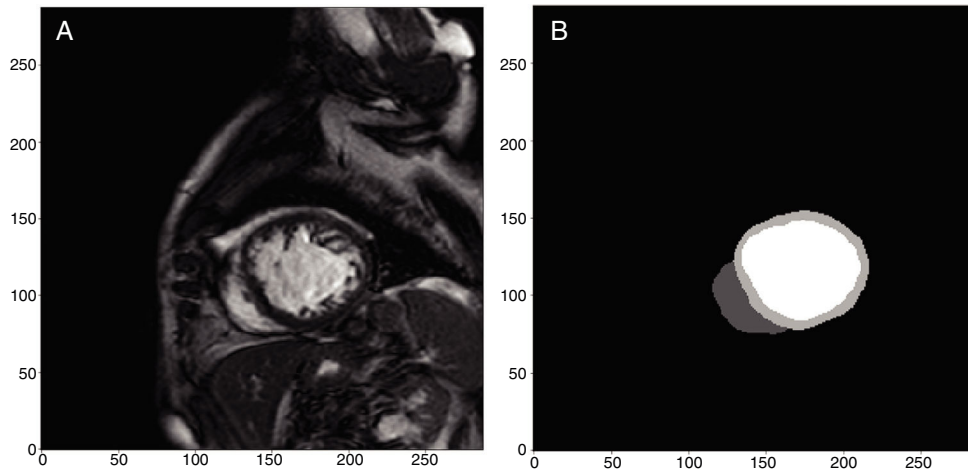


Figure 3. Metrics commonly used to evaluate supervised classification algorithms.



**Figure 4.** Example of automatic segmentation and identification of the left and right ventricle through deep learning performed in our department from images obtained with a 1.5-T Philips Achieva resonance system.<sup>12</sup> From original images (A), the free neural network<sup>13</sup> was able to identify and segment the left and right ventricles (B, white color for the left ventricle, light gray color for the myocardium of the left ventricle, and dark gray color for the right ventricle).

## LEGAL AND ETHICAL ASPECTS OF ARTIFICIAL INTELLIGENCE

When AI is applied to the world of medicine, issues can arise that are not strictly related to the quality of the ML and DL algorithms. These are connected to legal problems associated with the automatic processing of personal data or how to apply ML in day-to-day practice.

Practically all AI techniques studied to date require a certain amount of data for the training and validation of predictive models. In our case, the information used is particularly sensitive because it generally involves patients' personal and clinical data. Therefore, the proper use of this information from a legal point of view is a critical aspect.

Personal data protection is a fundamental right enshrined in the Spanish Constitution. However, such data may be used under certain conditions if a service is being provided to society. If both the organization collecting the information and that processing the data are fiscally resident in Spain, they must, since 2018, comply with Organic Law 3/2018 of December 5th, 2018, on the Protection of Personal Data and the Guarantee of Digital Rights.<sup>19</sup> According to this legislation, the use of personal data is allowed for statistical purposes, as long as the results are aggregated. It is also necessary to adopt measures to anonymize and protect these data, and a number of rights of the individuals concerned is recognized, such as to be informed, have access, and be allowed rectification. Therefore, given that the use of cloud computing services is very common in AI due to their flexibility and greater computing power, the largest providers have adapted their data-processing policies to provide a completely secure service. In addition, these companies adhere to the CISPE (Cloud Infrastructure Services Providers in Europe) Code of Conduct to guarantee that their data protection standards comply with the legislation. Likewise, in situations where an AI model is used as a diagnostic or decision-making tool, the patient must be informed and provide consent.

The subtle discrimination inherent in the provision of health services can also be a problem when AI models are being developed and applied.<sup>15</sup> For example, age is considered when scarce resources are being rationed, such as heart transplants. Such a consideration can lead to self-fulfilling prophecies: if physicians withdraw care for patients due to their advanced age, ML systems may conclude that the care of older patients is always fatal. On the other hand, it is also possible for ML models to help to resolve disparities in the provision of medical care if algorithms can be constructed to compensate for known biases or identify pressing areas of research.

There is broad discussion about how AI will affect the workflow of medical staff. AI might not eliminate jobs, but rather displace medical practice tasks. Routine and tiring jobs could be performed by machines to free up time and allow medical professionals to carry out more complex and sensitive tasks. However, the use of automatic diagnostic tools can lead to problems of bias in decision-making and in the assignment of responsibility in the case of error. Physicians may be inclined to review a diagnostic test in a more relaxed way if an algorithm has previously returned a negative result. The design of decision-making systems that involve both machines and humans is a crucial aspect for AI, and there are intermediate models between conventional clinical practice and fully automated systems, each with different characteristics of cost and time efficiency, risk, and interpretability.<sup>20</sup> The implementation of these systems will be an area with a fundamental role for clinicians as validators.

## ARTIFICIAL INTELLIGENCE IN CARDIOLOGY

Examples of AI using ML or DL are becoming more and more common in cardiology.<sup>21</sup> This section presents the evolution of the contributions of AI to the different application areas of cardiology (figure 5).

The supplementary material for this article includes a list of the contributions of the techniques described in several areas of cardiology in the last 5 full years and from January to March 2019, as well as a detailed description of the publications that we consider of most interest and that will help to explain the ML and DL models (table 2).<sup>22-40</sup> Many of the publications reviewed are the result of common initiatives in AI, such as challenges promoted in conferences<sup>24,41</sup> and datathons. In datathons, which are public and open competitions, people interested in the subject, from beginners to experts, work together with common databases to achieve an established goal. Participants must find a solution to a defined objective in the datasets provided. The most accurate and creative solutions are awarded a prize and made publicly accessible.

### Cardiac arrhythmias

One of the most widespread applications of ML in cardiology is the prediction of cardiac arrhythmias. Numerous studies address

predictive models for atrial fibrillation development, including paroxysmal, due to its impact and clinical implications, using supervised learning with predictive ML systems composed of different subprocesses: signal preprocessing, extraction of significant variables, and classification algorithms.<sup>22,42</sup> Similar ML models have been developed to improve telemonitoring alarm management<sup>24</sup> and to predict the occurrence of ventricular arrhythmias,<sup>43,44</sup> the response to invasive ablation procedures such as cryoablation,<sup>2,3</sup> and even mortality after resuscitated cardiac arrest.<sup>25</sup> DL techniques have been used with great success to detect distinct types of arrhythmias through direct analysis of images or electrocardiographic signals.<sup>26</sup> Another application has been to use unsupervised learning to identify phenotypes to classify hypertrophic cardiomyopathies with different arrhythmic risk.<sup>45</sup>

### Ischemic heart disease

AI has been applied to standard electronic medical records from primary care to predict the risk of cardiovascular disease in the general population and has been shown to be superior to the traditionally used risk scales.<sup>27</sup> In addition, supervised learning techniques have been applied to the prognostic prediction of stable ischemic heart disease,<sup>46</sup> coronary syndrome,<sup>28</sup> and mortality of patients with myocardial infarction by analyzing the results of individual hospitals<sup>29</sup> or large registries such as SWEDEHEART.<sup>30</sup> The results are disparate due to the sizes of the samples, which shows that ML techniques obtain better results with large sample sizes.

### Heart failure

ML systems could allow the optimization of avoidable hospitalizations due to heart failure by more accurately identifying patients susceptible to cardiac decompensation after hospital discharge than classic risk scales<sup>31,47</sup>; these results contradict initial experience in this field<sup>48</sup> and highlight the need to adjust the methodologies of the systems.<sup>49</sup> In addition, initial work has addressed the usefulness of AI as a management system in the telemonitoring of patients with HF. Studies have shown that ML is feasible<sup>50</sup> and that it might improve the clinical course of these patients.<sup>32</sup> Another important area is heart transplant, with ML systems applied to predict the probability of death or transplant in

patients on the waiting list or transplant success.<sup>33</sup> Notably, the clinical response to cardiac resynchronization can be predicted using ML systems.<sup>51</sup>

An excellent example of a combination of ML and DL methodologies in this field is the prediction of diastolic dysfunction through analysis of echocardiographic data.<sup>52</sup> This research, along with other examples that have used unsupervised learning,<sup>34,53</sup> shows that these procedures can facilitate the standardization and interpretation of complex heart diseases, such as the diagnosis of heart failure with preserved systolic function, enhancing decision-making.

Another study conducted with unsupervised techniques and a large amount of data from the Swedish Heart Failure Registry was able to identify 4 phenotypes with different clinical courses and therapeutic responses.<sup>35</sup>

### Cardiac imaging

Examples of AI using imaging data are beginning to become popular and will be responsible for a new revolution in the world of cardiac imaging. ML techniques using data generated from cardiac imaging quantification have been successfully used to, for example, predict cardiovascular mortality from extensive echocardiographic databases.<sup>36</sup> Similar ML models have been developed to differentiate the echocardiographic patterns of the physiological ventricular hypertrophy typical of athletes from the findings of familial hypertrophic cardiomyopathy.<sup>37</sup>

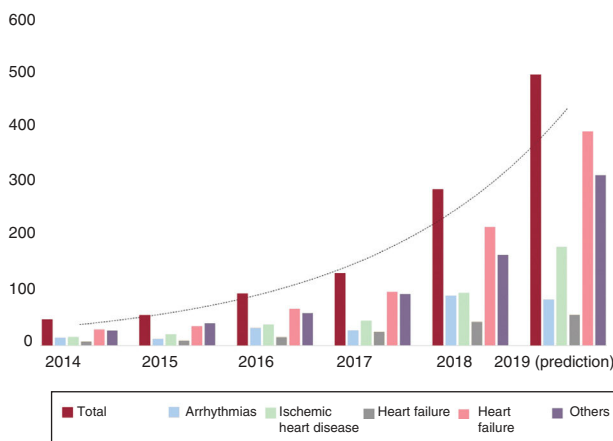
Apart from the ML methods, DL has been used to directly analyze images in several different application domains. Most of the efforts have focused on segmentation tasks for cardiac tissue and anatomical structures (eg, the endocardium; [figure 4](#)), which is usually the previous step in other study types, such as those involving injury detection or disease classification. The first study purely conducted with DL dates from 2013 and involved segmentation of the left ventricle from echocardiography.<sup>54</sup> Since 2009, datathon-style contests have systematically used a cardiac imaging tool, mainly with the objective of left and right ventricular segmentation and based on different images, typically magnetic resonance imaging and computed tomography.<sup>14,41,55</sup> In recent years, there have been repeated publications on NN architectures that have improved the state of the technique in terms of heart segmentation, in both 2-dimensions<sup>56,57</sup> and 3-dimensions.<sup>41,58,59</sup> There have even been comparative scientific studies of the automated determination of the endocardium and calculation of the left ventricular ejection fraction from DL algorithms directly applied to DICOM images vs manual tracing, with excellent concordance and speed.<sup>38</sup>

In the field of echocardiographic imaging, advances have already been made to implement a fully automatic interpretation, through the identification of viewpoints, image segmentation, the quantification of structures and functions, and disease detection.<sup>60</sup>

Other results related to nonsegmentation tasks obtained in recent years include calculation of the fractional flow reserve from coronary computed tomography images,<sup>39</sup> measurement of calcium in coronary arteries,<sup>61</sup> quantification and characterization of coronary and carotid artery tissues,<sup>62,63</sup> improvement or generation of cardiac images,<sup>64</sup> detection of stenosis and atherosclerosis,<sup>65</sup> and differentiation of constrictive pericarditis from restrictive cardiomyopathy.<sup>40</sup>

### Other applications

AI is being widely used in other application domains, which reflects its versatility. The search for patterns encompasses a multitude of functionalities, from the prediction of cardiovascular



**Figure 5.** Temporal changes in the publications indexed in PubMed on artificial intelligence, machine learning, and deep learning according to the area of interest in cardiology. The details of the publications by area of interest and the search methodology are described in the supplementary material.



**Table 2**  
 Relevant contributions of artificial intelligence to the various application areas of cardiology

Reference, y	Area	Application	Technique	Method	Results
Ebrahimzadeh et al. <sup>22</sup> (2018)	Arrhythmias	Prediction of paroxysmal AF from heart rate variability	Supervised learning	Data: 106 signals from 53 pairs of electrocardiograms for training Algorithms: KNN, SVM, NN	Sensitivity (100%), specificity (95%), accuracy (98%)
Budzianowski et al. <sup>23</sup> (2018)	Arrhythmias	Prediction of AF recurrence after pulmonary vein cryoablation	Supervised learning	Data: 118 patients with 56 clinical, laboratory, and procedural variables from each patient Algorithms: GB, SVM, oversampling	Identification of 7 predictors confirming univariate statistical analysis
Eerikainen et al. <sup>24</sup> (2016)	Arrhythmias	Classification of cardiac arrhythmia alarms in telemetry	Supervised learning	Data: PhysioNet/Computing in Cardiology Challenge 2015 Algorithm: RF	True positive, 95%; False positive, 83%
Nanayakkara et al. <sup>25</sup> (2018)	Arrhythmias	Prediction of hospital mortality in patients with resuscitated cardiac arrest from a registry	Supervised learning	Data: ANZICS registry, 39 566 patients Algorithms: LR, GB, SVM, NN, RF, combination (RF, SVM, GM)	Area under the curve of the best algorithm: 0.87 (vs 0.80 from the APACHE III scale and 0.81 from the ANZROD)
Yildirim et al. <sup>26</sup> (2018)	Arrhythmias	Detection of up to 17 types of arrhythmias from ECG	Supervised learning	Data: 1000 ECG signal fragments from the MIT-BIH Arrhythmia database Algorithm: convolutional NN	Accuracy (91%)
Weng et al. <sup>27</sup> (2017)	Cardiovascular risk	Prediction of cardiovascular events at 10 years from electronic medical records	Supervised learning	Data: 378 256 individuals; demographic data, medical records, medical prescriptions, and biological tests Algorithms: RF, LR, GB, NN	Area under the curve of the best algorithm: 0.76 (vs 0.72 from the ACC/AHA risk prediction scales)
Huang et al. <sup>28</sup> (2017)	Ischemic heart disease	Prediction of major cardiac adverse events in patients with acute coronary syndrome from electronic medical records	Supervised learning	Data: 2930 patients and 268 variables Algorithms: SVM and RF together with subsampling and oversampling techniques	Area under the curve of the best algorithm: 0.672 (significant improvement vs the GRACE scale, + 4.8%)
Shouval et al. <sup>29</sup> (2017)	Ischemic heart disease	30-day mortality prediction after myocardial infarction from a registry	Supervised learning	Data: 2782 patients with 54 variables Algorithms: naive Bayes, LR, RF, PART, ADABOOST, ADTree	Area under the curve of the best algorithm: 0.91 (vs 0.87 from the GRACE scale and 0.82 from the TIMI scale). Identification of 15 predictor variables
Wallert et al. <sup>30</sup> (2017)	Ischemic heart disease	2-year survival prediction after myocardial infarction from populational, clinical, and echocardiographic data	Supervised learning	Data: SWEDHEART registry, 51 943 patients and 54 variables Algorithms: SVM, LR, RF, and Boosted C5.0	Area under the curve of the best algorithm (0.84), negative predictive value (97%)
Shameer et al. <sup>31</sup> (2017)	Heart failure	Prediction of rehospitalizations for heart failure from electronic medical records	Supervised learning	Data: MOUNT SINAI HEART FAILURE COHORT: 1068 hospitalized patients and 4205 variables Algorithm: naive Bayes	Area under the curve of 0.78 and accuracy of 84% (previous models with areas under the curve of 0.6–0.7)
Inan et al. <sup>32</sup> (2018)	Heart failure	Remote monitoring of patients with heart failure to enable treatment adjustment and avoid hospitalizations from electrocardiographic variables from intelligent electronic devices	Unsupervised learning	Data: 45 patients (13 decompensated) Algorithm: graph similarity score	Significant differences between groups (compensated vs decompensated)
Medved et al. <sup>33</sup> (2017)	Heart failure	Diagnostic prediction in patients on the heart transplant waiting list	Supervised learning	Data: UNOS Registry (United Network for Organ Sharing), 27.444 patients Algorithm: NN	F1 scores of 0.674, 0.680, and 0.680 for prediction at 180, 365, and 730 days, respectively

**Table 2** (Continued)

Relevant contributions of artificial intelligence to the various application areas of cardiology

Reference, y	Area	Application	Technique	Method	Results
Sanchez-Martinez et al. <sup>34</sup> (2018)	Heart failure	Characterization of heart failure with preserved ejection function from echocardiographic variables	Unsupervised learning	Data: 156 patients and 22 variables (20 corresponding to velocity traces on echocardiography) Algorithm: clustering	True positive (71%)
Ahmad et al. <sup>35</sup> (2018)	Heart failure	Prediction of course and phenotyping of heart failure patients	Supervised and unsupervised learning	Data: 44 886 patients with demographic, clinical, and pharmacological data Algorithms: RF and cluster analysis	Identification of 4 phenotypes with significantly different courses and therapeutic responses
Samad et al. <sup>36</sup> (2018)	Cardiac imaging	Mortality prediction from echocardiographic variables and electronic medical records	Supervised learning	Data: quantified echocardiographic variables (331.317 echoes) and electronic medical record variables (171.510 patients) Algorithms: RF	Area under the curve of the best algorithm: 0.89 (vs 0.61 from the Framingham scale, 0.74 from the ACC/AHA scale, 0.79 from the Charlson scale)
Narula et al. <sup>37</sup> (2016)	Cardiac imaging	Identification of physiological vs pathological ventricular hypertrophy from echocardiographic variables	Supervised learning	Data: 77 athletes and 62 people with familial hypertrophic cardiomyopathy Algorithms: SVM, RF, NN	Sensitivity (96%) and specificity (77%) higher than those of E/A, $\dot{e}$ , and longitudinal strain
Knackstedt et al. <sup>38</sup> (2015)	Cardiac imaging	Comparative study between automatic vs manual quantification of echocardiographic left ventricular function	Supervised learning	Data: 255 patients Algorithm: automatic quantification using a deep learning algorithm for DICOM images	Feasible automated measurement (98%). Concordance, 0.83
Tesche et al. <sup>39</sup> (2018)	Cardiac imaging	Calculation of the fractional flow reserve from coronary tomography images	Supervised learning	Data: coronary CT from 85 patients Algorithm: SVM	Sensitivity (79%) and specificity (93%)
Sengupta et al. <sup>40</sup> (2016)	Other applications	Differentiation between constrictive pericarditis and restrictive cardiomyopathy from clinical and echocardiographic variables	Supervised learning	Data: clinical and echocardiographic of 50 patients with constrictive pericarditis and 44 with restrictive cardiomyopathy Algorithms: RF, KNN, SVM	Area under the curve of the algorithm 0.96 (vs 0.82 from mitral annular velocity and 0.64 from longitudinal strain)

ACC/AHA, American College of Cardiology/American Heart Association; AF, atrial fibrillation; ECG, electrocardiography; GB, gradient boosting; KNN, k-nearest neighbors; LR, logistic regression; NN, neural network; RF, random forest; SVMs, support vector machines.

risk through fundus analysis<sup>4</sup> to that of acute renal failure after cardiac surgery, for example.<sup>66</sup> These patterns can better identify groups of patients with dissimilar risk, and their incorporation into clinical practice can help to eliminate uncertainties and improve clinical outcomes.

## CONCLUSIONS

Although AI is often viewed as a futuristic and distant concept, the truth is that this technology is already in use in all types of areas, including cardiology. Due to the digitization of large amounts of data, the development of ML algorithms, and improvements in computer power in recent decades, AI can provide excellent opportunities for task automation, the application of precision medicine, or research progress through the detection of complex patterns in medical databases.<sup>67</sup> A particular case is that of medical image analysis, with DL techniques having undergone a true revolution and their application to the area of cardiology already yielding excellent results. However, there is still a long way to go before these techniques can be widely applied to clinical practice. Essential prerequisites are large databases with high-quality information and the evaluation and integration of AI into realistic clinical contexts; therefore, the understanding of AI

and its applications in our field is essential for the present and future development of cardiology.

## ACKNOWLEDGMENTS

The authors of this article would like to thank Antonio Sánchez, Rafael Vidal, Manuel Jiménez-Navarro, and Purificación Galindo for reviewing the manuscript prior to its submission to *Revista Española de Cardiología*. Their comments and advice were very helpful and no doubt contributed to the drafting of this article.

## FUNDING

This work was funded by the Carlos III Health Institute (Ministry of Science, Innovation, and Universities) within the CIBERCV, through which P.I. Dorado-Díaz and J. Sampedro-Gómez were hired.

## CONFLICTS OF INTEREST

V. Vicente-Palacios works for Philips Healthcare Ibérica. The remaining researchers have no conflicts of interest to declare.

## APPENDIX. SUPPLEMENTARY DATA

Supplementary data associated with this article can be found in the online version, at <https://doi.org/10.1016/j.rec.2019.05.014>

## REFERENCES

- Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep*. 2014;16:441.
- Samuel AL. Some studies in machine learning using the game of checkers. *IBM J Res Dev*. 1959;3:210-229.
- LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436-444.
- Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nat Biomed Eng*. 2018;2:158-164.
- Medrano IH, Guizarro JT, Belda C, et al. Savana: Re-using electronic health records with artificial intelligence. *International Journal of Interactive Multimedia and Artificial Intelligence*. 2018;4:8-12.
- Gonzalez-Mansilla A, Martinez-Legazpi P, Prieto A, et al. Valve area and the risk of overestimating aortic stenosis. *Heart*. 2019;105:911-919.
- Yoo JS, Kim JB, Jung SH, et al. Mitral durability after robotic mitral valve repair: analysis of 200 consecutive mitral regurgitation repairs. *J Thorac Cardiovasc Surg*. 2014;148:2773-2779.
- Miller RA, Pople Jr HE, Myers JD. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med*. 1982;307:468-476.
- Melero-Alegria JJ, Cascon M, Romero A, et al. SALMANTICOR study. Rationale and design of a population-based study to identify structural heart disease abnormalities: a spatial and machine learning analysis. *BMJ Open*. 2019;9:e024605.
- Rajkumar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med*. 2019;380:1347-1358.
- Murphy KP. *Machine learning: a probabilistic perspective*. Cambridge, Massachusetts: MIT press; 2012. Available at: [https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning\\_%20A%20Probabilistic%20Perspective%205BMurphy%202012-08-24%5D.pdf](https://doc.lagout.org/science/Artificial%20Intelligence/Machine%20learning/Machine%20Learning_%20A%20Probabilistic%20Perspective%205BMurphy%202012-08-24%5D.pdf)
- Barreiro-Perez M, Tundidor-Sanz E, Martin-Garcia A, et al. First magnetic resonance managed by a cardiology department in the Spanish public healthcare system. Experience and difficulties of an innovative model. *Rev Esp Cardiol*. 2018;71:365-372.
- Baumgartner CF, Koch LM, Pollefeys M, Konukoglu E. An exploration of 2 D and 3 D deep learning techniques for cardiac MR image segmentation. Available at: <https://arxiv.org/abs/1709.04496>
- Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal*. 2017;42:60-88.
- Char DS, Shah NH, Magnus D. Implementing machine learning in health care – addressing ethical challenges. *N Engl J Med*. 2018;378:981-983.
- Gijbsberts CM, Groenewegen KA, Hoefler IE, et al. Race/ethnic differences in the associations of the Framingham risk factors with carotid IMT and cardiovascular events. *PLoS One*. 2015;10:e0132321.
- Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: Addressing ethical challenges. *PLoS Med*. 2018;15:e1002689.
- Sturm I, Lapuschkin S, Samek W, Muller KR. Interpretable deep neural networks for single-trial EEG classification. *J Neurosci Methods*. 2016;274:141-145.
- Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. *Boletín Oficial del Estado*. 2018. Available at: <https://www.boe.es/eli/es/lo/2018/12/05/3/con>
- Mesko B, Hetenyi G, Gyorffy Z. Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Serv Res*. 2018;18:545.
- Al'Aref SJ, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J*. 2019;40:1975-1986.
- Ebrahimzadeh E, Kalantari M, Joulani M, Shahraki RS, Fayaz F, Ahmadi F. Prediction of paroxysmal atrial fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on HRV signal. *Comput Methods Programs Biomed*. 2018;165:53-67.
- Budzianowski J, Hiczkiewicz J, Burchardt P, et al. Predictors of atrial fibrillation early recurrence following cryoballoon ablation of pulmonary veins using statistical assessment and machine learning algorithms. *Heart Vessels*. 2019;34:352-359.
- Eerikainen LM, Vanschoren J, Rooijackers MJ, Vullings R, Aarts RM. Reduction of false arrhythmia alarms using signal selection and machine learning. *Physiol Meas*. 2016;37:1204-1216.
- Nanayakkara S, Fogarty S, Tremeer M, et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. *PLoS Med*. 2018;15:e1002709.
- Yildirim O, Plawiak P, Tan RS, Acharya UR. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Comput Biol Med*. 2018;102:411-420.
- Weng SF, Rejs J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One*. 2017;12:e0174944.
- Huang Z, Chan TM, Dong W. MACE prediction of acute coronary syndrome via boosted resampling classification using electronic medical records. *J Biomed Inform*. 2017;66:161-170.
- Shouval R, Hadanny A, Shlomo N, et al. Machine learning for prediction of 30-day mortality after ST elevation myocardial infarction: An acute coronary syndrome Israeli survey data mining study. *Int J Cardiol*. 2017;246:7-13.
- Wallert J, Tomasoni M, Madison G, Held C. Predicting two-year survival versus non-survival after first myocardial infarction using machine learning and Swedish national register data. *BMC Med Inform Decis Mak*. 2017;17:99.
- Shameer K, Johnson KW, Yahi A, et al. Predictive modeling of hospital readmission rates using electronic medical record-wide machine learning: a case-study using Mount Sinai Heart Failure cohort. *Pac Symp Biocomput*. 2017;22:276-287.
- Inan OT, Baran Pouyan M, Javaid AQ, et al. Novel wearable seismocardiography and machine learning algorithms can assess clinical status of heart failure patients. *Circ Heart Fail*. 2018;11:e004313.
- Medved D, Ohlsson M, Högglund P, Andersson B, Nugues P, Nilsson J. Improving prediction of heart transplantation outcome using deep learning techniques. *Sci Rep*. 2018;8:3613.
- Sanchez-Martinez S, Duchateau N, Erdei T, et al. Machine learning analysis of left ventricular function to characterize heart failure with preserved ejection fraction. *Circ Cardiovasc Imaging*. 2018;11:e007138.
- Ahmad T, Lund LH, Rao P, et al. Machine learning methods improve prognostication, identify clinically distinct phenotypes, and detect heterogeneity in response to therapy in a large cohort of heart failure patients. *J Am Heart Assoc*. 2018. <http://dx.doi.org/10.1161/JAHA.117.008081>.
- Samad MD, Ulloa A, Wehner GJ, et al. Predicting survival from large echocardiography and electronic health record datasets: optimization with machine learning. *JACC Cardiovasc Imaging*. 2018;12:681-689.
- Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-learning algorithms to automate morphological and functional assessments in 2 D echocardiography. *J Am Coll Cardiol*. 2016;68:2287-2295.
- Knackstedt C, Bekkers SC, Schummers G, et al. Fully automated versus standard tracking of left ventricular ejection fraction and longitudinal strain: The FAST-EFs Multicenter Study. *J Am Coll Cardiol*. 2015;66:1456-1466.
- Tesche C, De Cecco CN, Baumann S, et al. Coronary CT angiography-derived fractional flow reserve: machine learning algorithm versus computational fluid dynamics modeling. *Radiology*. 2018;288:64-72.
- Sengupta PP, Huang YM, Bansal M, et al. Cognitive machine-learning algorithm for cardiac imaging: a pilot study for differentiating constrictive pericarditis from restrictive cardiomyopathy. *Circ Cardiovasc Imaging*. 2016. <http://dx.doi.org/10.1161/CIRCIMAGING.115.004330>.
- Bernard O, Lalonde A, Zotti C, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? *IEEE Trans Med Imaging*. 2018;37:2514-2525.
- Xiong Z, Liu T, Tse G, et al. A machine learning aided systematic review and meta-analysis of the relative risk of atrial fibrillation in patients with diabetes mellitus. *Front Physiol*. 2018;9:835.
- Li Q, Rajagopalan C, Clifford GD. Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Trans Biomed Eng*. 2014;61:1607-1613.
- Figuera C, Irueta U, Morgado E, et al. Machine learning techniques for the detection of shockable rhythms in automated external defibrillators. *PLoS One*. 2016;11:e0159654.
- Lyon A, Ariga R, Mincholé A, et al. Distinct ECG phenotypes identified in hypertrophic cardiomyopathy using machine learning associate with arrhythmic risk markers. *Front Physiol*. 2018;9:213.
- Betancur J, Otaki Y, Motwani M, et al. Prognostic value of combined clinical and myocardial perfusion imaging data using machine learning. *JACC Cardiovasc Imaging*. 2018;11:1000-1009.
- Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of machine learning techniques for heart failure readmissions. *Circ Cardiovasc Qual Outcomes*. 2016;9:629-640.
- Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-day all-cause readmissions in patients hospitalized for heart failure: comparison of machine learning and other statistical approaches. *JAMA Cardiol*. 2017;2:204-209.
- Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial intelligence in cardiology. *J Am Coll Cardiol*. 2018;71:2668-2679.
- Fallahzadeh R, Pedram M, Ghasemzadeh H. SmartSock: a wearable platform for context-aware assessment of ankle edema. *Conf Proc IEEE Eng Med Biol Soc*. 2016;2016:6302-6306.
- Kalscheur MM, Kipp RT, Tattersall MC, et al. Machine learning algorithm predicts cardiac resynchronization therapy outcomes: lessons from the COMPANION Trial. *Circ Arrhythm Electrophysiol*. 2018;11:e005499.
- Salem Omar AM, Shameer K, Narula S, et al. Artificial intelligence-based assessment of left ventricular filling pressures from 2-dimensional cardiac ultrasound images. *JACC Cardiovasc Imaging*. 2018;11:509-510.
- Lancaster MC, Salem Omar AM, Narula S, Kulkarni H, Narula J, Sengupta PP. Phenotypic clustering of left ventricular diastolic function parameters: patterns and prognostic relevance. *JACC Cardiovasc Imaging*. 2019;12(7 Pt 1):1149-1161.
- Carneiro G, Nascimento JC. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. *IEEE Trans Pattern Anal Mach Intell*. 2013;35:2592-2607.
- Slomka PJ, Dey D, Sitek A, Motwani M, Berman DS, Germano G. Cardiac imaging: working towards fully-automated machine analysis & interpretation. *Expert Rev Med Devices*. 2017;14:197-212.
- Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal*. 2016;30:108-119.

57. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal.* 2017;35:159–171.
58. Dou Q, Yu L, Chen H, et al. 3 D deeply supervised network for automated segmentation of volumetric medical images. *Med Image Anal.* 2017;41:40–54.
59. Zheng Q, Delingette H, Duchateau N, Ayache N. 3-D Consistent and robust segmentation of cardiac images by deep learning with spatial propagation. *IEEE Trans Med Imaging.* 2018;37:2137–2148.
60. Zhang J, Gajjala S, Agrawal P, et al. Fully automated echocardiogram interpretation in clinical practice. *Circulation.* 2018;138:1623–1635.
61. Wolterink JM, Leiner T, De Vos BD, Van Hamersvelt RW, Viergever MA, Isgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Med Image Anal.* 2016;34:123–136.
62. Abdolmanafi A, Duong L, Dahdah N, Cheriet F. Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography. *Biomed Opt Express.* 2017;8:1203–1220.
63. Lekadir K, Galimzianova A, Betriu A, et al. A convolutional neural network for automatic characterization of plaque composition in carotid ultrasound. *IEEE J Biomed Health Inform.* 2017;21:48–55.
64. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A deep cascade of convolutional neural networks for dynamic MR image reconstruction. *IEEE Trans Med Imaging.* 2018;37:491–503.
65. Zreik M, Lessmann N, Van Hamersvelt RW, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal.* 2018;44:72–85.
66. Lee HC, Yoon HK, Nam K, et al. Derivation and validation of machine learning approaches to predict acute kidney injury after cardiac surgery. *J Clin Med.* 2018. <http://dx.doi.org/10.3390/jcm7100322>.
67. De la Torre-Hernández. Edelman ER. From nonclinical research to clinical trials and patient-registries: challenges and opportunities in biomedical research. *Rev Esp Cardiol.* 2017;70:1121–1133.

**Anexo 2. Artículo enviado. "Machine Learning to Predict Stent Reestenosis Based on Daily Demographic, Clinical and Angiographic Characteristics"**

Anexo 2. Artículo enviado. PDF 25 caras.

# MACHINE LEARNING TO PREDICT STENT RESTENOSIS BASED ON DAILY DEMOGRAPHIC, CLINICAL AND ANGIOGRAPHIC CHARACTERISTICS

Jesús Sampedro-Gómez<sup>\*1,2</sup>, ENG  
P. Ignacio Dorado-Díaz<sup>\*1,2</sup>, ENG  
Víctor Vicente-Palacios<sup>3</sup>, ENG, PHD  
Antonio Sánchez-Puente<sup>1</sup>, PHYS, PHD  
Manuel Jiménez-Navarro<sup>2,4</sup>, MD, PHD  
J. Alberto San Roman<sup>2,5</sup>, MD, PHD  
Purificación Galindo-Villardón<sup>6</sup>, BSTT, PHD  
Pedro L Sanchez<sup>†1,2</sup>, MD, PHD  
Francisco Fernández-Avilés<sup>2,7</sup>, MD, PHD

From the:

<sup>1</sup> Department of Cardiology, Hospital Universitario de Salamanca-IBSAL, Universidad de Salamanca, Spain

<sup>2</sup> CIBERCV, Instituto de Salud Carlos III, Spain

<sup>3</sup> Philips Healthcare

<sup>4</sup> UGC Área del Corazón, Hospital Virgen de la Victoria-IBIMA, Universidad de Málaga, Spain

<sup>5</sup> Instituto de Ciencias del Corazón, Hospital Clínico Universitario de Valladolid, Spain

<sup>6</sup> Department of Statistics, Universidad de Salamanca, Spain

<sup>7</sup> Department of Cardiology, Hospital General Universitario Gregorio Marañón-IiSGM, Universidad Complutense de Madrid, Spain

\*Jesús Sampedro-Gómez and P. Ignacio Dorado-Díaz equally contributed to this study as first authors

**Brief Title:** *Stent restenosis prediction using machine learning*

**Potential Conflicts of Interest:** None to disclose.

**Data Recruitment:** Gracia-3 trial

**Address for Correspondence:**

†Pedro L Sanchez, MD, PhD

Cardiology Department. Hospital Universitario de Salamanca-IBSAL

Paseo de San Vicente 58-187

37007 Salamanca, Spain

Telephone: 34-923291100 (ext 55356)

E-mail: [pedrolsanchez@secardiologia.es](mailto:pedrolsanchez@secardiologia.es)

## **ABSTRACT**

**OBJECTIVES:** This study sought to use machine-learning (ML) to discriminate stent restenosis (SR) compared to predictive clinical scores of SR. In order to develop an easily-applicable model, we performed our predictions without any additional variables than those obtained in daily practice.

**BACKGROUND:** There is growing interest in precision medicine techniques as ML, which can deliver individually adapted medical care.

**METHODS:** The dataset, obtained from the GRACIA-3 trial, consisted of 263 patients with demographic, clinical and angiographic characteristics; 23 of them presented SR at 12-months after stent implantation. A methodology to work with small unbalanced datasets was used and state-of-the-art ML classifiers were trained.

**RESULTS:** Our best performing model was developed with an extremely randomized trees (ERT) classifier; which significantly outperformed (0.77; area under the receiver-operating-characteristic curve [AUC-ROC]) three clinical scores; PRESTO-1 (0.58 AUC-ROC), PRESTO-2 (0.58 AUC-ROC), and TLR (0.62 AUC-ROC). Precision-recall curves for SR and non-SR classes gave a more accurate picture of the ERT model's performance showing an efficient algorithm (0.97 AUC) for non-SR, returning high precision as well as high recall. For a threshold considered as optimal, out of 1,000 patients undergoing stent implantation, our ERT model would correctly predict 181 (18%) more cases compared to the best clinical score. The most important variables ranked according to their contribution to the predictions were diabetes,  $\geq 2$  vessel-coronary disease, post-PCI TIMI-flow, abnormal platelets, post-PCI thrombus, and abnormal cholesterol.

**CONCLUSIONS:** Applied immediately after stent implantation, a ML model better differentiates those patients who will present or not SR over current discriminators.

## **KEY WORDS**

Machine-learning • percutaneous coronary intervention • stent restenosis • myocardial infarction • Receiver-operator-characteristic • Extremely Randomized Trees

**ABSTRACT WORD COUNT: 250**



## **ABBREVIATIONS**

AUC = area under the curve

ERT = Extremely Randomized Trees

MI = myocardial infarction

ML = machine-learning

PCI = percutaneous coronary interventions

PR = Precision-Recall

ROC = receiver-operating-characteristic

SR = stent restenosis

STEMI = ST elevation acute myocardial infarction

## INTRODUCTION

Angioplasty and stenting percutaneous coronary interventions (PCI) have become routine practice for the revascularization of coronary vessels with significant obstructive atherosclerotic disease. Furthermore, PCI with stent implantation immediately performed is the gold-standard of treatment in patients with acute myocardial infarction (MI), reducing the rates of death and recurrent ischemia as compared to medical treatment<sup>1</sup>.

The long-term success of these procedures may be limited by stent restenosis (SR)<sup>2, 3</sup>, a pathological process leading to recurrent arterial narrowing at the site of PCI which may manifest as a new MI or force new-target-vessel revascularization. Identifying patients with SR is a major challenge as management of patients with recurrent SR represents a major therapeutic dilemma<sup>4</sup>.

There is growing interest in precision medicine techniques that can deliver individually adapted medical care<sup>5</sup>. One of the new emerging tools that can help to implement precision medicine is machine learning (ML): starting with patient-level observations, algorithms sift through vast number of variables, and look for combinations that reliably predict outcomes<sup>6-8</sup>. We hypothesized that ML algorithms may produce a model that discriminates SR for individual patients better than the widely used predictive clinical discriminators of SR<sup>9, 10</sup>. We combine the use of daily available demographic, clinical, and angiographic data with ML to test this hypothesis.

## **METHODS**

**Figure 1 (central figure)** guides through the phases we have followed to build our ML model: preparation of the model, training the model, and evaluating the model. In preparation of the model, we defined the task, selected the subjects of the study, discriminated input (features) and output (labels) and proceeded to the data collection and preparation. The set of processed examples came from the GRACIA-3 randomized clinical trial<sup>11</sup>, and was divided into two sets. The first, the training dataset, was used to build the model; the second, the test set, was used to assess how well the model performs. In the model training, the model was developed including a selection of features, the training of different ML classifiers and the application of cross-validation methodology. The parameters of the classification algorithm and the particular feature selection strategy were chosen by means of a hyperparameter tuning phase to further improve the training algorithm. Finally, in evaluating the model, the test set was run through the final ML model to estimate the performance of the model in a real-world scenario. Comparisons of the development ML model with classical predictive clinical risk scores of SR were also performed.

### **PREPARING TO BUILD THE MACHINE LEARNING MODEL**

#### **Task Definition**

The aim of our study was to predict 12-month follow-up SR in patients with ST-elevation acute MI undergoing PCI. In order to develop an easily-applicable ML model, we performed our prediction without any additional variables than those obtained in routine clinical practice.

We started our supervised ML model specifying our input variables (features) that should be mapped to a corresponding output (label)<sup>6</sup>. Input data consists in patient's demographic and clinical characteristics, quantitative coronary analysis (QCA) for assessment of coronary

artery dimensions before and after stent implantation, physical examination parameters, ejection fraction, and routine biochemical parameters available at the time of the PCI procedure. As for the corresponding outcome, we defined SR of the infarct-related lesion as a >50% narrowing of the lumen diameter in the target segment, defined as all portions of the vessel that received treatment within the stent zone including the proximal and distal 5-mm margins<sup>12</sup>, at 12-month follow-up.

### **Data collection**

The data used to train and validate the ML models comes from the GRACIA-3 trial<sup>11</sup>. Briefly, the GRACIA-3 trial was a 2x2 randomized, open-label, multicenter, clinical trial that compares the efficacy of the paclitaxel-eluting stent with the conventional bare-metal stent. Patients with STEMI were enrolled from 20 Spanish hospitals. To determine the incidence of SR, coronary angiography was performed at baseline and after 12 months of follow-up. All QCA for assessment of coronary artery dimensions before and after stent implantation were analyzed at an independent angiography core laboratory (ICICOR, Valladolid, Spain) with a well-validated quantitative computer-based system (Medis, Leesburg, Va). The rate of SR was assessed by an experienced reader who was blinded to and not directly involved in the stent-implantation project.

The GRACIA-3 consisted of 436 patients of whom 299 (69%) underwent 12-month angiographic follow-up. The dataset used for model development is composed by 263 out of 299 (88%) patients where 68 features related with demographic, clinical, stent characteristics, and angiographic data were available. This model development cohort was identified before merging input data with outcome to avoid selection bias.

The GRACIA-3 executive committee and the institutional Ethical Committee of the University Hospital of Salamanca approved the retrospective use of the identified data from the trial for this study (PI201902178).

### **Data preparation**

An important step in preparing any ML model consists in preprocessing the raw data as a set of features to be usable by ML algorithms. For this purpose, multicategory features were one-hot encoded in binary variables. Information related to the particular coronary segment where grouped in 3 features with the vessel where the PCI was performed (left anterior descending artery, left circumflex or right coronary artery). Missing values were filled with the median and mode of each continuous and categorical feature, respectively.

ML models conventionally rely on distinct sub-datasets: one for training and another test dataset for evaluating the model. Training datasets are the learning basis of the models and the test set is then run through the final model and used to evaluate competing models. Different approaches are required in data preparation when the dataset used to train the model is not large and has unbalanced outputs. The main goal is to maximize the amount of data used to train the model and estimate its future performance. This was our case with 23 out of 263 patients developing 12-month follow-up SR. To address this issue we performed k-fold cross validation methodology<sup>13</sup>, to randomly split the dataset into k equally sized parts: k-1 parts constituted our training dataset and the remaining one was used as test dataset for evaluating the model. In our model we used 10-fold cross-validation, a typical setup which guarantees a minimum number of minority class cases represented in both training and test (**Supplementary Data: Figure 1**).

## **TRAINING THE MACHINE-LEARNING MODEL**

### **Machine learning models and feature selection techniques**

We applied and compared the performance of six ML classifiers that are widely used in the literature<sup>8</sup>: random forest (RF), extremely randomized trees (ERT), gradient boosting (GB), support vector machine classifier (SVC), L2-regularized logistic regression (LR), and non-regularized logistic regression (LR\_NOREG). These six ML classifiers were trained with 2 different feature selection techniques, making a total of 12 different comparable models. The first feature selection method was based in univariate selection where ANOVA tests were used to select those features that have the strongest relationship with the output variable, presence of SR at 12-months after stent implantation. The second feature selection method was based in feature importance where using a RF classifier we obtained a score for each feature of the dataset; the higher the score more important is the feature towards the output variable.

Model development codes were executed in Python and for the implementation of k-fold splitting, feature selection, and ML classifiers, the open source library scikit-Learn was used<sup>14</sup>. The developed code used to train the models can be consulted in <https://github.com/IA-Cardiologia-husa/Restenosis>. Also, a Google Colaboratory [notebook](#) with the developed model and instructions to make predictions is released as open source.

### **Cross-validation scheme and hyperparameter tuning**

All models were trained and evaluated with the aforementioned cross-validation scheme formed by stratified 10-folds and 20 repetitions. Tuning the hyperparameters to improve the performance of ML classifiers was also performed. Model hyperparameters are the properties that govern the entire training process; e.g. the number of features filtered by the feature selection technique or the depth of decision trees used by the ensemble classifiers. Automated hyperparameter setting included the best feature selection strategy according to best performance in also a stratified cross-validation which guarantee a minimum number of minority class cases and measuring the performance as the area under the receiver operating

characteristic (AUC-ROC) curve, which is the same metric we further use to evaluate our final models. This approach is computationally costly due to its highly iterative nature, but it is affordable when dealing with small datasets and allows the use of all the dataset without holding-out a part of it to test the generalization error<sup>15</sup>. The optimization of the parameters was performed with a randomized searching algorithm<sup>16</sup>. The fixed values of not optimized hyperparameters and the ranges of optimized ones for each classification and feature selection algorithm can be consulted in **Supplementary Data: Table 1**.

## **EVALUATING THE MACHINE-LEARNING MODEL**

### **Evaluation scheme**

Once training was complete, all ML classifiers were evaluated with the aforementioned 10-folds cross-validation scheme with 20 repetitions. We used the test dataset we set aside to evaluate our models. Each time we evaluate a new instance of our model using this data, we followed all the steps previously described for training the model, including a new feature selection and hyperparameter tuning.

### **Comparison of ML models with predictive clinical risk scores of SR**

We compared the performance of the developed ML algorithms to standard predictive multivariate logistic regression models: PRESTO-1, PRESTO-2 and TLR risk scores.

Briefly, PRESTO scores were developed from the PRESTO (Prevention of Restenosis with Tranilast and its Outcomes) trial of 1,312 patients<sup>17</sup>. Investigators constructed 2 risk scores: PRESTO-1 used preprocedural variables (female gender, vessel size, lesion length, diabetes, smoking status, type C lesion, any previous PCI, and unstable angina) that are simple to assess and have frequently been reported to be strong predictors of SR; PRESTO-2 considered significant univariate clinical and angiographic predictors of SR identified from the PRESTO dataset (treated diabetes mellitus, nonsmoker, vessel size, lesion length, type C

lesion, ostial location, and previous PCI). Both scores achieved an AUC-ROC of 0.63 on the PRESTO population.

TLR risk score (age<60, prior PCI, unprotected left main PCI, saphenous vein graft PCI, minimum stent diameter, and total stent length) used data from the EVENT (Evaluation of Drug-Eluting Stents and Ischemic Events) registry of 8,829 patients<sup>18</sup>. TLR achieved a 0.68 AUC-ROC on its independent validation cohort.

For the assessment of external validation of PRESTO and TLR scores, we used the odds ratios calculated in their respective original cohort studies and the integer scores reported by their authors, following TRIPOD statement<sup>19</sup>. We computed the risk scores using the same cross-validation scheme and splits described in the model development, obtaining 200 randomized test subsamples (10-fold with 20 repetitions).

### **Evaluation Metrics**

The differences in patient's demographic, clinical characteristics, quantitative coronary analysis, physical examination, ejection fraction, and routine biochemical parameters for each partition were compared using  $\chi^2$  tests for categorical variables and ANOVA for continuous variables.

In a binary decision problem as our study, our classifiers labeled examples as either positive (SR) or negative (non-SR). The decisions made by the classifiers were represented in contingency tables with four categories (**Supplementary Data; Table 2**): true positives were examples correctly labeled as positives, false positives refer to negative examples incorrectly labeled as positive; true negatives correspond to negatives correctly labeled as negative and; finally, false negatives refer to positive examples incorrectly labeled as negative; where precision is defined as the number of true positives over all the predictive positives, recall or sensitivity as the number of true positives over the real number of positives; specificity as the



number of true negatives over the real number of negatives; and negative predictive value as the number of true negatives over the number of predicted negatives.

ROC curve analysis was used to assess the predictive capacity of each ML model and predictive clinical risk scores in two different ways. First, we concatenated the predictions of all the test sets and constructed a ROC curve. Second, we computed and saved the AUC-ROC of each test set of the outer cross-validation scheme. We used the first strategy to plot the ROC curves for illustration purposes and the second one to compute the mean scores, confidence intervals and t-test significance between ML models and predictive clinical risk scores<sup>20</sup>. We calculated these confidence intervals and t-test significance taking into account that the results over each k-fold are not completely independent<sup>21</sup>.

We further computed the Precision-Recall (PR) curves. Precision-Recall curves are also commonly used for binary decision problems in ML as they can give a more intuitive insight of the behavior of the model when the data used is imbalanced<sup>22, 23</sup>. We show the PR curves referred to the SR and non-SR classes in different plots for better interpretation of the models results. In the non-SR class, precision and recall are equivalent to the negative predictive value and specificity.

The classification performance at particular cutoff thresholds based on the ROC and PR curves were also evaluated according to its sensitivity, specificity, precision, and negative predictive value.

Finally, we compute feature importance of the best model by measuring how AUC decreases when a feature is not available through the method known as Permutation Importance or Mean Decrease Accuracy (MDA)<sup>24</sup>. The method consists in replacing each feature on the test dataset with random noise-feature column and measuring the performance

for the ML model. The weight of the features with positive impact in the predictive model is scaled to 1. The algorithm is also available in the Python open source library, eli5<sup>25</sup>.

## RESULTS

### Characteristics of the study population

The continuous and categorical data input of the patients used for ML model development are shown in the **Supplementary Data: Table 3 and 4**, respectively. Twenty-three patients (8.7%) presented angiographic binary SR at 12 months follow-up of randomization. In the bivariate analysis, patients who experienced SR were significantly more likely to have the following characteristics: diabetes, abnormal cholesterol,  $\geq 2$  vessel-coronary disease, total number of stents implanted in the PCI, lower post-PCI minimal luminal diameter, higher post-PCI percent diameter stenosis, higher post-PCI thrombus and lower post-PCI TIMI flow.

### Comparison of prediction models

**Figure 2** shows the ROC curves for the 6 ML models in combination with the univariate ANOVA feature selection. The ANOVA selection performed better than RF-based feature importance selection. The best ML final model with a mean AUC-ROC of 0.77 (95%: CI, 0.66-0.89) was obtained with the ERT classifier. No significant differences between the ERT model with other ML models were observed, except for SVC (P=0,003).

The prediction accuracy of the ERT model achieved statistically significant improvement compared with those of predictive clinical risk scores (P<0,001; **Figure 3**); PRESTO-1 (AUC-ROC 0.57 [0.43-0.71]), PRESTO-2 (AUC-ROC 0.58 [0.43-0.72]), and TLR (AUC-

ROC 0.62 [0.50-0.74]). The risk scores based on the exact and rounded odd-ratios (OR) showed very similar results.

PR curves gave a more accurate picture of the models performance. These curves, taken from the same learned models, highlight the visual difference between binary decision (SR vs. non-SR) spaces. PR curves for SR prediction showed precision and recall below 0.5; whether, the high AUC-PR for non-SR prediction represents both high precision and high recall, where high precision relates to a low false positive rate and high recall relates to a low false negative rate (**Figure 4**). This difference exists because in this domain the number of negative examples greatly exceeds the number of positive instances; and provides our ETR algorithm an excellent negative predictive value. AUC-PR for the ETR model was 0.96 (0.96-0.99) with a significant ( $P=0.003$ ) advantage over PRESTO-1, PRESTO-2 and TLR scores.

### **Curves interpretation and classification analysis**

In order to better assess the clinical significance of our results, we compared the classification performance of our ETR model with the best performed clinical risk score (TLR) at three particular cut-off thresholds (**Table 1**), based on the ROC and PR curves (**Figures 2 & 3**). This is a critical requirement for a predictive score, where depending on the case we should minimize the number of false negatives or false positives. For a threshold considered optimal for high precision (first operating point), out of 1,000 MI patients undergoing stent implantation, our ETR model would be able to correctly predict 181 more cases compared to the TLR score and is able to achieve a 89.6% of true negatives (71.5% for the TLR score) and detect a group of 32 patients with 50% risk of SR who should be routinely scheduled for 12-months angiographic follow-up. In contrast, for a maximum negative predictive value threshold (third operating point), our model is able to detect a group

of 318 patients with 5% risk of SR avoiding unnecessary follow-up workload and lowering costs.

### **Variable importance**

**Figure 5** shows the importance of the first 6 features that were found as main predictors for the ERT model: diabetes,  $\geq 2$  coronary vessel disease, post-PCI TIMI flow, abnormal platelets, post-PCI thrombus and abnormal total cholesterol. Diabetes, only included in the PRESTO-2 score, appeared to be more important for the SR prediction than other cardiovascular risk factors or habits as smoking or alcohol consumption. Among the angiographic variables, post-PCI TIMI flow was the most important feature of the top predictors. Other features not shown in the figure presented values lower than 0.01 (scaled to the main feature importance) or negative values (variables that have negative contribution to the classification task).

This variable importance selection is different from the feature selection used for the computation of the generalized metric score of the model. During that process, the number of features, their particular selection and the rest of the model hyperparameters are chosen exclusively with the data of the training set in the nested cross-validation scheme. If we calculated the generalization score with a classifier and features optimized with the entire dataset, we would be leaking information from the test set to the model, incurring in overfitting and thus, the generalization score would be over optimistic. This kind of miscalculation is not uncommon in the application of cross-validation and as example, in this study, the AUC-ROC curve for the ERT model would have reached a value of 0.82.

## DISCUSSION

Using ML, we developed an ETR model to predict SR based on variables obtained in routine clinical practice in a retrospective analysis of the GRACIA 3 trial. Compared to standard predictive multivariate logistic regression models as the PRESTO-1, PRESTO-2 and LTR risk scores; our ETR model increased prediction by 20%.

Ideally, treatment recommendations for patients will include individualized outcome estimates. In this regard, our ML model is of clinical and economic relevance; as we could avoid unnecessary follow-up and decrease costs for patients with predicted non-risk of SR and; by contrast, improve outcomes in patients with high risk of SR stabilising 12-month ischemic tests or catheterization to avoid adverse clinical outcomes. Although SR may be associated with a recurrence of stable angina symptoms, it is recognized that up to 1 out of 3 patients present with MI or unstable angina amenable to repeat catheter intervention<sup>26, 27</sup>.

The GRACIA 3 trial is unique in developing this predictive model for several reasons. First of all, it provides the opportunity to incorporate feature (or variable) input regarding the use of bare metal stents versus drug eluting stents in a clinical context, ST elevation acute MI, where the benefit of new drug eluting stents is still discussed and under comparison<sup>28</sup>. Thus, our predictive model would be specifically useful in this specific scenario. Second, the number of paired angiographic studies performed in the GRACIA 3 trial (60,3%) is considerable and similar to surveillance studies in large cohort of patients where follow-up catheterization was also scheduled<sup>29</sup>. Finally, the data come from a randomized clinical trial, involving 20 Spanish hospitals, where all angiograms were analyzed at an independent angiographic core laboratory. Thus, the methodology used to define the presence of SR was of maximum rigor and quality and represents an excellent opportunity to develop a ML SR predictive model.

ML algorithms are a growing trend in the analysis of biomedical data through their ability to find complex patterns that typically appear in this field<sup>6</sup>. This approach has been used to study several cardiac pathologies<sup>7, 8</sup>. In the case of SR, to our knowledge, the only study that has used a ML approach was developed by Cui et al<sup>30</sup>. In their study, a set of 6 plasma metabolites were identified to predict SR with an excellent AUC-ROC curve of 0.94. However, these biomarkers are not usually measurable in daily cardiac practice. Thus, developing a model to predict SR based on variables obtained in routine clinical practice can be a valuable tool for improving the risk-stratification of conventional patients, providing individually tailored follow-up and treatment.

ML models work efficiently when they are provided with large amount of data and balanced classes. These conditions are not usually fulfilled when dealing with data consisting in patients that will or not develop pathology, as it is the case of the SR after a regular PCI procedure. In this sense, ROC and PR curves are commonly used to present results for binary decision problems in ML. When the class distribution is close to being uniform, ROC curves have many desirable properties. However, when dealing with a highly skewed dataset as this study, PR curves should also be used. An important difference between ROC and PR is the visual representation of the curves. The goal in ROC space is to be in the upper-left-hand corner and when one looks at the ETR model ROC curve in our study it appears to be fairly close to optimal<sup>23</sup>. In PR space the goal is to be in the upper-right-hand corner, and the PR curve shows that the ETR model is very optimal when predicting non-SR, but there is still significant room for improvement when predicting SR. Thus, in our example, PR curves give a more accurate picture of our algorithm's performance.

Scenarios like the one shown in this study, small datasets with imbalanced target classes, are common in the clinical field, where some pathologies are rare, and the data collection is costly. For this reason, we used a consistent methodology based on cross-validation to

evaluate the generalization of our results<sup>21</sup>. However, the use of cross-validation and steps like the fine tuning of the algorithms, the data preprocessing and specially, the feature selection, must be handled with care, since they can lead to a leak of information from the test to the training sets and thus, an over optimistic outcome while trying to maximize a metric. We also performed an interpretation of the broadly used ROC curve in conjunction with the not so common PR curves for the positive and negative SR classes. This kind of assess is necessary in order to compare models and postulate their possible impact in the clinical practice, evaluating their strong and weak points.

## **Limitations**

As we have previously stated, processing of small and unbalanced datasets is a recurrent problem in ML methodology. Different guidelines suggest using much bigger cohorts to both develop new models and validate existing ones. However, this is not always possible when prospective study subjects imply invasive procedures, like routine post-PCI coronary angiography, which are not currently recommended in clinical practice. Although we used different methodology approaches to address these drawbacks, our results show some unavoidable shortcomings.

When comparing our ML algorithm with the existing clinical scores (PRESTO-1, PRESTO-2 and TLR), we were comparing the external performance of these clinical models to the internal performance of our ML algorithm. Thus, to avoid this unfair advantage for the internal validation of our algorithm, we used for the clinical risk scores the odds ratios calculated in their respective original cohort studies and the integer scores reported by their authors, following TRIPOD statement.

Another inherent problem shared by classic statistics and newer ML modelling is the possible presence of bias in the dataset used to fit the algorithm. The GRACIA 3 trial is a

good source of data, due to the use of registers from 20 different Spanish hospitals. However, the data was collected 10 years ago, and it is restricted to the population of Spain. Current PCI data and wider cohorts may lead to additional conclusions and better generalization of the predictions. For this purpose, we provide our developed ML algorithms as free software for other investigators.

Finally, a well-known limitation of ML algorithms is the lack of a straightforward interpretation of the model, besides the scoring of the contribution of the features used to train the model. Depending on the ML algorithm, the construction of a manually assessed risk score may not be possible and thus, the implementation of these tools in the daily clinical practice may require the use of computers and software applications.



## **SOURCES OF FUNDING**

This study was supported by the Spanish Cardiovascular Network (CIBERCV) from the Institute of Health Carlos III (Spanish Ministry of Economy and Competitiveness), from the Fondo de Investigacion Sanitaria (FIS: PI040308, PI040235, PI040361, PI042035, PI042686, PI040166, PI040276, PI 040306, PI040309, PI040350, PI040478, and PI0402037) from the Institute of Health Carlos III (Spanish Ministry of Economy and Competitiveness).

## References

1. Szummer K, Jernberg T and Wallentin L. From Early Pharmacology to Recent Pharmacology Interventions in Acute Coronary Syndromes: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2019;74:1618-1636.
2. Califf RM, Fortin DF, Frid DJ, Harlan WR, 3rd, Ohman EM, Bengtson JR, Nelson CL, Tcheng JE, Mark DB and Stack RS. Restenosis after coronary angioplasty: an overview. *J Am Coll Cardiol*. 1991;17:2B-13B.
3. Byrne RA, Joner M and Kastrati A. Stent thrombosis and restenosis: what have we learned and where are we going? The Andreas Gruntzig Lecture ESC 2014. *Eur Heart J*. 2015;36:3320-31.
4. Alfonso F, Byrne RA, Rivero F and Kastrati A. Current treatment of in-stent restenosis. *J Am Coll Cardiol*. 2014;63:2659-73.
5. Ginsburg GS and Phillips KA. Precision Medicine: From Science To Value. *Health Aff (Millwood)*. 2018;37:694-701.
6. Rajkomar A, Dean J and Kohane I. Machine Learning in Medicine. *N Engl J Med*. 2019;380:1347-1358.
7. Westcott RJ and Tcheng JE. Artificial Intelligence and Machine Learning in Cardiology. *JACC Cardiovasc Interv*. 2019;12:1312-1314.
8. Dorado-Diaz PI, Sampedro-Gomez J, Vicente-Palacios V and Sanchez PL. Applications of artificial intelligence in cardiology: the future is already here. *Rev Esp Cardiol*. 2019.
9. Singh M, Gersh BJ, McClelland RL, Ho KKL, Willerson JT, Penny WF and Holmes DR. Clinical and Angiographic Predictors of Restenosis After Percutaneous Coronary Intervention. *Circulation*. 2004;109:2727-2731.
10. Stolker JM, Kennedy KF, Lindsey JB, Marso SP, Pencina MJ, Cutlip DE, Mauri L, Kleiman NS and Cohen DJ. Predicting Restenosis of Drug-Eluting Stents Placed in Real-World Clinical Practice. *Circulation: Cardiovascular Interventions*. 2010;3:327-334.
11. Sanchez PL, Gimeno F, Ancillo P, Sanz JJ, Alonso-Briales JH, Bosa F, Santos I, Sanchis J, Bethencourt A, Lopez-Messa J, de Prado AP, Alonso JJ, San Roman JA and Fernandez-Aviles F. Role of the paclitaxel-eluting stent and tirofiban in patients with ST-elevation myocardial infarction undergoing postfibrinolysis angioplasty: the GRACIA-3 randomized clinical trial. *Circ Cardiovasc Interv*. 2010;3:297-307.
12. Kuntz RE and Baim DS. Defining coronary restenosis. Newer clinical and angiographic paradigms. *Circulation*. 1993;88:1310-23.
13. Rodriguez JD, Perez A and Lozano JA. Sensitivity analysis of kappa-fold cross validation in prediction error estimation. *IEEE Trans Pattern Anal Mach Intell*. 2010;32:569-75.

14. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M and Duchesnay E. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12:2825-2830.
15. Varma S and Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*. 2006;7:91.
16. Bergstra J and Bengio Y. Random search for hyper-parameter optimization. *J Mach Learn Res*. 2012;13:281-305.
17. Singh M, Gersh BJ, McClelland RL, Ho KK, Willerson JT, Penny WF and Holmes DR, Jr. Clinical and angiographic predictors of restenosis after percutaneous coronary intervention: insights from the Prevention of Restenosis With Tranilast and Its Outcomes (PRESTO) trial. *Circulation*. 2004;109:2727-31.
18. Stolker JM, Kennedy KF, Lindsey JB, Marso SP, Pencina MJ, Cutlip DE, Mauri L, Kleiman NS, Cohen DJ and Investigators E. Predicting restenosis of drug-eluting stents placed in real-world clinical practice: derivation and validation of a risk model from the EVENT registry. *Circ Cardiovasc Interv*. 2010;3:327-34.
19. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, Vickers AJ, Ransohoff DF and Collins GS. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med*. 2015;162:W1-73.
20. Airola A, Pahikkala T, Waegeman W, De Baets B and Salakoski T. An experimental comparison of cross-validation techniques for estimating the area under the ROC curve. *Computational Statistics & Data Analysis*. 2011;55:1828-1844.
21. Bouckaert RR. Choosing between two learning algorithms based on calibrated tests. *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*. 2003:51-58.
22. Saito T and Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One*. 2015;10:e0118432.
23. Davis J and Goadrich M. The relationship between precision-recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning (ICML-06)*. 2006:233-240.
24. Breiman L. Random forests. *Machine Learning*. 2001;45:5-32.
25. <https://eli5.readthedocs.io/en/latest/index.html>.
26. Chen MS, John JM, Chew DP, Lee DS, Ellis SG and Bhatt DL. Bare metal stent restenosis is not a benign clinical entity. *Am Heart J*. 2006;151:1260-4.
27. Dangas GD, Claessen BE, Caixeta A, Sanidas EA, Mintz GS and Mehran R. In-stent restenosis in the drug-eluting stent era. *J Am Coll Cardiol*. 2010;56:1897-907.

28. Feinberg J, Nielsen EE, Greenhalgh J, Hounscome J, Sethi NJ, Safi S, Gluud C and Jakobsen JC. Drug-eluting stents versus bare-metal stents for acute coronary syndrome. *Cochrane Database Syst Rev*. 2017;8:CD012481.

29. Mindrescu C, Brener SJ, Guerchicoff A, Fahy M, Parise H, Mehran R and Stone GW. Impact of scheduled angiographic follow-up in patients treated with primary percutaneous coronary intervention for ST-segment elevation myocardial infarction. *J Interv Cardiol*. 2013;26:319-24.

30. Cui S, Li K, Ang L, Liu J, Cui L, Song X, Lv S and Mahmud E. Plasma Phospholipids and Sphingolipids Identify Stent Restenosis After Percutaneous Coronary Intervention. *JACC Cardiovasc Interv*. 2017;10:1307-1316.

## FIGURE LEGENDS

**Figure 1 (central figure).** Overview of the phases we have followed to build our machine-learning model.

**Figure 2. Receiver operating characteristic curves for machine learning models.** ERT: Extremely Randomized Trees; RF: Random Forest; LR: Logistic Regression with L2 regularization; LR\_NOREG: Logistic Regression without regularization penalty; GB: Gradient Boosting Classifier; SVC: Support Vector Machine Classifier. The model with best mean AUC was the ERT, only with significant difference with the SVC classifier.

**Figure 3. Receiver operating characteristic curves comparing the best machine learning model with those of the classical risk scores.** Risk scores with (*OR*) suffix were calculated with the exact odd-ratios of each risk score model. Using area under curve (AUC), the extremely randomized trees (ERT) model was statistically significant superior to those derived from standard multivariate logistic regression models: PRESTO-1, PRESTO-2 and TLR risk scores.

**Figure 4. Precision-recall curves for the prediction of stent restenosis (precision) or no stent restenosis (negative predictive value) of machine learning models and risk scores comparators.** In Precision-Recall (PR) curves the goal is to be in the upper-right-hand corner and the PR curve for the Extremely Randomized Trees classifier (ERT) appears to be close to optimal.

**Figure 5. Bar chart of feature importance.** The importance of each of the features in predicting 12-months follow-up stent restenosis with an Extremely Randomized Trees

classifier. The most important features were diabetes,  $\geq 2$  coronary vessel disease, post-PCI TIMI flow.

## TABLES

**Table 1.** Values of true positives (TP), true negatives (TN), false positives (FP), false negatives (FN), sensitivity, specificity, precision and negative predictive value (NPV) with 3 different thresholds for the ERT model and the TLR score. Values are scaled for a population of 1,000 patients with the same prevalence of stent restenosis as the GRACIA-3 study.

Model - point	Sensitivity	Specificity	Precision	NPV	TP	TN	FP	FN
ERT - 1	0,18	0,98	0,49	0,93	16	896	16	72
TLR - 1	0,18	0,78	0,07	0,91	16	715	197	72
ERT - 2	0,56	0,86	0,28	0,95	49	784	128	38
TLR - 2	0,56	0,58	0,11	0,93	49	533	379	39
ERT - 3	0,81	0,53	0,14	0,97	70	487	426	17
TLR - 3	0,81	0,33	0,10	0,95	70	301	611	17

## Anexo 3. Código fuente desarrollado (Python) para la construcción del modelo *machine learning* de predicción de reestenosis

[https://github.com/PedroDorado/Restenosis\\_IA\\_Gracia3](https://github.com/PedroDorado/Restenosis_IA_Gracia3)

- **lib\_ml.py**: librería para importar funciones que se necesitan y funciones desarrolladas para hacer la validación cruzada anidada.
- **model\_zoo.py**: lista de modelos de scikit-learn utilizados
- **script\_nested.py**: para hacer la validación cruzada anidada en el conjunto de datos que le pasemos y calcular el error de generalización.
- **train\_final\_model.ipynb**: para entrenar el modelo final con todo el conjunto de datos.
- **resultados.ipynb**: para sacar las gráficas y resultados

Anexo 3. Código fuente Python modelo reestenosis. PDF 32 caras.

Branch: master ▾

Find file

Copy path

## Restenosis\_IA\_Gracia3 / lib\_ml.py

 PedroDorado Add files via upload

bb23e2b 2 days ago

1 contributor

Raw Blame History



121 lines (99 sloc) 5.67 KB

```
1 import pandas as pd
2 import numpy as np
3 import matplotlib.pyplot as plt
4 import scipy.stats
5 import time
6
7 from sklearn.ensemble import RandomForestClassifier, ExtraTreesClassifier, GradientBoostingClassifier
8 from sklearn.linear_model import LogisticRegression
9 from sklearn.svm import SVC
10
11 from sklearn.feature_selection import SelectFromModel, SelectKBest, f_classif
12 from sklearn.model_selection import train_test_split, RandomizedSearchCV
13 from sklearn.metrics import confusion_matrix, roc_auc_score, precision_recall_fscore_support, accuracy_score, roc_curve, auc, precision_reca
14 from sklearn.model_selection import RepeatedKFold, RepeatedStratifiedKFold, StratifiedKFold, KFold
15 from sklearn.model_selection import cross_val_predict, cross_val_score, cross_validate
16 from sklearn import preprocessing
17 from sklearn.pipeline import Pipeline
18
19 from eli5.sklearn import PermutationImportance
20 from eli5 import explain_weights, explain_weights_df, show_weights
21
22 import multiprocessing as mp
23
24 import warnings
25 warnings.filterwarnings("ignore")
26
27 ### train_split_nested
28 ## Performs a single train-test on a binary classification problem with an hyperparameter-optimized
29 ## model through an inner k-fold iteration over the train SelectKBest
30 # Parameters:
31 # - X, y: data points and labels
32 # - train_index, test_index: self explanatory
33 # - n_split: used to save results when train_split_nested is embedded in an outer k-fold cross-validation loop
34 # - model: scikit-learn classification model to optimize and use
35 # - param_grid: scikit-learn styled hyperparameter dictionary
36 # - inner_rkf: scikit-learn cross_validation generator
37 # Returns:
38 # - df_res_test: dataframe with predictions for each data point.
39 #           Includes: index;
40 #           probability of being in class 1;
41 #           train (0) or test (1); n_split (for outer cross-validation purpose);
42 #           seed given to the inner kf generator
43 # - df_res_inner: dataframe with results over the inner cv loop, used for training evaluation purposes.
44 #           Includes: ROC-AUC over the training set of the best model trained in the inner cv loop;
45 #           ROC-AUC over the test set of the best model trained in the inner cv loop;
46 #           String with construction of the best inner cv model_selection
47 #           Features used for the best model (for feature selection evaluation purposes)
48 def train_split_nested(X, y, train_index, test_index, n_split, model, param_grid, inner_rkf):
49
50     X_train = X.loc[train_index,:]
51     y_train = y[train_index]
52     X_test = X.loc[test_index]
53     y_test = y[test_index]
54
55     gridcv = RandomizedSearchCV(estimator = model,
56                               param_distributions= param_grid,
57                               scoring = 'roc_auc', cv = inner_rkf, refit = True, return_train_score = True,
58                               n_jobs = 1, random_state = 0, n_iter = 20)
```



```

59 gridcv.fit(X_train, y_train)
60
61 inner_train, inner_test = gridcv.cv_results_['mean_train_score'][gridcv.best_index_], gridcv.cv_results_['mean_test_score'][gridcv.best
62
63 y_prob_test = gridcv.best_estimator_.predict_proba(X_test)
64 y_prob_train = gridcv.best_estimator_.predict_proba(X_train)
65
66 outer_test = roc_auc_score(y_test, y_prob_test[:,1])
67 outer_train = roc_auc_score(y_train, y_prob_train[:,1])
68
69 print('TRAIN-TEST SPLIT ', n_split)
70 print('INNER TRAIN: %.3f, INNER TEST: %.3f' % (inner_train, inner_test))
71 print('OUTER TRAIN: %.3f, OUTER TEST: %.3f' % (outer_train, outer_test))
72
73 df_res_test = pd.DataFrame()
74 df_res_test['idx'] = test_index
75 df_res_test['proba'] = y_prob_test[:,1]
76 df_res_test['GT'] = list(y_test)
77 df_res_test['train_test'] = 1
78
79 df_res_train = pd.DataFrame()
80 df_res_train['idx'] = train_index
81 df_res_train['proba'] = y_prob_train[:,1]
82 df_res_train['GT'] = list(y_train)
83 df_res_train['train_test'] = 0
84
85 df_res_test = df_res_test.append(df_res_train)
86
87 df_res_test['split'] = n_split
88
89 df_res_inner = pd.DataFrame(columns = ['inner_train', 'inner_test', 'model', 'variables', 'split'])
90 df_res_inner.loc[0, 'inner_train'] = inner_train
91 df_res_inner.loc[0, 'inner_test'] = inner_test
92 df_res_inner['model'] = str(gridcv.best_estimator_)
93 df_res_inner['variables'] = str(list(np.array(list(X))[gridcv.best_estimator_.named_steps['fs'].get_support()])))
94 df_res_inner['split'] = n_split
95
96 return df_res_test, df_res_inner
97
98 ### k_fold_nested
99 ## function to perform a nested cv stratified k-fold iteration in a binary classification problem and save predictions.
100 ## Model and hyperparameters are optimized on an inner loop.
101 ## Parameters: see train_split_nested
102 ## Returns: dataframe with concatenated results of each train_test split of the outer k-fold.
103 def k_fold_nested(X, y, n_folds, seed, model, param_grid, inner_rkf):
104     print( '** STARTING K-FOLD WITH SEED % d **' % (seed))
105     res_out = pd.DataFrame(columns = ['idx', 'proba', 'GT', 'train_test', 'split', 'seed'])
106     res_inner = pd.DataFrame(columns = ['inner_train', 'inner_test', 'model', 'variables', 'split', 'seed'])
107     skf = StratifiedKFold(n_splits = n_folds, random_state = seed, shuffle = True)
108     n = 1
109     for train_index, test_index in skf.split(X, y):
110         res_split_out, res_split_in = train_split_nested(X, y, train_index,
111                                                         test_index, n, model, param_grid, inner_rkf=inner_rkf)
112         n=n+1
113         res_split_out['seed'] = seed
114         res_split_in['seed'] = seed
115         res_out = res_out.append(res_split_out)
116         res_inner = res_inner.append(res_split_in)
117
118     print( '** SEED % d FINISHED**' % (seed))
119
120     return res_out, res_inner


```

Branch: master ▾

Find file

Copy path

## Restenosis\_IA\_Gracia3 / model\_zoo.py

 PedroDorado correcciones

e3a7b62 31 minutes ago

1 contributor

Raw Blame History



108 lines (90 sloc) 7.16 KB

```

1  from lib_ml import *
2
3  clf1 = LogisticRegression(random_state=0, penalty = 'l2', n_jobs = 1)
4  model1 = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
5                   ('scaler', preprocessing.StandardScaler()),
6                   ('fs', SelectKBest(score_func=f_classif)),
7                   ('clf', clf1)])
8  param_grid1 = {'fs__k': [2,4], 'clf__C': [0.1,1]}
9  model_dict1 = {'clf_name': 'median_standar_anova_lr',
10               'clf_fun': model1 ,
11               'param_grid': param_grid1}
12
13  ## LOGISTIC REGRESSION MODELS
14  model_lr = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
15                     ('scaler', preprocessing.StandardScaler()),
16                     ('fs', SelectKBest(score_func=f_classif)),
17                     ('clf', LogisticRegression(random_state=0, penalty = 'l2', n_jobs = 1))])
18  param_grid_lr = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__C': [0.1,1,10,100,1000]}
19  model_dict_lr = {'clf_name': 'median_standar_anova_lr', 'clf_fun': model_lr, 'param_grid': param_grid_lr}
20
21  model_lr2 = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
22                      ('scaler', preprocessing.StandardScaler()),
23                      ('fs', SelectFromModel(threshold=-np.inf, estimator = ExtraTreesClassifier(n_estimators = 50, random_state = 0, max_featu
24                      ('clf', LogisticRegression(random_state=0, penalty = 'l2', n_jobs = 1))])
25  param_grid_lr2 = {'fs__max_features': [2,4,6,8,10,12,14,16,18,20], 'clf__C': [0.1,1,10,100,1000]}
26  model_dict_lr2 = {'clf_name': 'median_standar_ef_lr', 'clf_fun': model_lr2, 'param_grid': param_grid_lr2}
27
28  model_lr_noreg = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
29                            ('scaler', preprocessing.StandardScaler()),
30                            ('fs', SelectKBest(score_func=f_classif)),
31                            ('clf', LogisticRegression(random_state=0, penalty = 'none', solver = 'newton-cg', n_jobs = 1))])
32  param_grid_lr_noreg = {'fs__k': [2,4,6,8,10,12,14,16,18,20]}
33  model_dict_lr_noreg = {'clf_name': 'median_standar_anova_lr_noreg', 'clf_fun': model_lr_noreg ,
34                       'param_grid': param_grid_lr_noreg}
35
36  model_lr_noreg2 = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
37                             ('scaler', preprocessing.StandardScaler()),
38                             ('fs', SelectFromModel(threshold=-np.inf, estimator = ExtraTreesClassifier(n_estimators = 50, random_state = 0, max_featu
39                             ('clf', LogisticRegression(random_state=0, penalty = 'none', solver = 'newton-cg', n_jobs = 1))])
40  param_grid_lr_noreg2 = {'fs__max_features': [2,4,6,8,10,12,14,16,18,20]}
41  model_dict_lr_noreg2 = {'clf_name': 'median_standar_ef_lr_noreg', 'clf_fun': model_lr_noreg2 ,
42                        'param_grid': param_grid_lr_noreg2}
43
44  # RANDOM FOREST MODELS
45  model_rf = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
46                      ('fs', SelectKBest(score_func=f_classif)),
47                      ('clf', RandomForestClassifier(random_state=0, n_jobs = 1, n_estimators = 500, max_features = 1))])
48  param_grid_rf = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__max_depth': [2,3,4,5,10]}
49  model_dict_rf = {'clf_name': 'median_anova_rf', 'clf_fun': model_rf, 'param_grid': param_grid_rf}
50
51  model_rf2 = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
52                       ('fs', SelectFromModel(threshold=-np.inf, estimator = ExtraTreesClassifier(n_estimators = 50, random_state = 0, max_featu
53                       ('clf', RandomForestClassifier(random_state=0, n_jobs = 1, n_estimators = 500, max_features = 1))])
54  param_grid_rf2 = {'fs__max_features': [2,4,6,8,10,12,14,16,18,20], 'clf__max_depth': [2,3,4,5,10]}
55  model_dict_rf2 = {'clf_name': 'median_ef_rf', 'clf_fun': model_rf2, 'param_grid': param_grid_rf2}
56
57  # EXTREMELY RANDOMIZED TREES MODELS
58  model_ef = Pipeline([('mv', preprocessing.Imputer(strategy='median')),

```

```

59         ('fs', SelectKBest(score_func=f_classif)),
60         ('clf', ExtraTreesClassifier(random_state = 0, n_jobs = 1, n_estimators = 500,
61                                     max_features = 1)))
62 param_grid_ef = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__max_depth': [2,3,4,5,10]}
63 model_dict_ef = {'clf_name': 'median_anova_ef', 'clf_fun': model_ef, 'param_grid': param_grid_ef}
64
65
66 model_ef2 = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
67                      ('fs', SelectFromModel(threshold=-np.inf, estimator = ExtraTreesClassifier(n_estimators = 50, random_state = 0, max_featu
68                      ('clf', ExtraTreesClassifier(random_state = 0, n_jobs = 1, n_estimators = 500,
69                                                     max_features = 1)))])
70 param_grid_ef2 = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__max_depth': [2,3,4,5,10]}
71 model_dict_ef2 = {'clf_name': 'median_ef_ef', 'clf_fun': model_ef2, 'param_grid': param_grid_ef2}
72
73 # GRADIENT BOOSTING MODELS
74 model_gb = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
75                      ('fs', SelectKBest(score_func=f_classif)),
76                      ('clf', GradientBoostingClassifier(random_state = 0, n_estimators = 500, max_features = 1)))])
77 param_grid_gb = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__max_depth': [2,3,4,5,10]}
78 model_dict_gb = {'clf_name': 'median_anova_gb', 'clf_fun': model_gb, 'param_grid': param_grid_gb}
79
80 model_gb2 = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
81                      ('fs', SelectFromModel(threshold=-np.inf, estimator = ExtraTreesClassifier(n_estimators = 50, random_state = 0, max_featu
82                      ('clf', GradientBoostingClassifier(random_state = 0, n_estimators = 500, max_features = 1)))])
83 param_grid_gb2 = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__max_depth': [2,3,4,5,10]}
84 model_dict_gb2 = {'clf_name': 'median_ef_gb', 'clf_fun': model_gb2, 'param_grid': param_grid_gb2}
85
86 # SUPPORT VECTOR MACHINES MODELS
87 model_sv = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
88                      ('scaler', preprocessing.StandardScaler()),
89                      ('fs', SelectKBest(score_func=f_classif)),
90                      ('clf', SVC(random_state = 0, kernel = 'rbf', probability = True))])
91 param_grid_sv = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__C': [0.1,1,10,100,1000],
92                  'clf__gamma': [0.1, 1, 10, 100]}
93 model_dict_sv = {'clf_name': 'median_escaler_anova_svc', 'clf_fun': model_sv, 'param_grid': param_grid_sv}
94
95 model_sv2 = Pipeline([('mv', preprocessing.Imputer(strategy='median')),
96                      ('scaler', preprocessing.StandardScaler()),
97                      ('fs', SelectFromModel(threshold=-np.inf, estimator = ExtraTreesClassifier(n_estimators = 50, random_state = 0, max_featur
98                      ('clf', SVC(random_state = 0, kernel = 'rbf', probability = True))])
99 param_grid_sv2 = {'fs__k': [2,4,6,8,10,12,14,16,18,20], 'clf__C': [0.1,1,10,100,1000],
100                  'clf__gamma': [0.1, 1, 10, 100]}
101 model_dict_sv2 = {'clf_name': 'median_escaler_ef_svc', 'clf_fun': model_sv2, 'param_grid': param_grid_sv2}
102
103
104
105 model_zoo = [model_dict_lr, model_dict_lr2, model_dict_lr_noreg, model_dict_lr_noreg2,
106              model_dict_rf, model_dict_rf2, model_dict_ef, model_dict_ef2,
107              model_dict_gb, model_dict_gb2, model_dict_sv, model_dict_sv2]


```

Branch: master ▾

Find file

Copy path

## Restenosis\_IA\_Gracia3 / script\_nested.py

 PedroDorado Add files via upload

bb23e2b 2 days ago

1 contributor

Raw Blame History



45 lines (36 sloc) 1.69 KB

```
1 from model_zoo import *
2 import mkl
3 mkl.set_num_threads(1)
4
5 # Script to run train_split_nested with outer cv k-folds and n repetitions.
6 # Inner cv is done with k-1 folds with 1 repetition.
7 # Each k-fold is run on a processor core.
8
9 # Returns dataframes with generalization results over test and training subsets
10 # and scores obtained in the inner cv k-folds
11
12 folds = 5
13 repetitions = 1
14 threads = 2
15
16 if __name__ == '__main__':
17     data = pd.read_excel('./data/clean_df.xls')
18     X = data.copy().drop('label', axis = 1)
19     y = data.loc[:, 'label']
20
21     # model_dict = model_zoo
22
23     inner_rkf = StratifiedKFold(n_splits = folds-1, random_state= 0, shuffle= True)
24     for model_dict in model_zoo:
25         print('**** TRAINING ' + model_dict['clf_name'] + ' ****')
26         res_out = pd.DataFrame(columns = ['idx', 'proba', 'GT', 'train_test', 'split', 'seed'])
27         res_inn = pd.DataFrame(columns = ['inner_train', 'inner_test', 'model', 'variables', 'split', 'seed'])
28         model = model_dict['clf_fun']
29         param_grid = model_dict['param_grid']
30
31         folds = 10
32         pool = mp.Pool(threads)
33         res = [pool.apply_async(k_fold_nested, args = (X, y, folds, seed, model, param_grid, inner_rkf))
34                for seed in range(repetitions)]
35         pool.close()
36         for p in res:
37             df_out, df_inn = p.get()
38             res_out = res_out.append(df_out)
39             res_inn = res_inn.append(df_inn)
40         res_out.to_csv('./results/out_' + model_dict['clf_name'] + '_' + str(folds) + '.csv')
41         res_inn.to_csv('./results/in_' + model_dict['clf_name'] + '_' + str(folds) + '.csv', sep=';')
42         print('**** ' + model_dict['clf_name'] + ' FINISHED **** \n')
43
44     print('FINISHED')
```

Branch: master ▾

Find file Copy path

## Restenosis\_IA\_Gracia3 / train\_final\_model.ipynb

 PedroDorado Add files via upload

bb23e2b 2 days ago

1 contributor



Raw Blame History



189 lines (188 sloc) 4.73 KB

```
In [8]: from model_zoo import *
import pickle
```

```
In [3]: data = pd.read_excel('./data/clean_df.xls')
```

```
In [4]: X = data.copy().drop('label', axis = 1)
y = data.loc[:, 'label']
```

```
In [5]: ## With a model from model_zoo, optimizes its hyperparamters with 10 fold CV and train the best estimator
## trained with the entire dataset.

## Model is saved in $clf_name.sav pickle
```

```
In [7]: model_dict = model_zoo[0]
param_grid = model_dict['param_grid']
model = model_dict['clf_fun']
model_name = model_dict['clf_name']

rkf = StratifiedKFold(n_splits=10, random_state=0, shuffle=True)
```

```
In [9]: gridcv = RandomizedSearchCV(estimator = model, param_distributions= param_grid, scoring = 'roc_auc', cv =
rkf,
refit = True, return_train_score = True, n_jobs = 4, random_state = 0, n_iter
= 20, verbose = 1)
```

```
In [10]: gridcv.fit(X,y)
```

```
Fitting 10 folds for each of 4 candidates, totalling 40 fits
[Parallel(n_jobs=4)]: Using backend LokyBackend with 4 concurrent workers.
[Parallel(n_jobs=4)]: Done 33 out of 40 | elapsed: 3.2s remaining: 0.6s
[Parallel(n_jobs=4)]: Done 40 out of 40 | elapsed: 3.3s finished
```

```
Out[10]: RandomizedSearchCV(cv=StratifiedKFold(n_splits=10, random_state=0, shuffle=True),
error_score='raise-deprecating',
estimator=Pipeline(memory=None,
steps=[('mv', Imputer(axis=0, copy=True, missing_values='NaN', strategy='median', verbose=0)), ('scaler', StandardScaler(copy=True, with_mean=True, with_std=True)), ('fs', SelectKBest(k=10, score_func=<function f_classif at 0x00000206D7136950>)), ('clf', LogisticRegression(C=1.0, class_weight=None, ... penalty='l2', random_state=0, solver='warn', tol=0.0001, verbose=0, warm_start=False))]),
fit_params=None, iid='warn', n_iter=20, n_jobs=4,
param_distributions={'fs_k': [2, 4], 'clf_C': [0.1, 1]},
pre_dispatch='2*n_jobs', random_state=0, refit=True,
return_train_score=True, scoring='roc_auc', verbose=1)
```

```
In [18]: best_model = gridcv.best_estimator_
filename = './models/' + model_name + '_model.sav'
```

```
In [19]: pickle.dump(best_model, open(filename, 'wb'))
```

```
In [25]: np.array(list(X))[best_model.named_steps['fs'].get_support(indices = True)]
```

```
Out[25]: array(['diabetes', 'angio.illvess2', 'timi_post', 'lab.cho.a'],
dtype='<U19')'
```

In [ ]:

In [ ]:

Branch: master Restenosis\_IA\_Gracia3 / resultados.ipynb Find file Copy path

PedroDorado notebook resultados f0cff36 1 hour ago

1 contributor

1.04 MB Download History 🗨 🗑

In [114]:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import scipy
from scipy import interp, stats
import os

from sklearn.model_selection import RepeatedKFold, RepeatedStratifiedK
Fold, cross_val_predict, cross_val_score, cross_validate
from sklearn.metrics import confusion_matrix, roc_auc_score, precision
_recall_fscore_support, accuracy_score, roc_curve, auc, precision_recal
l_curve, average_precision_score
```

In [115]:

```
def pred_metric(y_prob, y_resp):

    y_ordered = [x for _, x in sorted(zip(y_prob, y_resp))]
    y_prob_ordered = [x for x, _ in sorted(zip(y_prob, y_resp))]
    sens = []
    spec = []
    prec = []
    nprv = []
    for i in range(len(y_ordered)):
        tn = y_ordered[0:i].count(0)
        fn = y_ordered[0:i].count(1)
        tp = y_ordered[i:].count(1)
        fp = y_ordered[i:].count(0)
        # print(i, tn, fn, tp, fp)
        if ((tp+fn) != 0):
            sens.append(tp / (tp+fn))
        else:
            sens.append(1)

        if ((tn+fp) != 0):
            spec.append(tn / (tn+fp))
        else:
            spec.append(1)

        if ((tp+fp) != 0):
            prec.append(tp / (tp+fp))
```

```

    else:
        prec.append(1)

    if((tn+fn)!=0):
        nprv.append(tn/(tn+fn))
    else:
        nprv.append(1)

    return np.array(prec), np.array(sens), np.array(spec), np.array(np
rv)
#     return np.array(sens), np.array(spec), np.array(prec), np.array(
nprv), y_ordered, y_prob_ordered

```

```

def pred_metric_ch(y_prob, y_resp):
    Prec1, Rec1, Spe1, NPV1 = [], [], [], []
    for i in np.sort(y_prob.unique()):
        cm = confusion_matrix(y_resp, y_prob >= i)
        TN, TP, FN, FP = cm[0,0], cm[1,1], cm[1,0], cm[0,1]
        if((TP + FN)!=0):
            Rec1 = Rec1 + [TP / (TP + FN)]
        else:
            Rec1 = Rec1+[1]

        if ((TP + FP)!=0):
            Prec1 = Prec1 + [TP / (TP + FP)]
        else:
            Prec1 = Prec1 + [1]

        if ((TN + FP)!=0):
            Spe1 = Spe1 + [TN / (TN + FP)]
        else:
            Spe1 = Spe1 + [1]

        if((TN+FN)!=0):
            NPV1 = NPV1 + [TN / (TN + FN)]
        else:
            NPV1 = NPV1 + [1]
    return np.array(Prec1), np.array(Rec1), np.array(Spe1), np.array(N
PV1)

```

In [116]:

```

## LOAD RESULTS
files = ['median_standar_anova_lr', 'median_anova_rf', 'median_anova_e
rt', 'median_anova_gb', 'median_escaler_anova_svc', 'median_standar_an
ova_lr_noreg']
files_out = files.copy()
for i,file in enumerate(files):
    files_out[i] = 'out_' + file + '_10.csv'

```

In [117]:

```
files_out
```

Out[117]:

```
['out_median_standar_anova_lr_10.csv',  
 'out_median_anova_rf_10.csv',  
 'out_median_anova_ert_10.csv',  
 'out_median_anova_gb_10.csv',  
 'out_median_escaler_anova_svc_10.csv',  
 'out_median_standar_anova_lr_noreg_10.csv']
```

In [118]:

```
df_lr = pd.read_csv('./results/out_median_standar_anova_lr_10.csv').drop('Unnamed: 0', axis = 1).tail(52600)  
df_rf = pd.read_csv('./results/out_median_anova_rf_10.csv').drop('Unnamed: 0', axis = 1).tail(52600)  
df_ert = pd.read_csv('./results/out_median_anova_ert_10.csv').drop('Unnamed: 0', axis = 1).tail(52600)  
df_svc = pd.read_csv('./results/out_median_escaler_anova_svc_10.csv').drop('Unnamed: 0', axis = 1).tail(52600)  
df_gb = pd.read_csv('./results/out_median_anova_gb_10.csv').drop('Unnamed: 0', axis = 1).tail(52600)  
df_lr_noreg = pd.read_csv('./results/out_median_standar_anova_lr_noreg_10.csv').drop('Unnamed: 0', axis = 1).tail(52600)
```

In [119]:

```
def rocs(df_lr):  
    rocs_lr = []  
    for seed in df_lr.seed.unique():  
        for split in df_lr.split.unique():  
            rocs_lr = rocs_lr + [roc_auc_score(df_lr.loc[(df_lr.seed == seed) & (df_lr.split == split) & (df_lr.train_test == 1)], 'GT'),  
                                df_lr.loc[(df_lr.seed == seed) & (df_lr.split == split) & (df_lr.train_test == 1), 'proba'])  
        ]  
    return rocs_lr  
  
def avg_pr(df_lr):  
    avg_pr_lr = []  
    for seed in df_lr.seed.unique():  
        for split in df_lr.split.unique():  
            avg_pr_lr = avg_pr_lr + [average_precision_score(df_lr.loc[(df_lr.seed == seed) & (df_lr.split == split) & (df_lr.train_test == 1)], 'GT'),  
                                    df_lr.loc[(df_lr.seed == seed) & (df_lr.split == split) & (df_lr.train_test == 1), 'proba'])  
        ]  
    return avg_pr_lr
```



```

def avg_pr_noest(df_lr):
    avg_pr_lr = []
    for seed in df_lr.seed.unique():
        for split in df_lr.split.unique():
            avg_pr_lr = avg_pr_lr + [average_precision_score(df_lr.loc
                [((df_lr.seed == seed) & (df_lr.split == split) & (df_lr.train_test ==
                1)), 'GT'],
                1 - df_lr.loc[(df_lr.se
                ed == seed) & (df_lr.split == split) & (df_lr.train_test == 1), 'prob
                a'], pos_label = 0)]
    return avg_pr_lr

```

In [120]:

```

rocs_lr = rocs(df_lr)
rocs_rf = rocs(df_rf)
rocs_ert = rocs(df_ert)
rocs_svc = rocs(df_svc)
rocs_gb = rocs(df_gb)
rocs_lr_noreg = rocs(df_lr_noreg)

```

In [121]:

```

def mean_confidence_interval(data, confidence=0.95):
    a = 1.0 * np.array(data)
    n = len(a)
    m, se = np.mean(a), scipy.stats.sem(a)
    h = se * scipy.stats.t.ppf((1 + confidence) / 2., n-1)
    return m, m-h, m+h

def mean_confidence_interval_corrected(data, confidence=0.95):
    a = 1.0 * np.array(data)
    n = len(a)
    m = np.mean(a)
    var = np.var(a)
    var_bess = var * n / (n - 1)
    S = np.sqrt(var_bess * (1/n + 1/9))
    h = S * scipy.stats.t.ppf((1 + confidence) / 2., n-1)
    return m, m-h, m+h

def t_test_corregido(rocs1, rocs2, confidence = 0.95):
    n = len(rocs1)
    dif = np.array(rocs1) - np.array(rocs2)
    media = dif.mean()
    var_bess = dif.var() * n / (n-1)

    std_error = np.sqrt(var_bess * ((1 / n) + 1/9))
    t_statistic = media/std_error

    pvalue = scipy.stats.t.sf(np.absolute(t_statistic), df= n-1)
    return media, pvalue

```

In [122]:

```
t_test_corregido(rocs_ert, rocs_lr)
```

Out[122]:

```
(0.030555555555555544, 0.2023766132967137)
```

In [123]:

```
df_presto = pd.read_excel('./results/df_presto.xls')
df_presto1_acc = pd.read_excel('./results/df_presto1_acc.xls')
df_presto2_acc = pd.read_excel('./results/df_presto2_acc.xls')
df_TLR = pd.read_excel('./results/df_TLR.xls')
```

In [124]:

```
rocs_presto1 = []
rocs_presto2 = []
rocs_TLR = []
rocs_presto1_acc = []
rocs_presto2_acc = []
rocs_TLR_acc = []
for seed in df_lr.seed.unique():
    for split in df_lr.split.unique():
        idx = list((df_lr.loc[(df_lr.seed == seed) & (df_lr.split == split) & (df_lr.train_test == 1)], 'idx'))
        rocs_presto1 = rocs_presto1 + [roc_auc_score(df_presto.loc[idx, 'reestenosis'], df_presto.loc[idx, 'presto1'])]
        rocs_presto2 = rocs_presto2 + [roc_auc_score(df_presto.loc[idx, 'reestenosis'], df_presto.loc[idx, 'presto2'])]
        rocs_TLR = rocs_TLR + [roc_auc_score(df_TLR.loc[idx, 'reestenosis'], df_TLR.loc[idx, 'TLR_score'])]
        rocs_presto1_acc = rocs_presto1_acc + [roc_auc_score(df_presto1_acc.loc[idx, 'reestenosis'], df_presto1_acc.loc[idx, 'presto1'])]
        rocs_presto2_acc = rocs_presto2_acc + [roc_auc_score(df_presto2_acc.loc[idx, 'reestenosis'], df_presto2_acc.loc[idx, 'presto2'])]
        rocs_TLR_acc = rocs_TLR_acc + [roc_auc_score(df_TLR.loc[idx, 'reestenosis'], df_TLR.loc[idx, 'TLR_acc_score'])]
```

In [125]:

```
print('PRESTO 1: ', mean_confidence_interval_corrected(rocs_presto1))
print('PRESTO 2: ', mean_confidence_interval_corrected(rocs_presto2))
print('TLR: ', mean_confidence_interval_corrected(rocs_TLR))
print('PRESTO 1 acc: ', mean_confidence_interval_corrected(rocs_presto1_acc))
print('PRESTO 2 acc: ', mean_confidence_interval_corrected(rocs_presto2_acc))
print('TLR acc: ', mean_confidence_interval_corrected(rocs_TLR_acc))

presto1_conf = mean_confidence_interval_corrected(rocs_presto1)
presto2_conf = mean_confidence_interval_corrected(rocs_presto2)
TLR_conf = mean_confidence_interval_corrected(rocs_TLR)
```

```

prestol_acc_conf = mean_confidence_interval_corrected(rocs_prestol_acc
)
prestol2_acc_conf = mean_confidence_interval_corrected(rocs_presto2_acc
)
TLR_acc_conf = mean_confidence_interval_corrected(rocs_TLR_acc)
PRESTO 1: (0.5672395833333334, 0.42978712900171223, 0.704692037664954
5)
PRESTO 2: (0.5802083333333333, 0.4371469845278454, 0.7232696821388213
)
TLR: (0.6174131944444444, 0.49840496303296705, 0.7364214258559217)
PRESTO 1 acc: (0.5696354166666667, 0.42548465436756233, 0.71378617896
57711)
PRESTO 2 acc: (0.5610069444444444, 0.4138755754414542, 0.708138313447
4346)
TLR acc: (0.6144444444444445, 0.49644129927028835, 0.7324475896186006
)

```

In [126]:

```

print('LR: ', mean_confidence_interval(rocs_lr))
print('RF: ', mean_confidence_interval(rocs_rf))
print('ert: ', mean_confidence_interval(rocs_ert))
print('SVC: ', mean_confidence_interval(rocs_svc))
print('GB: ', mean_confidence_interval(rocs_gb))
print('LR_noreg: ', mean_confidence_interval(rocs_lr_noreg))

lr_conf = mean_confidence_interval(rocs_lr)
rf_conf = mean_confidence_interval(rocs_rf)
ert_conf = mean_confidence_interval(rocs_ert)
svc_conf = mean_confidence_interval(rocs_svc)
gb_conf = mean_confidence_interval(rocs_gb)
lr_noreg_conf = mean_confidence_interval(rocs_lr_noreg)
LR: (0.7417534722222223, 0.7176171797358428, 0.7658897647086018)
RF: (0.7518402777777777, 0.7259682375983111, 0.7777123179572444)
ert: (0.7723090277777778, 0.7480912723541336, 0.7965267832014219)
SVC: (0.5312152777777777, 0.49851098206738326, 0.5639195734881721)
GB: (0.6783680555555555, 0.6473385050355935, 0.7093976060755174)
LR_noreg: (0.6988194444444444, 0.6734754140167117, 0.7241634748721771
)

```

In [127]:

```

print('LR: ', mean_confidence_interval_corrected(rocs_lr))
print('RF: ', mean_confidence_interval_corrected(rocs_rf))
print('ert: ', mean_confidence_interval_corrected(rocs_ert))
print('SVC: ', mean_confidence_interval_corrected(rocs_svc))
print('GB: ', mean_confidence_interval_corrected(rocs_gb))
print('LR_noreg: ', mean_confidence_interval_corrected(rocs_lr_noreg))

lr_conf = mean_confidence_interval_corrected(rocs_lr)

```

```

rf_conf = mean_confidence_interval_corrected(rocs_rf)
ert_conf = mean_confidence_interval_corrected(rocs_ert)
svc_conf = mean_confidence_interval_corrected(rocs_svc)
gb_conf = mean_confidence_interval_corrected(rocs_gb)
lr_noreg_conf = mean_confidence_interval_corrected(rocs_lr_noreg)

LR: (0.7417534722222223, 0.6254420279909381, 0.8580649164535066)
RF: (0.7518402777777777, 0.6271643624448107, 0.8765161931107448)
ert: (0.7723090277777778, 0.6556050182057345, 0.8890130373498211)
SVC: (0.5312152777777777, 0.37361510497619416, 0.6888154505793611)
GB: (0.6783680555555555, 0.5288383862055075, 0.8278977249056034)
LR_noreg: (0.6988194444444444, 0.5766879785884229, 0.820950910300466)

```

In [128]:

```

# t-test corregido, devuelve media de la diferencia entre las dos mues
tras y p-valor
t_test_corregido(rocs_ert, rocs_svc)

```

Out[128]:

```

(0.24109375, 0.0028584540146056993)

```

In [129]:

```

t_test_corregido(rocs_ert, rocs_TLR)

```

Out[129]:

```

(0.15489583333333334, 0.022011844785024097)

```

In [130]:

```

## CURVAS
#Curvas roc. Comparación de modelos

plt.figure(figsize=(15,10))
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.0])

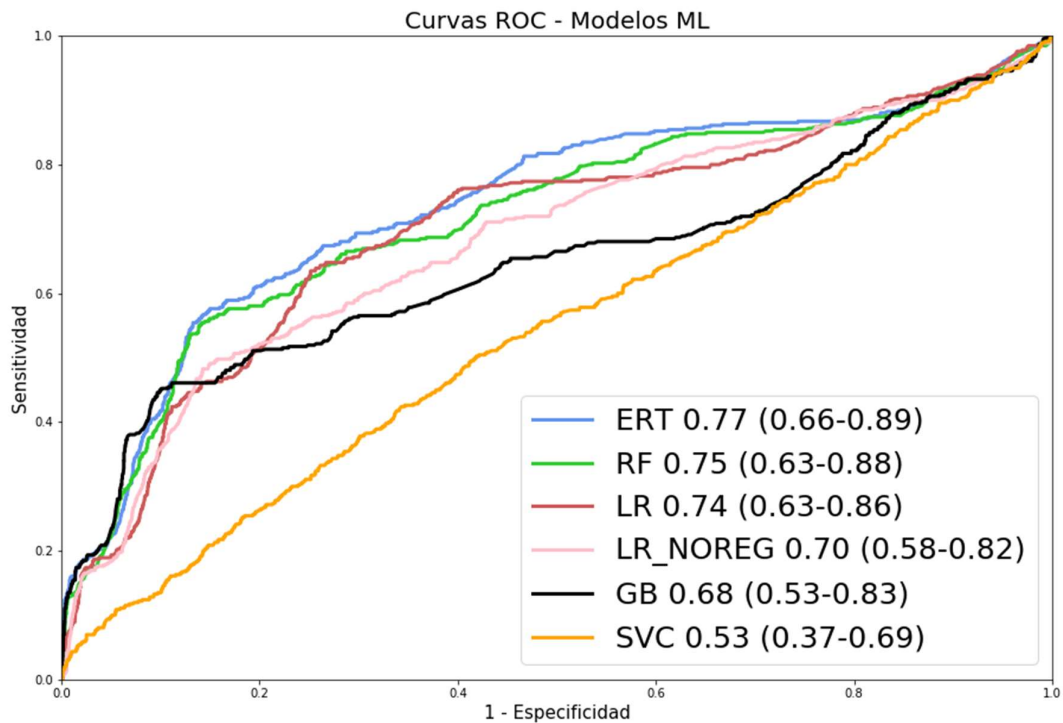
model_names = ['ert', 'rf', 'lr', 'lr_noreg', 'gb', 'svc']
colors = ['cornflowerblue', 'limegreen', 'indianred', 'pink', 'black',
'orange']
model_zoo = [eval('df_' + name) for name in model_names]
roc_CI = [eval(name + '_conf') for name in model_names]

for i, model in enumerate(model_zoo):
    df_model = model.loc[((model.train_test == 1)), :]
    fpr1, tpr1, thresholds1 = roc_curve(df_model['GT'],df_model['proba
'])
    plt.plot(fpr1, tpr1, lw=3, alpha=1, color = colors[i],
            label = model_names[i].upper() + ' %0.2f (%0.2f-%0.2f)'
% (roc_CI[i][0], roc_CI[i][1], roc_CI[i][2]))

plt.xlabel('1 - Especificidad', fontsize = 15)
plt.ylabel('Sensitividad', fontsize = 15)
plt.legend(loc="lower right", fontsize = 25)
plt.title('Curvas ROC - Modelos ML', fontsize = 20)

```

```
plt.savefig("./images/roc_modelos.png")
```



In [131]:

```
### Curvas roc. Mejor modelo y scores
```

```
plt.figure(figsize=(15,10))  
plt.xlim([0.0, 1.0])  
plt.ylim([0.0, 1.0])
```

```
model_names = ['prestol', 'presto2', 'TLR', 'prest1_acc', 'presto2_acc',  
'TLR_acc']  
colors = ['indianred', 'limegreen', 'brown', 'black', 'purple', 'orange']  
model_zoo = [df_lr, df_rf, df_ert, df_svc, df_gb, df_lr_noreg]  
roc_CI = [lr_conf, rf_conf, ert_conf, svc_conf, gb_conf, lr_noreg_conf]  
]
```

```
fpr1, tpr1, thresholds1 = roc_curve(df_ert.loc[df_ert.train_test == 1,  
'GT'],df_ert.loc[df_ert.train_test == 1,'proba'])  
plt.plot(fpr1, tpr1, lw=3, alpha=1, color = 'cornflowerblue',  
label = 'ERT %0.2f (%0.2f-%0.2f)' % (ert_conf[0], ert_conf[1]  
, ert_conf[2]))
```

```
fpr2, tpr2, thresholds2 = roc_curve(df_presto['reestenosis'],df_presto  
['prestol'])  
plt.plot(fpr2, tpr2, lw=3, alpha=1, color = 'indianred',  
label = 'PRESTO-1 %0.2f (%0.2f-%0.2f)' % (prestol_conf[0], pr  
estol_conf[1], prestol_conf[2]))
```

```

fpr5, tpr5, thresholds5 = roc_curve(df_presto1_acc['reestenosis'],df_p
resto1_acc['presto1'])
plt.plot(fpr5, tpr5, lw=2, alpha=1, color = 'indianred', linestyle='da
shed',
        label = 'PRESTO-1(OR) %0.2f (%0.2f-%0.2f)' % (presto1_acc_con
f[0], presto1_acc_conf[1], presto1_acc_conf[2]))

fpr3, tpr3, thresholds3 = roc_curve(df_presto['reestenosis'],df_presto
['presto2'])
plt.plot(fpr3, tpr3, lw=3, alpha=1, color = 'limegreen',
        label = 'PRESTO-2 %0.2f (%0.2f-%0.2f)' % (presto2_conf[0], pr
esto2_conf[1], presto2_conf[2]))

fpr6, tpr6, thresholds6 = roc_curve(df_presto2_acc['reestenosis'],df_p
resto2_acc['presto2'])
plt.plot(fpr6, tpr6, lw=2, alpha=1, color = 'limegreen',linestyle='das
hed',
        label = 'PRESTO-2(OR) %0.2f (%0.2f-%0.2f)' % (presto2_acc_con
f[0], presto2_acc_conf[1], presto2_acc_conf[2]))

fpr4, tpr4, thresholds4 = roc_curve(df_TLR['reestenosis'], df_TLR['TLR
_score'])
plt.plot(fpr4, tpr4, lw=3, alpha=1, color = 'orange',
        label = 'TLR %0.2f (%0.2f-%0.2f)' % (TLR_conf[0], TLR_conf[1]
, TLR_conf[2]))

fpr7, tpr7, thresholds7 = roc_curve(df_TLR['reestenosis'], df_TLR['TLR
_acc_score'])
plt.plot(fpr7, tpr7, lw=2, alpha=1, color = 'orange',linestyle='dashed
',
        label = 'TLR(OR) %0.2f (%0.2f-%0.2f)' % (TLR_acc_conf[0], TLR
_acc_conf[1], TLR_acc_conf[2]))

plt.xlabel('1 - Especificidad', fontsize = 15)
plt.ylabel('Sensitividad', fontsize = 15)
plt.legend(loc="lower right", fontsize = 20)
plt.title('Curvas ROC - Modelo ERT vs Scores de riesgo', fontsize = 20
)
plt.savefig("./images/roc_bestmodel_risk_no_threshold.png")

```

In [161]:

```
### Curvas roc. Mejor modelo y scores
```

```

plt.figure(figsize=(15,10))
plt.xlim([0.0, 1.0])
plt.ylim([0.0, 1.0])

model_names = ['presto1', 'presto2', 'TLR', 'prest1_acc', 'presto2_acc',
               ', 'TLR_acc']
colors = ['indianred', 'limegreen', 'brown', 'black', 'purple', 'orange']
model_zoo = [df_lr, df_rf, df_ert, df_svc, df_gb, df_lr_noreg]
roc_CI = [lr_conf, rf_conf, ert_conf, svc_conf, gb_conf, lr_noreg_conf]

points_rec = [0.18, 0.56, 0.81]
points_fpr = [0.018, 0.14, 0.466]

plt.scatter(points_fpr, points_rec, lw = 1, marker = "P", s=200, color = 'black',
            label = 'Puntos de corte seleccionados', zorder = 10)

fpr1, tpr1, thresholds1 = roc_curve(df_ert.loc[df_ert.train_test == 1,
'GT'],df_ert.loc[df_ert.train_test == 1,'proba'])
plt.plot(fpr1, tpr1, lw=3, alpha=1, color = 'cornflowerblue',
         label = 'ERT %0.2f (%0.2f-%0.2f)' % (ert_conf[0], ert_conf[1],
, ert_conf[2]))

fpr2, tpr2, thresholds2 = roc_curve(df_presto['reestenosis'],df_presto
['presto1'])
plt.plot(fpr2, tpr2, lw=3, alpha=1, color = 'indianred',
         label = 'PRESTO-1 %0.2f (%0.2f-%0.2f)' % (presto1_conf[0], pr
esto1_conf[1], presto1_conf[2]))

fpr5, tpr5, thresholds5 = roc_curve(df_presto1_acc['reestenosis'],df_p
resto1_acc['presto1'])
plt.plot(fpr5, tpr5, lw=2, alpha=1, color = 'indianred', linestyle='da
shed',
         label = 'PRESTO-1(OR) %0.2f (%0.2f-%0.2f)' % (presto1_acc_con
f[0], presto1_acc_conf[1], presto1_acc_conf[2]))

fpr3, tpr3, thresholds3 = roc_curve(df_presto['reestenosis'],df_presto
['presto2'])
plt.plot(fpr3, tpr3, lw=3, alpha=1, color = 'limegreen',
         label = 'PRESTO-2 %0.2f (%0.2f-%0.2f)' % (presto2_conf[0], pr
esto2_conf[1], presto2_conf[2]))

fpr6, tpr6, thresholds6 = roc_curve(df_presto2_acc['reestenosis'],df_p
resto2_acc['presto2'])

```

```

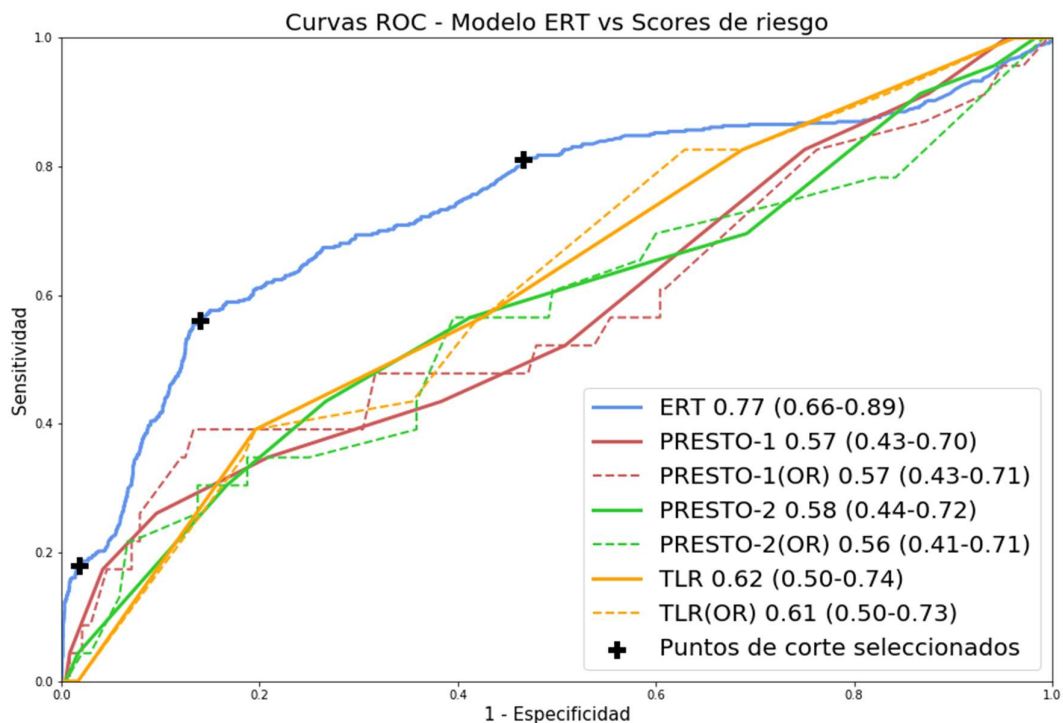
plt.plot(fpr6, tpr6, lw=2, alpha=1, color = 'limegreen',linestyle='dashed',
        label = 'PRESTO-2(OR) %0.2f (%0.2f-%0.2f)' % (presto2_acc_conf[0], presto2_acc_conf[1], presto2_acc_conf[2]))

fpr4, tpr4, thresholds4 = roc_curve(df_TLR['reestenosis'], df_TLR['TLR_score'])
plt.plot(fpr4, tpr4, lw=3, alpha=1, color = 'orange',
        label = 'TLR %0.2f (%0.2f-%0.2f)' % (TLR_conf[0], TLR_conf[1], TLR_conf[2]))

fpr7, tpr7, thresholds7 = roc_curve(df_TLR['reestenosis'], df_TLR['TLR_acc_score'])
plt.plot(fpr7, tpr7, lw=2, alpha=1, color = 'orange',linestyle='dashed',
        label = 'TLR(OR) %0.2f (%0.2f-%0.2f)' % (TLR_acc_conf[0], TLR_acc_conf[1], TLR_acc_conf[2]))

plt.xlabel('1 - Especificidad', fontsize = 15)
plt.ylabel('Sensibilidad', fontsize = 15)
plt.legend(loc="lower right", fontsize = 20)
plt.title('Curvas ROC - Modelo ERT vs Scores de riesgo', fontsize = 20)
plt.savefig("./images/roc_bestmodel_risk_threshold.png")

```



In [133]:



```

pr_lr = avg_pr(df_lr)
pr_rf = avg_pr(df_rf)
pr_ert = avg_pr(df_ert)
pr_svc = avg_pr(df_svc)
pr_gb = avg_pr(df_gb)
pr_lr_noreg = avg_pr(df_lr_noreg)

pr_lr_CI = mean_confidence_interval_corrected(pr_lr)
pr_ert_CI = mean_confidence_interval_corrected(pr_ert)
pr_rf_CI = mean_confidence_interval_corrected(pr_rf)
pr_svc_CI = mean_confidence_interval_corrected(pr_svc)
pr_gb_CI = mean_confidence_interval_corrected(pr_gb)
pr_lr_noreg_CI = mean_confidence_interval_corrected(pr_lr_noreg)

print(mean_confidence_interval_corrected(pr_lr))
print(mean_confidence_interval_corrected(pr_rf))
print(mean_confidence_interval_corrected(pr_ert))
print(mean_confidence_interval_corrected(pr_svc))
print(mean_confidence_interval_corrected(pr_gb))
print(mean_confidence_interval_corrected(pr_lr_noreg))

(0.389897750260787, 0.24340577208233893, 0.536389728439235)
(0.42619336089457577, 0.26591304670107185, 0.5864736750880797)
(0.4553348279832969, 0.28808388985395, 0.6225857661126437)
(0.2151564137809452, 0.10737602465187646, 0.3229368029100139)
(0.38427701545437104, 0.2245615726966478, 0.5439924582120943)
(0.3504742388229146, 0.20598432974805295, 0.4949641478977763)

```

In [134]:

```

pr_presto1 = []
pr_presto2 = []
pr_TLR = []
pr_presto1_acc = []
pr_presto2_acc = []
pr_TLR_acc = []
for seed in df_lr.seed.unique():
    for split in df_lr.split.unique():
        idx = list((df_lr.loc[((df_lr.seed == seed) & (df_lr.split
== split) & (df_lr.train_test == 1)), 'idx']))
        pr_presto1 = pr_presto1 + [average_precision_score(df_presto.loc[idx, 'reestenosis'], df_presto.loc[idx, 'presto1'])]
        pr_presto2 = pr_presto2 + [average_precision_score(df_presto.loc[idx, 'reestenosis'], df_presto.loc[idx, 'presto2'])]
        pr_TLR = pr_TLR + [average_precision_score(df_TLR.loc[idx, 'reestenosis'], df_TLR.loc[idx, 'TLR_score'])]
        pr_presto1_acc = pr_presto1_acc + [average_precision_score(df_presto1_acc.loc[idx, 'reestenosis'], df_presto1_acc.loc[idx, 'presto1'])]

```

```

pr_presto2_acc = pr_presto2_acc + [average_precision_score
(df_presto2_acc.loc[idx, 'reestenosis'], df_presto2_acc.loc[idx, 'presto2'])]

pr_TLR_acc = pr_TLR_acc + [average_precision_score(df_TLR.loc[idx, 'reestenosis'], df_TLR.loc[idx, 'TLR_acc_score'])]
pr_presto1_CI = mean_confidence_interval_corrected(pr_presto1)
pr_presto2_CI = mean_confidence_interval_corrected(pr_presto1)
pr_TLR_CI = mean_confidence_interval_corrected(pr_TLR)

print(mean_confidence_interval_corrected(pr_presto1))
print(mean_confidence_interval_corrected(pr_presto2))
print(mean_confidence_interval_corrected(pr_TLR))
print(mean_confidence_interval_corrected(pr_presto1_acc))
print(mean_confidence_interval_corrected(pr_presto2_acc))
print(mean_confidence_interval_corrected(pr_TLR_acc))

(0.2527153629813636, 0.12450434162258675, 0.3809263843401405)
(0.22884570083255387, 0.11859316736841989, 0.33909823429668784)
(0.16623038948669921, 0.11297454724074832, 0.2194862317326501)
(0.2900760068027138, 0.1519053762835594, 0.4282466373218682)
(0.24859465250549767, 0.12048697210095174, 0.3767023329100436)
(0.17472549281130095, 0.11884775156095177, 0.23060323406165012)

```

In [135]:

```

pr_lr_CI = mean_confidence_interval_corrected(pr_lr)
pr_ert_CI = mean_confidence_interval_corrected(pr_ert)

```

In [136]:

```

plt.figure(figsize = (10,10))
# plt.grid()
plt.title('Curvas PR para la clase restenosis - Modelos ML', fontsize = 20)

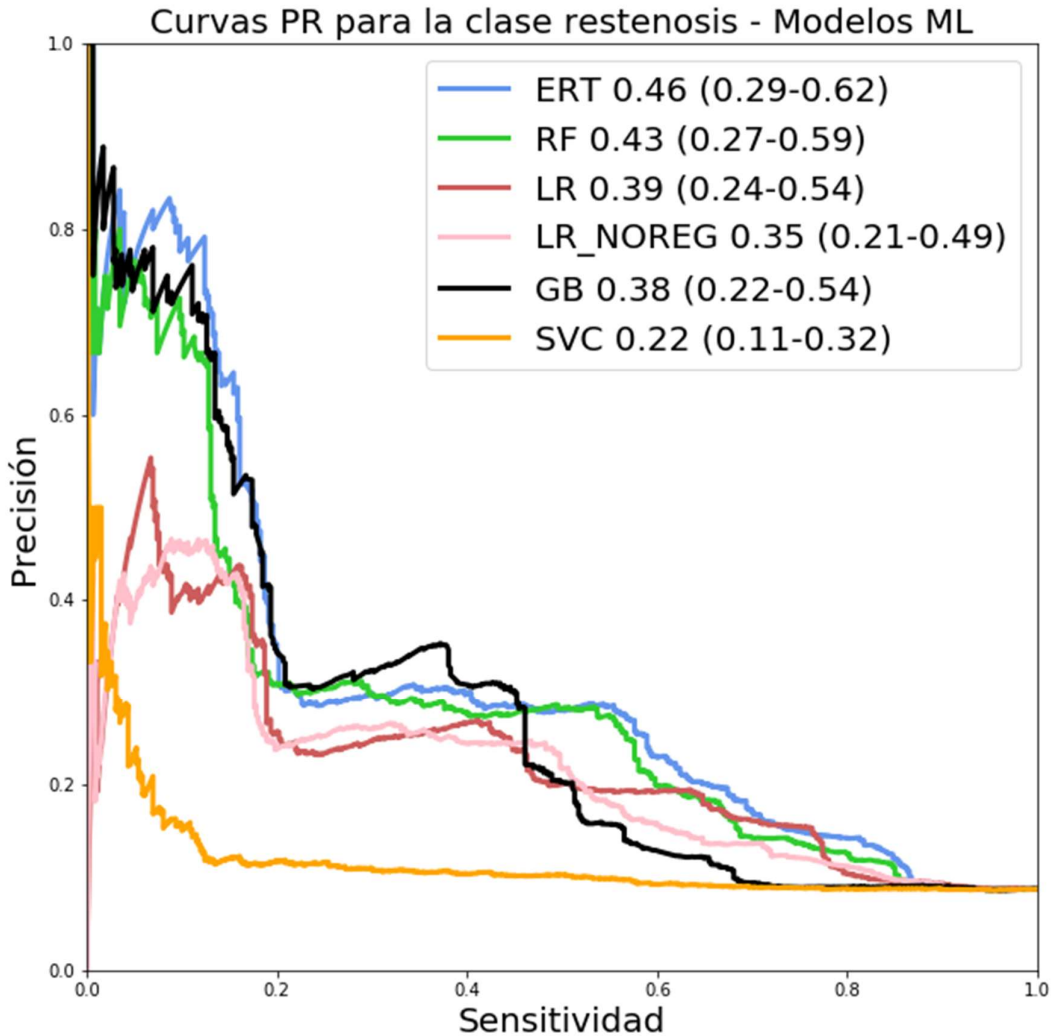
model_names = ['ert', 'rf', 'lr', 'lr_noreg', 'gb', 'svc']
colors = ['cornflowerblue', 'limegreen', 'indianred', 'pink', 'black', 'orange']
model_zoo = [eval('df_' + name) for name in model_names]
roc_CI = [eval('pr_' + name + '_CI') for name in model_names]

for i, model in enumerate(model_zoo):
    df_model = model.loc[((model.train_test == 1)), :]
    precision1, recall1, thresholds1 = precision_recall_curve(df_model.loc[:, 'GT'], df_model['proba'])
    plt.plot(recall1, precision1, lw=3, alpha=1, color = colors[i],
             label = model_names[i].upper() + ' %0.2f (%0.2f-%0.2f)'
             % (roc_CI[i][0], roc_CI[i][1], roc_CI[i][2]))

plt.ylabel('Precisión', fontsize = 20)
plt.xlabel('Sensitividad', fontsize = 20)
plt.legend(loc="best", fontsize = 20)

```

```
plt.ylim([0.0, 1.0])
plt.xlim([0.0, 1.0])
plt.savefig('./images/pr_est_modelos.png')
```



In [147]:

```
plt.figure(figsize = (10,10))
plt.title('Curvas PR para la clase restenosis\nModelo ERT vs Scores de
riesgo', fontsize = 20)
precision11, recall11, thresholds11 = precision_recall_curve(df_ert.loc[
df_ert.train_test ==1 , 'GT'], df_ert.loc[df_ert.train_test ==1 , 'proba
'], pos_label=1)
precision12, recall12, thresholds12 = precision_recall_curve(df_lr.loc[
df_lr.train_test ==1 , 'GT'], df_lr.loc[df_lr.train_test ==1 , 'proba
'], pos_label=1)

precision3, recall3, thresholds3 = precision_recall_curve(df_presto['r
eestenosis'], df_presto['presto1'], pos_label=1)
precision4, recall4, thresholds4 = precision_recall_curve(df_presto['r
eestenosis'], df_presto['presto2'], pos_label=1)
```

```

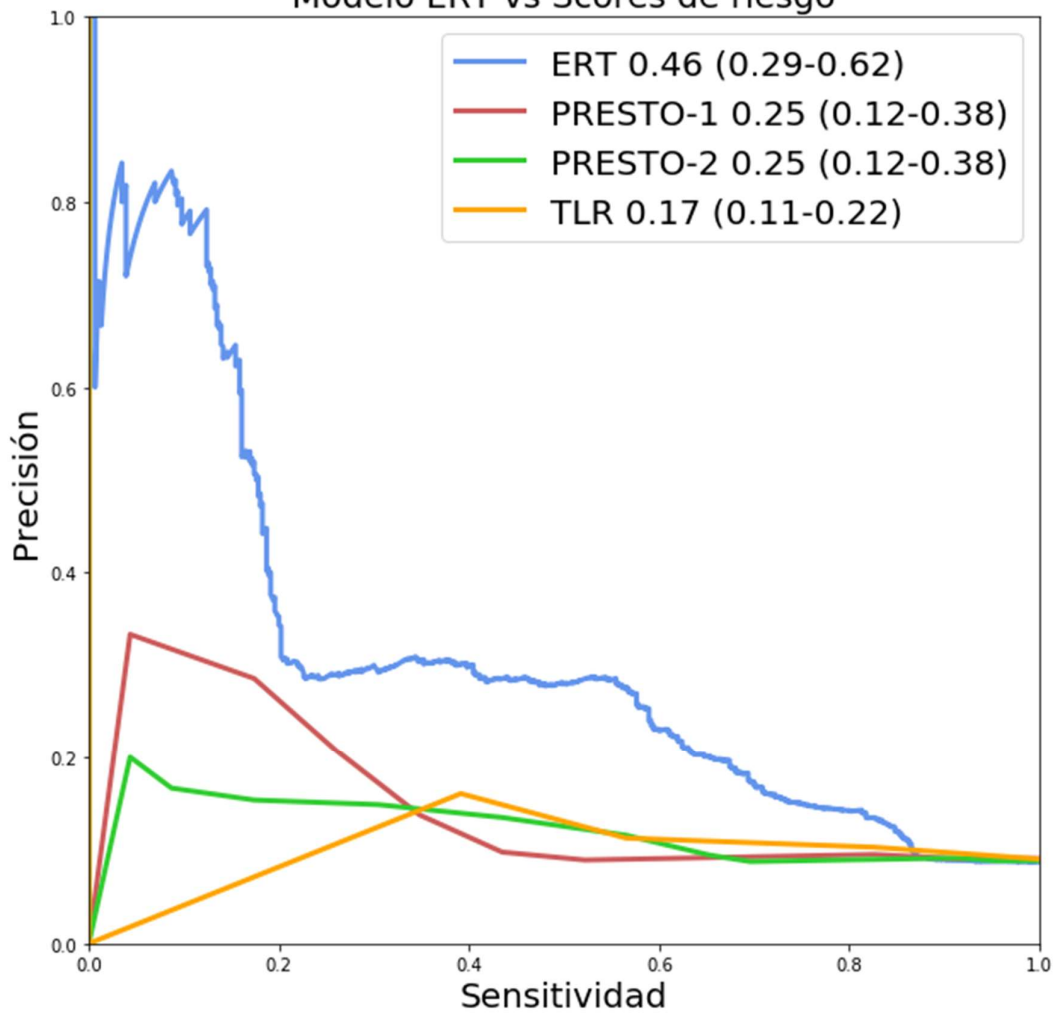
precision5, recall5, thresholds5 = precision_recall_curve(df_TLR['rees
tenosis'], df_TLR['TLR_score'], pos_label=1)

plt.plot(recall11, precision11, lw = 3, color = 'cornflowerblue', label = 'ERT %0.2f (%0.2f-%0.2f)' % (pr_ert_CI[0], pr_ert_CI[1], pr_ert_CI[2]))
plt.plot(recall13, precision3, lw = 3, color = 'indianred', label = 'PRESTO-1 %0.2f (%0.2f-%0.2f)' % (pr_presto1_CI[0], pr_presto1_CI[1], pr_presto1_CI[2]))
plt.plot(recall14, precision4, lw = 3, color = 'limegreen', label = 'PRESTO-2 %0.2f (%0.2f-%0.2f)' % (pr_presto2_CI[0], pr_presto2_CI[1], pr_presto2_CI[2]))
plt.plot(recall5, precision5, lw = 3, color = 'orange', label = 'TLR %0.2f (%0.2f-%0.2f)' % (pr_TLR_CI[0], pr_TLR_CI[1], pr_TLR_CI[2]))

plt.ylabel('Precisión', fontsize = 20)
plt.xlabel('Sensitividad', fontsize = 20)
plt.legend(loc="best", fontsize = 20)
plt.ylim([0.0, 1.0])
plt.xlim([0.0, 1.0])
plt.savefig('./images/pr_curves_est_no_threshold.png')

```

### Curvas PR para la clase restenosis Modelo ERT vs Scores de riesgo



In [146]:

```
plt.figure(figsize = (10,10))
plt.title('Curvas PR para la clase restenosis\nModelo ERT vs Scores de
riesgo', fontsize = 20)
precision11, recall11, thresholds11 = precision_recall_curve(df_ert.loc[
df_ert.train_test ==1 , 'GT'], df_ert.loc[df_ert.train_test ==1 , 'p
roba'], pos_label=1)
precision12, recall12, thresholds12 = precision_recall_curve(df_lr.loc
[df_lr.train_test ==1 , 'GT'], df_lr.loc[df_lr.train_test ==1 , 'proba
'], pos_label=1)

precision3, recall3, thresholds3 = precision_recall_curve(df_presto['r
eestenosis'], df_presto['presto1'], pos_label=1)
precision4, recall4, thresholds4 = precision_recall_curve(df_presto['r
eestenosis'], df_presto['presto2'], pos_label=1)
precision5, recall5, thresholds5 = precision_recall_curve(df_TLR['rees
tenosis'], df_TLR['TLR_score'], pos_label=1)
```

```

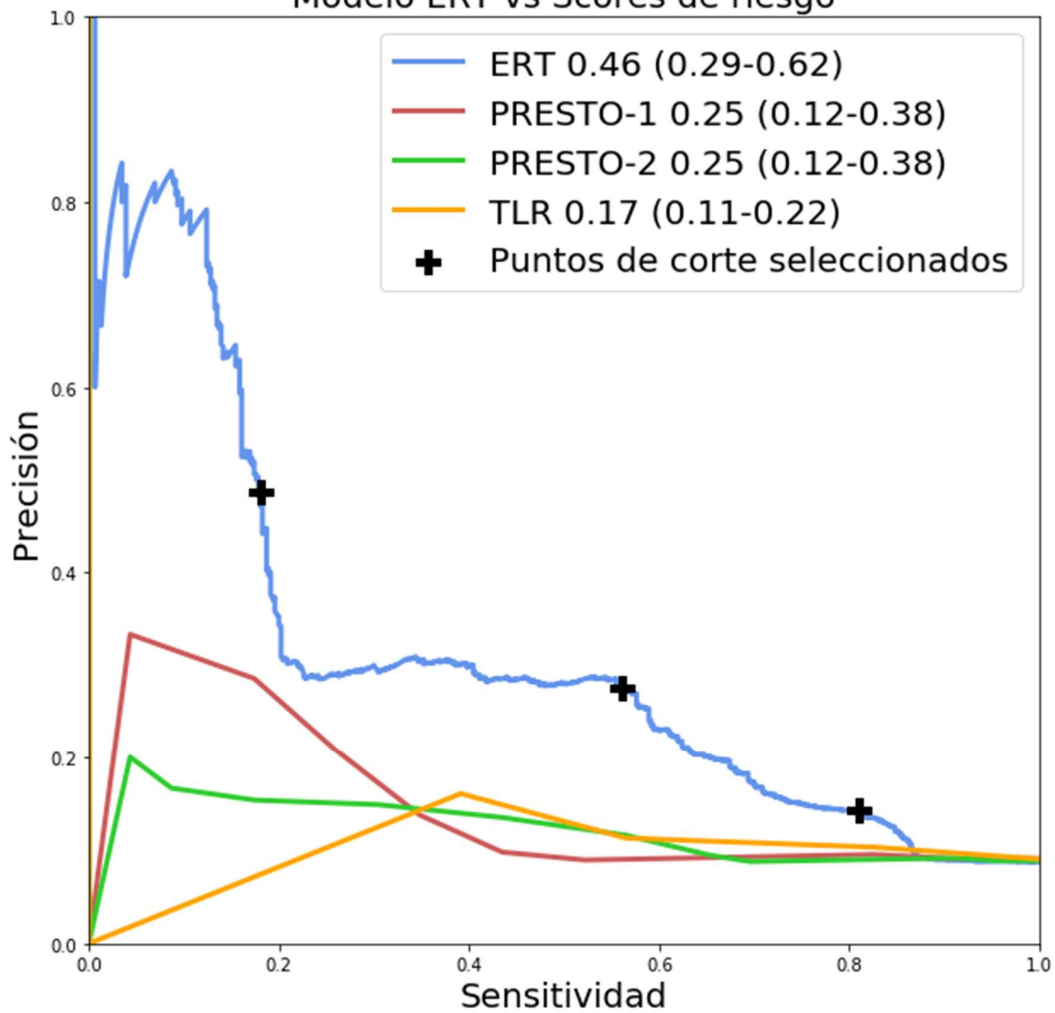
points_rec = [0.18, 0.56, 0.81]
points_prec = [0.488, 0.276, 0.143]
plt.scatter(points_rec, points_prec, marker = "P", s= 200, color = 'black',
            label = 'Puntos de corte seleccionados', zorder = 10)

plt.plot(recall11, precision11, lw = 3, color = 'cornflowerblue', label = 'ERT %0.2f (%0.2f-%0.2f)' % (pr_ert_CI[0], pr_ert_CI[1], pr_ert_CI[2]))
plt.plot(recall13, precision3, lw = 3, color = 'indianred', label = 'PRESTO-1 %0.2f (%0.2f-%0.2f)' % (pr_presto1_CI[0], pr_presto1_CI[1], pr_presto1_CI[2]))
plt.plot(recall14, precision4, lw = 3, color = 'limegreen', label = 'PRESTO-2 %0.2f (%0.2f-%0.2f)' % (pr_presto2_CI[0], pr_presto2_CI[1], pr_presto2_CI[2]))
plt.plot(recall15, precision5, lw = 3, color = 'orange', label = 'TLR %0.2f (%0.2f-%0.2f)' % (pr_TLR_CI[0], pr_TLR_CI[1], pr_TLR_CI[2]))

plt.ylabel('Precisión', fontsize = 20)
plt.xlabel('Sensitividad', fontsize = 20)
plt.legend(loc="best", fontsize = 20)
plt.ylim([0.0, 1.0])
plt.xlim([0.0, 1.0])
plt.savefig('./images/pr_curves_est_threshold.png')

```

### Curvas PR para la clase restenosis Modelo ERT vs Scores de riesgo



In [139]:

```
pr_norest_ert = avg_pr_noest(df_ert)
pr_norest_lr = avg_pr_noest(df_lr)
pr_norest_rf = avg_pr_noest(df_rf)
pr_norest_gb = avg_pr_noest(df_gb)
pr_norest_svc = avg_pr_noest(df_svc)
pr_norest_lr_noreg = avg_pr_noest(df_lr_noreg)
```

```
pr_norest_ert_CI = mean_confidence_interval_corrected(pr_norest_ert)
pr_norest_lr_CI = mean_confidence_interval_corrected(pr_norest_lr)
pr_norest_rf_CI = mean_confidence_interval_corrected(pr_norest_rf)
pr_norest_gb_CI = mean_confidence_interval_corrected(pr_norest_gb)
pr_norest_svc_CI = mean_confidence_interval_corrected(pr_norest_svc)
pr_norest_lr_noreg_CI = mean_confidence_interval_corrected(pr_norest_lr_noreg)
```

In [140]:

```
pr_norest_prestol = []
```

```

pr_norest_presto2 = []
pr_norest_TLR = []
pr_norest_presto1_acc = []
pr_norest_presto2_acc = []
pr_norest_TLR_acc = []
for seed in df_lr.seed.unique():
    for split in df_lr.split.unique():
        idx = list((df_lr.loc[(df_lr.seed == seed) & (df_lr.split
== split) & (df_lr.train_test == 1)], 'idx'))
        pr_norest_presto1 = pr_norest_presto1 + [average_precision
_score(df_presto.loc[idx, 'reestenosis'], 1 - df_presto.loc[idx, 'pres
to1'], pos_label = 0)]
        pr_norest_presto2 = pr_norest_presto2 + [average_precision
_score(df_presto.loc[idx, 'reestenosis'], 1 - df_presto.loc[idx, 'pres
to2'], pos_label = 0)]
        pr_norest_TLR = pr_norest_TLR + [average_precision_score(d
f_TLR.loc[idx, 'reestenosis'], 1 - df_TLR.loc[idx, 'TLR_score'], pos_l
abel = 0)]
        pr_norest_presto1_acc = pr_norest_presto1_acc + [average_p
recision_score(df_presto1_acc.loc[idx, 'reestenosis'], 1 - df_presto1_
acc.loc[idx, 'presto1'], pos_label = 0)]
        pr_norest_presto2_acc = pr_norest_presto2_acc + [average_p
recision_score(df_presto2_acc.loc[idx, 'reestenosis'], 1 - df_presto2_
acc.loc[idx, 'presto2'], pos_label = 0)]
        pr_norest_TLR_acc = pr_norest_TLR_acc + [average_precision
_score(df_TLR.loc[idx, 'reestenosis'], 1 - df_TLR.loc[idx, 'TLR_acc_sc
ore'], pos_label = 0)]
pr_norest_presto1_CI = mean_confidence_interval_corrected(pr_norest_pr
estol)
pr_norest_presto2_CI = mean_confidence_interval_corrected(pr_norest_pr
estol)
pr_norest_TLR_CI = mean_confidence_interval_corrected(pr_norest_TLR)

print(mean_confidence_interval_corrected(pr_norest_presto1))
print(mean_confidence_interval_corrected(pr_norest_presto2))
print(mean_confidence_interval_corrected(pr_norest_TLR))
print(mean_confidence_interval_corrected(pr_norest_presto1_acc))
print(mean_confidence_interval_corrected(pr_norest_presto2_acc))
print(mean_confidence_interval_corrected(pr_norest_TLR_acc))

(0.9307689379704229, 0.9052878214485528, 0.9562500544922931)
(0.9319939722144762, 0.90429344262248, 0.9596945018064724)
(0.9376750620638936, 0.915092585232936, 0.9602575388948512)
(0.9290269685957093, 0.8987782582316656, 0.9592756789597531)
(0.9252894146787249, 0.8930972078565029, 0.957481621500947)
(0.9400486243190431, 0.9168180842676855, 0.9632791643704007)

```

In [141]:

```
t_test_corregido(pr_norest_ert, pr_norest_TLR)
```



Out[141]:

```
(0.02620644983398148, 0.04169384092464181)
```

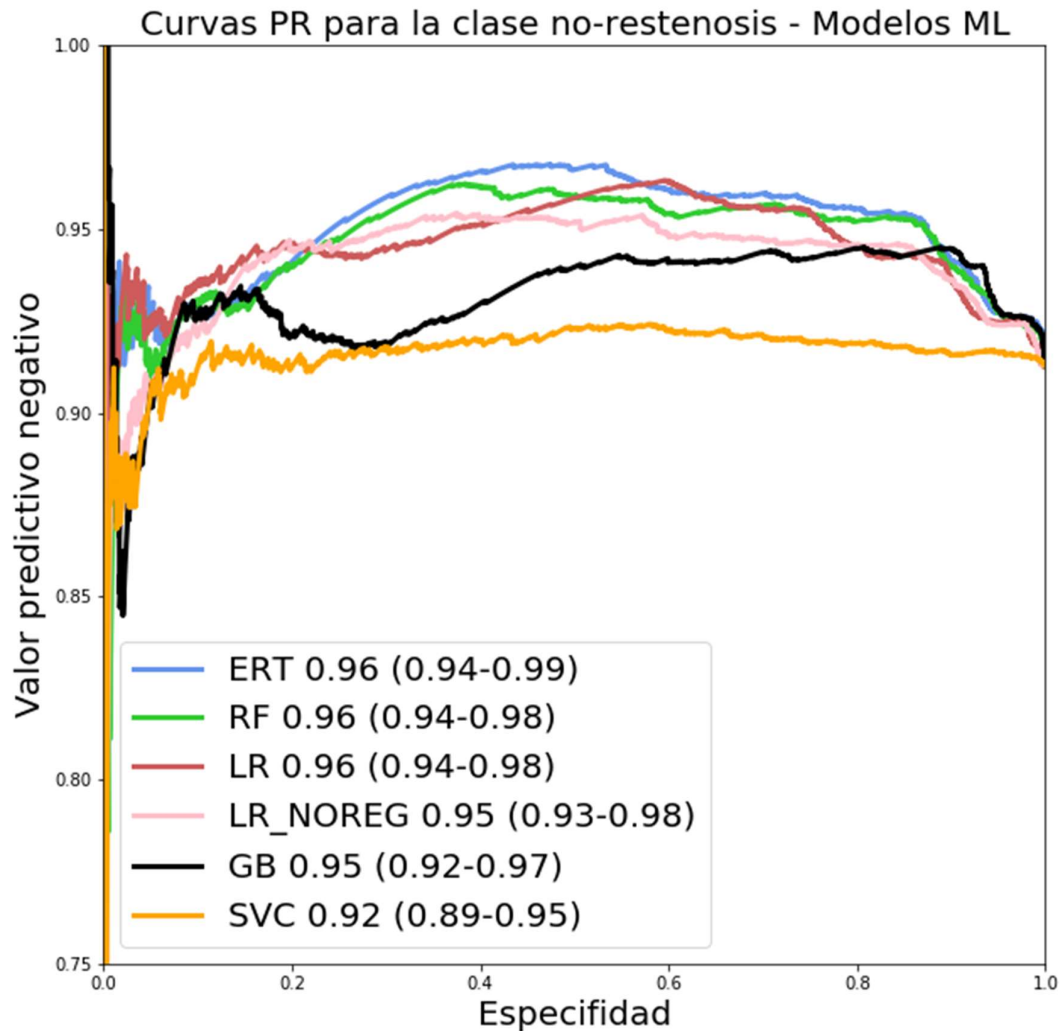
In [148]:

```
plt.figure(figsize = (10,10))
# plt.grid()
plt.title("Curvas PR para la clase no-restenosis - Modelos ML", fontsize = 20)

model_names = ['ert', 'rf', 'lr', 'lr_noreg', 'gb', 'svc']
colors = ['cornflowerblue', 'limegreen', 'indianred', 'pink', 'black',
'orange']
model_zoo = [eval('df_' + name) for name in model_names]
roc_CI = [eval('pr_norest_' + name + '_CI') for name in model_names]

for i, model in enumerate(model_zoo):
    df_model = model.loc[((model.train_test == 1)), :]
    Prec1, Rec1, Spel, NPV1 = pred_metric(df_model['proba'], df_model[
'GT'])
    plt.plot(Spel, NPV1, lw=3, alpha=1, color = colors[i],
            label = model_names[i].upper() + ' %0.2f (%0.2f-%0.2f)'
% (roc_CI[i][0], roc_CI[i][1], roc_CI[i][2]))

plt.ylabel('Valor predictivo negativo', fontsize = 20)
plt.xlabel('Especificidad', fontsize = 20)
plt.legend(loc="best", fontsize = 20)
plt.ylim([0.75, 1.0])
plt.xlim([0.0, 1.0])
plt.savefig('./images/npv_curves_models.png')
```



In [149]:

```
## Para curvas PR de scores
Prec1, Rec1, Spe1, NPV1 = pred_metric_ch(df_presto['presto1'], df_presto['reestenosis'])
Prec2, Rec2, Spe2, NPV2 = pred_metric_ch(df_presto['presto2'], df_presto['reestenosis'])
Prec3, Rec3, Spe3, NPV3 = pred_metric_ch(df_TLR['TLR_score'], df_TLR['reestenosis'])
```

```
## Para curvas PR de modelos
Prec4, Rec4, Spe4, NPV4 = pred_metric(df_ert.loc[df_ert.train_test == 1, 'proba'],
                                     df_ert.loc[df_ert.train_test == 1, 'GT'])
# Prec5, Rec5, Spe5, NPV5 = pred_metric(df_results['lr'], df_results['y_real'])
```

In [163]:

```
plt.figure(figsize = (10,10))
```

```

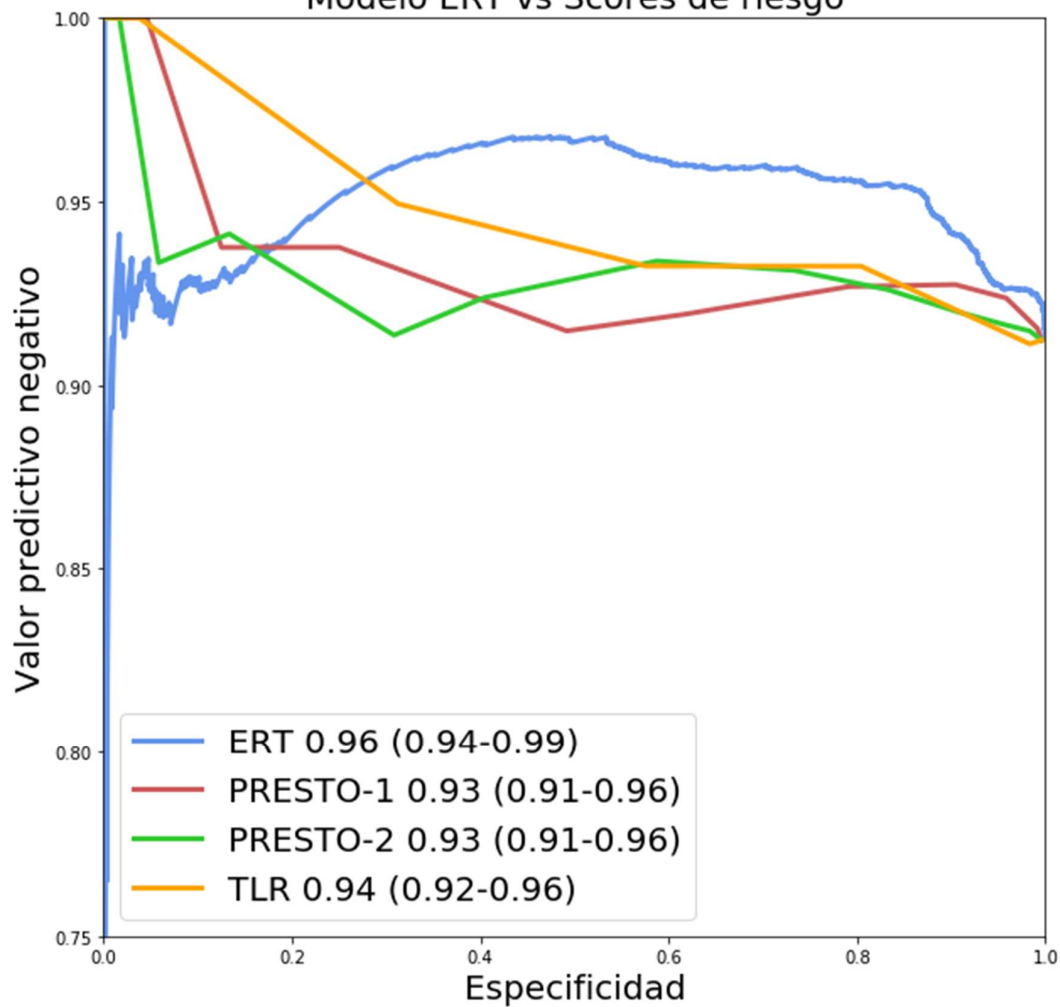
plt.title("Curvas PR para la clase no-restenosis\nModelo ERT vs Scores
de riesgo", fontsize = 20)

plt.plot(Spe4, NPV4, lw = 3, color = 'cornflowerblue', label = 'ERT %0
.2f (%0.2f-%0.2f)' % (pr_norest_ert_CI[0], pr_norest_ert_CI[1], pr_nor
est_ert_CI[2]))
plt.plot(Spe1, NPV1, lw = 3, color = 'indianred', label = 'PRESTO-1 %0
.2f (%0.2f-%0.2f)' % (pr_norest_presto1_CI[0], pr_norest_presto1_CI[1]
, pr_norest_presto1_CI[2]))
plt.plot(Spe2, NPV2, lw = 3,color = 'limegreen', label = 'PRESTO-2 %0.
2f (%0.2f-%0.2f)' % (pr_norest_presto2_CI[0], pr_norest_presto2_CI[1],
pr_norest_presto2_CI[2]))
plt.plot(Spe3, NPV3, lw = 3,color = 'orange', label = 'TLR %0.2f (%0.2
f-%0.2f)' % (pr_norest_TLR_CI[0], pr_norest_TLR_CI[1], pr_norest_TLR_C
I[2]))

plt.ylabel('Valor predictivo negativo', fontsize = 20)
plt.xlabel('Especificidad', fontsize = 20)
plt.legend(loc="best", fontsize = 20)
plt.ylim([0.75, 1.0])
plt.xlim([0.0, 1.0])
plt.savefig('./images/npv_curves_no_threshold.png')

```

Curvas PR para la clase no-restenosis  
Modelo ERT vs Scores de riesgo



In [162]:

```
plt.figure(figsize = (10,10))
plt.title("Curvas PR para la clase no-restenosis - Modelos ML", fontsize = 20)

plt.plot(Spe4, NPV4, lw = 3, color = 'cornflowerblue', label = 'ERT %0.2f (%0.2f-%0.2f)' % (pr_norest_ert_CI[0], pr_norest_ert_CI[1], pr_norest_ert_CI[2]))
plt.plot(Spe1, NPV1, lw = 3, color = 'indianred', label = 'PRESTO-1 %0.2f (%0.2f-%0.2f)' % (pr_norest_presto1_CI[0], pr_norest_presto1_CI[1], pr_norest_presto1_CI[2]))
plt.plot(Spe2, NPV2, lw = 3, color = 'limegreen', label = 'PRESTO-2 %0.2f (%0.2f-%0.2f)' % (pr_norest_presto2_CI[0], pr_norest_presto2_CI[1], pr_norest_presto2_CI[2]))
plt.plot(Spe3, NPV3, lw = 3, color = 'orange', label = 'TLR %0.2f (%0.2f-%0.2f)' % (pr_norest_TLR_CI[0], pr_norest_TLR_CI[1], pr_norest_TLR_CI[2]))
```

```

points_spe = [0.53, 0.86, 0.98]
points_npv = [0.967, 0.953, 0.925]
plt.scatter(points_spe, points_npv, marker = "P", s= 200, color = 'black', label = 'Puntos de corte seleccionados', zorder = 10)

plt.ylabel('Valor predictivo negativo', fontsize = 20)
plt.xlabel('Especificidad', fontsize = 20)
plt.legend(loc="best", fontsize = 20)
plt.ylim([0.75, 1.0])
plt.xlim([0.0, 1.0])
plt.savefig('./images/npv_curves_threshold.png')

```

In [152]:

```

## CUTOFFS
cm = confusion_matrix(df_ert.loc[df_ert.train_test == 1, 'GT'],
                      df_ert.loc[df_ert.train_test == 1, 'proba'] >= 0
                      .68) * 100 / 5260
TN, TP, FN, FP = cm[0,0], cm[1,1], cm[1,0], cm[0,1]
Prec = TP / (TP + FP)
Rec = TP / (TP + FN)
Spe = TN / (TN + FN)
NPV = TN / (TN + FP)
print(cm)
print(Prec, Rec, Spe, NPV)

[[9.12167300e+01 3.80228137e-02]
 [8.66920152e+00 7.60456274e-02]]
0.6666666666666666 0.008695652173913044 0.9132089836315188 0.999583333
3333333

```

In [153]:

```

## Cálculo de números con diferentes thresholds
from scipy import interpolate
f_tlr = interpolate.interpld(Rec1, Prec1)
f_ert = interpolate.interpld(Rec4, Prec4)

```

In [154]:

```

prev = 23 / (263)
REC = 0.81
PREC = float(f_ert(REC))
real_pos = 100*prev
TP = REC * real_pos
FN = real_pos - TP
pred_pos = TP / PREC
FP = TP / PREC - TP
TN = 100 - TP - FN - FP
NPV = TN / (TN + FN)
SPE = TN / (TN + FP)

```

In [155]:

```

SPE, PREC, NPV, TP, TN, FP, FN, 1-SPE

```

Out[155]:

```
(0.5337081030647165,
 0.14271481624744597,
 0.9670088945852643,
 7.083650190114069,
 48.703401040126224,
 42.55135181158481,
 1.6615969581749042,
 0.4662918969352835)
```

**FEATURE IMPORTANCES**

In [156]:

```
weights = pd.read_excel('./results/feature_importances2.xls')
```

In [157]:

```
list(weights.head(6) ['feature'])
```

Out[157]:

```
['diabetes',
 'angio.illvess2',
 'timi_post',
 'lab.plt.a',
 'thrombus_post',
 'lab.cho.a']
```

In [158]:

```
weights['weight_scaled'] = weights['weight'] / weights['weight'].max()
```

In [159]:

```
weights.head(11)
```

Out[159]:

	feature	weight	std	weight_scaled
0	diabetes	0.114200	0.019100	1.000000
1	angio.illvess2	0.060154	0.018909	0.526743
2	timi_post	0.058077	0.011878	0.508553
3	lab.plt.a	0.034912	0.001844	0.305713
4	thrombus_post	0.023578	0.013999	0.206462
5	lab.cho.a	0.016383	0.009905	0.143458
6	perc_diam_post	0.000000	0.000000	0.000000
7	obs_length_pre	0.000000	0.000000	0.000000
8	perc_diam_pre	0.000000	0.000000	0.000000
9	perc_area_sten_pre	0.000000	0.000000	0.000000
10	thrombus_pre	0.000000	0.000000	0.000000

In [160]:

```
plt.figure(figsize=(25,10))
```

```

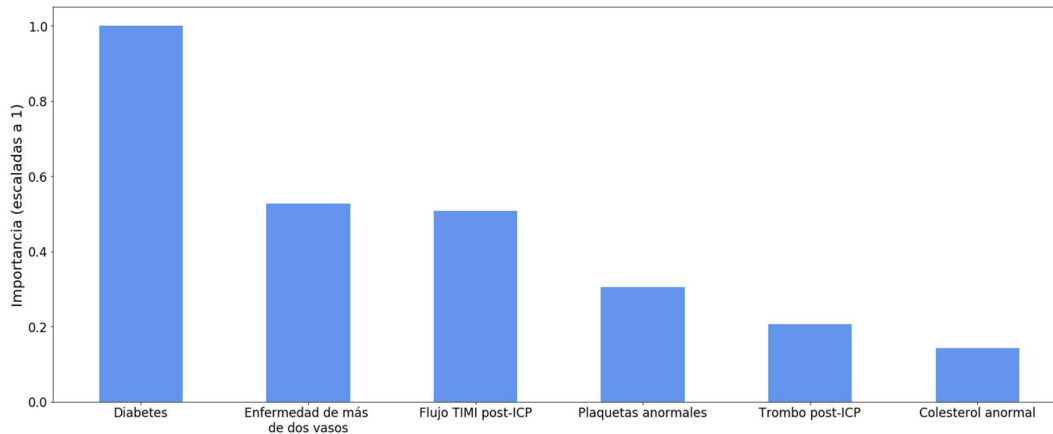
plt.xticks(rotation=0, fontsize = 17)
plt.yticks(rotation=0, fontsize = 17)

nombres = ['Diabetes', u'\u2265' + ' 2 vessel\ndisease', 'Post-PCI\nt
IMI flow', 'Abnormal\nplatelets',
           'Post-PCI\nthrombus', 'Abnormal\ntotal\ncholesterol']

nombres = ['Diabetes', 'Enfermedad de más \nde dos vasos', 'Flujo TIMI
post-ICP', 'Plaquetas anormales',
           'Trombo post-ICP', 'Colesterol anormal']

plt.bar(nombres, weights.loc[:5,'weight_scaled'], width = 0.5, color =
'cornflowerblue')
plt.xticks(rotation=0)
plt.ylabel('Importancia (escaladas a 1)', fontsize = 20)
plt.savefig('./images/importancias.png')

```



## **Anexo 4. Código fuente calculadoras desarrolladas en Python para predecir reestenosis del stent tras ICP**

### **Anexo 4.1 Calculadora Google Colab**

**restenosis\_risk\_model.py:** código Python en Google Colab con la calculadora.

**funcs\_bokeh.py:** librería para representar al paciente en las curvas ROC y PR.

<https://colab.research.google.com/drive/1jwGpC1rhkVKZi3sWdc-yKm2VxVsgMFUM#scrollTo=9LQCONSXRU56>

### **Anexo 4.2 Calculadora Tkinter**

**calculator.py:** código Python utilizando Tkinter para la calculadora.

**funcs\_bokeh.py:** librería para representar al paciente en las curvas ROC y PR.

[https://github.com/PedroDorado/Restenosis\\_IA\\_Gracia3/tree/master/calculator](https://github.com/PedroDorado/Restenosis_IA_Gracia3/tree/master/calculator)

Anexo 4. Código fuente Python calculadoras reestenosis. PDF 7 caras.



Branch: master ▾

Find file

Copy path

## Restenosis\_IA\_Gracia3 / calculator / funcs\_bokeh.py

 PedroDorado Add files via upload

bb23e2b 2 days ago

1 contributor

Raw Blame History



102 lines (83 sloc) 3.35 KB

```
1 import os
2 import numpy as np
3 import pandas as pd
4 import matplotlib.pyplot as plt
5 from sklearn.metrics import roc_curve, confusion_matrix, precision_recall_curve
6 import pickle
7
8 from bokeh.io import output_notebook, show
9 from bokeh.plotting import figure
10 from bokeh.models.markers import CircleX
11 from bokeh.util.browser import view
12 from bokeh.document import Document
13 from bokeh.embed import file_html
14 from bokeh.layouts import column, gridplot
15 from bokeh.models import Circle, ColumnDataSource, Div, Grid, Line, LinearAxis, Plot, Range1d, HoverTool
16 from bokeh.resources import INLINE
17
18 def pred_metric(y_prob,y_resp):
19
20     y_ordered = [x for _,x in sorted(zip(y_prob,y_resp))]
21     y_prob_ordered = [x for x,_ in sorted(zip(y_prob,y_resp))]
22     sens = []
23     spec = []
24     prec = []
25     nprv = []
26     for i in range(len(y_ordered)):
27         tn = y_ordered[0:i].count(0)
28         fn = y_ordered[0:i].count(1)
29         tp = y_ordered[i:].count(1)
30         fp = y_ordered[i:].count(0)
31     #     print(i, tn, fn, tp, fp)
32         if((tp+fn)!=0):
33             sens.append(tp/(tp+fn))
34         else:
35             sens.append(1)
36
37         if((tn+fp)!=0):
38             spec.append(tn/(tn+fp))
39         else:
40             spec.append(1)
41
42         if((tp+fp)!=0):
43             prec.append(tp/(tp+fp))
44         else:
45             prec.append(1)
46
47         if((tn+fn)!=0):
48             nprv.append(tn/(tn+fn))
49         else:
50             nprv.append(1)
51
52     return np.array(prec), np.array(sens), np.array(spec), np.array(nprv)
53
54
55 def point_scores(data, proba):
56     cm = confusion_matrix(data.loc[:, 'GT'],
57                           data.loc[:, 'proba'] >= proba) * 100 / 5260
58     TN, TP, FN, FP = cm[0,0], cm[1,1], cm[1,0], cm[0,1]
```


```
59     PREC = TP / (TP + FP)
60     REC = TP / (TP + FN)
61     SPE = TN / (TN + FP)
62     NPV = TN / (TN + FN)
63     return PREC, REC, SPE, NPV
64
65
66 def gracia_curves(data):
67     fpr1, tpr1, _ = roc_curve(data['GT'], data['proba'])
68     prec1, rec1, _ = precision_recall_curve(data['GT'], data['proba'])
69     _, _, spe1, npv1 = pred_metric(data['proba'], data['GT'])
70     return fpr1, tpr1, prec1, rec1, spe1, npv1
71
72
73 def figure_bokeh(curva, thresholds, x_label, y_label, x, y, x_p, y_p, colors, legend_pos,
74                 range_x = (0,1), range_y=(0,1)):
75     patchx = []
76     patchy_el = [0, 0, 1, 1]
77     patchy = []
78     alphas = [0.3, 0.3, 0.3, 0.3]
79     for i in range(1, len(thresholds), 1):
80         patchx = patchx + [[thresholds[i-1], thresholds[i], thresholds[i], thresholds[i-1]]]
81         patchy = patchy + [patchy_el]
82
83     p = figure(plot_width=700, plot_height=550, title="PUNTO SOBRE LA CURVA " + curva,
84              x_axis_label=x_label, y_axis_label = y_label,
85              x_range=range_x, y_range=range_y, toolbar_location=None, tools = "")
86
87     for i in range(len(patchx)):
88         p.patch(patchx[i], patchy[i], fill_color = colors[i], alpha = alphas[i], line_width = 0)
89
90     linea = p.line(x, y, line_width = 3, legend = "MODELO ERT", name = 'CURVA')
91     punto = p.x(x_p, y_p, size = 10, line_width = 6, alpha = 1, color = 'black',
92               legend = 'PUNTO: ' + x_label + ' = %0.2f; ' % x_p + y_label + ' = %0.2f' % y_p,
93               name = 'PUNTO')
94     p.add_tools(HoverTool(tooltips=[(x_label, '@x'), (y_label, '@y')], renderers = [linea], mode='vline'))
95
96     p.legend.location = legend_pos
97     p.legend.border_line_width = 3
98     p.legend.border_line_color = "navy"
99     p.legend.border_line_alpha = 0.5
100
101     return p
```

Branch: master ▾

Find file

Copy path

Restenosis\_IA\_Gracia3 / calculator / restenosis\_risk\_model.py / Jump to ▾

 PedroDorado Add files via upload

715c122 now

1 contributor

Raw Blame History



72 lines (54 sloc) 2.84 KB

● You're using jump to definition to discover and navigate code. [Opt out](#) Beta[Learn more](#) or [give us feedback](#)

```
1 # -*- coding: utf-8 -*-
2 """Restenosis_risk_model.ipynb
3
4 Automatically generated by Colaboratory.
5
6 Original file is located at
7     https://colab.research.google.com/drive/1jwGpC1rhkVKZi3sWdc-yKm2VxVsgMFUM
8
9 Notebook to run the model developed for the study:
10
11 ### "MACHINE LEARNING TO PREDICT STENT RESTENOSIS BASED ON DAILY DEMOGRAPHIC, CLINICAL AND ANGIOGRAPHIC CHARACTERISTICS"
12
13 In order to run the notebook:
14 - Write the inputs in the next cell. In case one of the features is not available, keep the values that are already written.
15 - Click on "Runtime -> Run all". Wait the notebook to run all cells. Check the results at the end of it.
16 """
17
18 # Data inputs
19 diabetes = 1
20 vessels_gt2 = 0
21 thrombust_post_pci = 0
22 timi_post_pci = 3
23 abnormal_cholesterol = 0
24 abnormal_platelets = 0
25
26 import os
27 if os.path.exists('./finalized_model.sav') == False:
28     !wget -nv https://raw.githubusercontent.com/IA-Cardiologia-husa/Restenosis/master/models/finalized_model.sav
29
30 if os.path.exists('./probs.csv') == False:
31     !wget -nv https://raw.githubusercontent.com/IA-Cardiologia-husa/Restenosis/master/models/probs.csv
32
33 if os.path.exists('./funcs_bokeh.py') == False:
34     !wget -nv https://raw.githubusercontent.com/IA-Cardiologia-husa/Restenosis/master/calculator/funcs_bokeh.py
35
36 from funcs_bokeh import *
37 model = pickle.load(open('finalized_model.sav', 'rb'))
38 model = model.named_steps['clf']
39 X = np.array([diabetes, vessels_gt2, thrombust_post_pci, timi_post_pci,
40              abnormal_cholesterol, abnormal_platelets]).reshape(1,6)
41 proba = model.predict_proba(X)
42
43 data = pd.read_csv('probs.csv', sep = ';')
44 fpr1, tpr1, prec1, rec1, spe1, npv1 = gracia_curves(data)
45 PREC, REC, SPE, NPV = point_scores(data, proba[0][1])
46
47 """### PREDICTING PROBABILITY OF RESTENOSIS"""
48
49 print('GRACIA-3 restenosis risk score: %2f' % (proba[0][1]))
50 print('For this score, the model gives the next metrics: ')
51 print('Specificity = %2f' % (SPE))
52 print('Sensitivity = %2f' % (REC))
53 print('Precision = %2f' % (PREC))
54 print('Negative Predictive Value = %2f' % (NPV))
55
56 """### PLOTTING YOUR POINT OVER THE GRACIA-3 MODEL ROC CURVE"""
```

```
57
58 thresholds_pr = [0, 0.18, 0.56, 0.81,1]
59 thresholds_fpr = [0, 0.018, 0.14, 0.466,1]
60 thresholds_npv = [0, 0.53, 0.86, 0.98,1]
61 colors = ["red", "orange", "yellow", "green"]
62
63 output_notebook()
64 fig1 = figure_bokeh('ROC', thresholds=thresholds_fpr, x_label = 'FPR', y_label = 'TPR',
65                   x = fpr1, y = tpr1, x_p = 1-SPE, y_p = REC, colors = colors, legend_pos = 'bottom_right')
66 fig2 = figure_bokeh('P-R', thresholds=thresholds_pr, x_label = 'REC', y_label = 'PREC',
67                   x = rec1, y = prec1, x_p = REC, y_p = PREC, colors = colors, legend_pos = 'top_right')
68 fig3 = figure_bokeh('NPV-SPE', thresholds=thresholds_npv, x_label = 'SPE', y_label = 'NPV',
69                   x = spe1, y = npv1, x_p = SPE, y_p = NPV, colors = colors[:-1], legend_pos = 'bottom_right', range_y = (0.85,1))
70 show(column(fig1, fig2, fig3))
71
```

Branch: master ▾

Find file

Copy path

## Restenosis\_IA\_Gracia3 / calculator / calculator.py

 PedroDorado Add files via upload

bb23e2b 2 days ago

1 contributor

Raw Blame History



174 lines (152 sloc) 7.01 KB

```
1 import pandas as pd
2 import numpy as np
3 import sklearn
4 import pickle
5 from tkinter import *
6 from PIL import ImageTk, Image
7 from funcs_bokeh import *
8 from bokeh.io import export_png
9
10 class MyWindow:
11     def __init__(self, win):
12         self.lb10=Label(win, text='Calculadora riesgo restenosis tras PCI', font='Helvetica 14 bold')
13         self.lb11=Label(win, text='Diabetes', font='Helvetica 10 bold')
14         self.lb12=Label(win, text='Enf. en >=2 vasos', font='Helvetica 10 bold')
15         self.lb13=Label(win, text='Trombo post PCI', font='Helvetica 10 bold')
16         self.lb14=Label(win, text='TIMI post PCI', font='Helvetica 10 bold')
17         self.lb15=Label(win, text='Colesterol tot. anormal', font='Helvetica 10 bold')
18         self.lb16=Label(win, text='Plaquetas anormales', font='Helvetica 10 bold')
19         self.res1=Label(win, text='Score restenosis', font='Helvetica 10 bold')
20         self.SPE=Label(win, text='Especificidad', font='Helvetica 10 bold')
21         self.REC=Label(win, text='Sensibilidad', font='Helvetica 10 bold')
22         self.PRE=Label(win, text='Exhaustividad', font='Helvetica 10 bold')
23         self.NPV=Label(win, text='Valor predictivo negativo', font='Helvetica 10 bold')
24
25         self.num1 = IntVar()
26         self.num2 = IntVar()
27         self.num3 = IntVar()
28         self.num4 = IntVar()
29         self.num5 = IntVar()
30         self.num6 = IntVar()
31
32         self.num1.set(0)
33         self.num2.set(0)
34         self.num3.set(0)
35         self.num4.set(3)
36         self.num1.set(0)
37         self.num1.set(0)
38
39         rad1 = Radiobutton(window,text='Sí', variable=self.num1, value=1)
40         rad2 = Radiobutton(window,text='No', variable=self.num1, value=0)
41         rad1.place(x=250, y=200)
42         rad2.place(x=300, y=200)
43         rad3 = Radiobutton(window,text='Sí', variable=self.num2, value=1)
44         rad4 = Radiobutton(window,text='No', variable=self.num2, value=0)
45         rad3.place(x=250, y=250)
46         rad4.place(x=300, y=250)
47
48
49         rad50 = Radiobutton(window,text='0', variable=self.num3, value=0)
50         rad51 = Radiobutton(window,text='1', variable=self.num3, value=1)
51         rad52 = Radiobutton(window,text='2', variable=self.num3, value=2)
52         rad53 = Radiobutton(window,text='3', variable=self.num3, value=3)
53         rad54 = Radiobutton(window,text='4', variable=self.num3, value=4)
54         rad55 = Radiobutton(window,text='5', variable=self.num3, value=5)
55         rad50.place(x=250, y=300)
56         rad51.place(x=280, y=300)
57         rad52.place(x=310, y=300)
58         rad53.place(x=340, y=300)
```

```

59     rad54.place(x=370, y=300)
60     rad55.place(x=400, y=300)
61
62
63     rad7 = Radiobutton(window, text='1', variable=self.num4, value=1)
64     rad8 = Radiobutton(window, text='2', variable=self.num4, value=2)
65     rad81 = Radiobutton(window, text='3', variable=self.num4, value=3)
66     rad7.place(x=250, y=350)
67     rad8.place(x=300, y=350)
68     rad81.place(x=350, y=350)
69     rad9 = Radiobutton(window, text='Sí', variable=self.num5, value=1)
70     rad10 = Radiobutton(window, text='No', variable=self.num5, value=0)
71     rad9.place(x=250, y=400)
72     rad10.place(x=300, y=400)
73     rad11 = Radiobutton(window, text='Si', variable=self.num6, value=1)
74     rad12 = Radiobutton(window, text='No', variable=self.num6, value=0)
75     rad11.place(x=250, y=450)
76     rad12.place(x=300, y=450)
77     self.resr=Entry()
78     self.lbl10.place(x=55,y=5)
79     self.lbl11.place(x=100, y=200)
80     self.lbl12.place(x=100, y=250)
81     self.lbl13.place(x=100, y=300)
82     self.lbl14.place(x=100, y=350)
83     self.lbl15.place(x=100, y=400)
84     self.lbl16.place(x=100, y=450)
85     self.b1=Button(win, text='Ejecutar modelo', command=self.add, font='Helvetica 14 bold')
86     self.b1.place(x=100, y=500)
87     self.res1.place(x=100, y=550)
88     self.resr.place(x=300, y=550)
89     def add(self):
90         model = pickle.load(open('../models/finalized_model.sav', 'rb'))
91         model = model.named_steps['clf']
92         X = np.array([self.num1.get(), self.num2.get(), self.num3.get(), self.num4.get(), self.num5.get(), self.num6.get()]).reshape(1,6)
93         proba = model.predict_proba(X)
94         data = pd.read_csv('../models/probs.csv', sep = ';')
95         PREC, REC, SPE, NPV = point_scores(data, proba[0][1])
96         self.resr.delete(0, 'end')
97         result=proba[0][1]
98         fpr1, tpr1, prec1, rec1, spe1, npv1 = gracia_curves(data)
99         thresholds_pr = [0, 0.18, 0.56, 0.81,1]
100        thresholds_fpr = [0, 0.018, 0.14, 0.466,1]
101        thresholds_npv = [0, 0.53, 0.86, 0.98,1]
102        colors = ["red", "orange", "yellow", "green"]
103        fig1 = figure_bokeh('ROC', thresholds=thresholds_fpr,
104                            x_label = '1 - Especificidad', y_label = 'Sensitividad',
105                            x = fpr1, y = tpr1, x_p = 1-SPE, y_p = REC, colors = colors, legend_pos = 'bottom_right')
106        fig2 = figure_bokeh('P-R', thresholds=thresholds_pr, x_label = 'Sensitividad',
107                            y_label = 'Exhaustividad', x = rec1, y = prec1,
108                            x_p = REC, y_p = PREC, colors = colors,
109                            legend_pos = 'top_right')
110        fig3 = figure_bokeh('NPV-SPE', thresholds=thresholds_npv,
111                            x_label = 'Especificidad',
112                            y_label = 'Valor Predictivo Negativo',
113                            x = spe1, y = npv1, x_p = SPE, y_p = NPV,
114                            colors = colors[::-1], legend_pos = 'bottom_right',
115                            range_y = (0.85,1))
116        export_png(fig1, filename="fig1.png")
117        export_png(fig2, filename="fig2.png")
118        export_png(fig3, filename="fig3.png")
119
120        load2 = Image.open("fig1.png")
121        render2 = ImageTk.PhotoImage(load2)
122        img2 = Label(window, image=render2)
123        img2.image = render2
124        img2.place(x=550, y=40)
125
126        load3 = Image.open("fig2.png")
127        render3 = ImageTk.PhotoImage(load3)
128        img3 = Label(window, image=render3)
129        img3.image = render3
130        img3.place(x=550, y=350)
131

```

```
132     load4 = Image.open("fig3.png")
133     render4 = ImageTk.PhotoImage(load4)
134     img4 = Label(window, image=render4)
135     img4.image = render4
136     img4.place(x=900, y=350)
137
138     self.SPE.place(x=920, y=100)
139     self.REC.place(x=920, y=150)
140     self.PRE.place(x=920, y=200)
141     self.NPV.place(x=920, y=250)
142
143     self.RSPE=Entry()
144     self.RREC=Entry()
145     self.RPRE=Entry()
146     self.RNPV=Entry()
147
148     self.RSPE.place(x=1100, y=100)
149     self.RREC.place(x=1100, y=150)
150     self.RPRE.place(x=1100, y=200)
151     self.RNPV.place(x=1100, y=250)
152
153     self.RSPE.insert(END, str(SPE))
154     self.RREC.insert(END, str(REC))
155     self.RPRE.insert(END, str(PRE))
156     self.RNPV.insert(END, str(NPV))
157
158     self.resr.insert(END, str(result))
159 window=Tk()
160 mywin=MyWindow(window)
161 window.title('Calculadora riesgo restenosis tras ICP')
162 window.geometry("1300x650+10+10")
163
164 load1 = Image.open("USAL2.jpg")
165 render1 = ImageTk.PhotoImage(load1)
166 img1 = Label(window, image=render1)
167 img1.image = render1
168 img1.place(x=10, y=40)
169
170
171
172
173 window.mainloop()
```

## **Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)**

Se ha realizado una búsqueda de publicaciones en la base de datos PubMed por dos grupos de términos, tomando todas las publicaciones de los últimos 5 años completos y de enero a marzo 2019 (ambos meses inclusive), que contuvieran en el título o *abstract* uno o más términos del grupo 1 y uno o más términos del grupo 2.

- Grupo 1: “machine learning”, “deep learning”, “neural network”, “neural network(s)”, convolutional, “artificial intelligence”.
- Grupo 2: cardiac, cardiology, "heart failure", coronary, ventricle, aortic, atrial, fibrillation, myocardium, pericardium, endocardium, "heart disease", cardiotoxicity, arrhythmia, ischemic.

La instrucción al buscador de PubMed fue:

("machine learning"[Title/Abstract] OR "deep learning"[Title/Abstract] OR "neural network"[Title/Abstract] OR "neural networks"[Title/Abstract] OR convolutional[Title/Abstract] OR "artificial intelligence"[Title/Abstract]) AND (cardiac[Title/Abstract] OR cardiology[Title/Abstract] OR "heart failure"[Title/Abstract] OR coronary[Title/Abstract] OR ventricle[Title/Abstract] OR aortic[Title/Abstract] OR atrial[Title/Abstract] OR fibrillation[Title/Abstract] OR myocardium[Title/Abstract] OR pericardium[Title/Abstract] OR endocardium[Title/Abstract] OR "heart disease"[Title/Abstract] OR cardiotoxicity[Title/Abstract] OR arrhythmia[Title/Abstract]) OR ischemic[Title/Abstract])

De la lista inicial obtenida (1.111 publicaciones) se ha hecho un filtrado manual para excluir las publicaciones de revisión y no estrictamente relacionadas con el área de la cardiología. La lista de publicaciones originales final (673 publicaciones) se clasificó en 5 áreas, en función de la aparición de los siguientes términos en título y *abstract* y la revisión de cada artículo:

- Arritmias: arrhythmia, fibrillation, ECG
- Cardiopatía Isquémica: risk, ischem\*, hemodynam\*, acute



- Insuficiencia cardiaca: insufficiency, failure, transplant
- Imagen cardiaca: image, imaging, MR, MRI, CT, CTA, tomography, echocard\*
- Otros: cardiotoxicity, pericardium, myocardium, endocard\*, disease, surgery, surgical, percutaneous, structural, congenital, hypertension, systolic, perfusion.

**Anexo 5 - Tabla 1. Publicaciones inteligencia artificial en arritmias cardiacas**

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Referencia	Título	Modalidad	Observación
Li et al. (2014) <sup>80</sup>	Ventricular fibrillation and tachycardia classification using a machine learning approach	ML	Identificación de fibrilación ventricular o taquicardia ventricular a partir de datos del ECG
Das et al. (2014) <sup>104</sup>	ECG Beats Classification Using Mixture of Features	DL	Clasificación en 5 tipos de latidos a partir de datos del ECG
Mahajan et al. (2014) <sup>105</sup>	Performance of an automatic arrhythmia classification algorithm: comparison to the ALTITUDE electrophysiologist panel adjudications	ML	Clasificación y adjudicación de episodios arritmicos en portadores de DAI
Caravaca et al. (2014) <sup>106</sup>	Application of machine learning techniques to analyse the effects of physical exercise in ventricular fibrillation	ML, DL	Análisis de la influencia del ejercicio previo durante eventos de fibrilación ventricular en un modelo animal
Daqrouq et al. (2014) <sup>107</sup>	Neural network and wavelet average framing percentage energy for atrial fibrillation classification	DL	Clasificación de fibrilación auricular
Ebrahimzadeh et al. (2014) <sup>108</sup>	A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals	ML, DL	Predictor de muerte súbita a partir de análisis de la variabilidad de frecuencia cardiaca en el ECG
Behar et al. (2014) <sup>109</sup>	A comparison of single channel fetal ECG extraction methods	DL	Estimación de la frecuencia cardiaca fetal a partir de la monitorización de mujeres embarazadas
Li et al. (2014) <sup>110</sup>	A machine learning approach to multi-level ECG signal quality classification	ML	Clasificación de arritmias a partir del ECG
Sunkaria et al. (2014) <sup>111</sup>	An ANN-based HRV classifier for cardiac health prognosis	ML	Clasificador de salud cardiovascular basada en análisis de variabilidad de frecuencia cardiaca en el ECG
Prakosa et al. (2014) <sup>112</sup>	Cardiac electrophysiological activation pattern estimation from images using a patient-specific database of synthetic image sequences	ML	Estimación de los patrones de activación electromecánica a partir del análisis de contractilidad
Beudel et al. (2014) <sup>113</sup>	Small-world characteristics of EEG patterns in post-anoxic encephalopathy	DL	Predicción de encefalopatía anóxica a partir del análisis de electroencefalogramas tras parada cardiaca resucitada
Nguyen et al. (2014) <sup>114</sup>	Neural network approach for non-invasive detection of hyperglycemia using electrocardiographic signals	DL	Clasificación de estados de hiperglucemia o normoglucemia a partir del análisis electrocardiográficos
Wander et al. (2014) <sup>115</sup>	A combined segmenting and non-segmenting approach to signal quality estimation for ambulatory photoplethysmography	ML	Optimización de la medida de frecuencia cardiaca a partir de ondas ópticas de sistemas de monitorización móviles
Juhola et al. (2014) <sup>116</sup>	On computation of calcium cycling anomalies in cardiomyocytes data	ML	Clasificación de señales para detectar picos en las corrientes de Ca <sup>2+</sup> en cardiomiocitos
Kennedy et al. (2015) <sup>117</sup>	Using Time Series Analysis to Predict Cardiac Arrest in a PICU	ML	Modelo de predicción de parada cardiaca en pacientes pediátricos ingresados en unidades de críticos
Kora et al. (2015) <sup>118</sup>	Improved Bat algorithm for the detection of myocardial infarction	ML	Sistema de detección de infarto de miocardio utilizando en análisis de ECG
Naseri et al. (2015) <sup>119</sup>	Electrocardiogram signal quality assessment using an artificially reconstructed target lead	DL	Optimización de las señales electrocardiográficas resultantes del datatón: Computing-in-Cardiology Challenge 2011

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Attin et al. (2015) <sup>120</sup>	Electrocardiogram characteristics prior to in-hospital cardiac arrest	ML	Predictores de parada cardiaca resucitada en el ECG
Zhang et al. (2015) <sup>121</sup>	Automatic recognition of cardiac arrhythmias based on the geometric patterns of Poincaré plots	ML	Clasificador automático de arritmias cardiacas
He et al. (2015) <sup>122</sup>	Combining multiple ECG features does not improve prediction of defibrillation outcome compared to single features in a large population of out-of-hospital cardiac arrests	ML, DL	Predicción de éxito de desfibrilación combinando múltiples variables
Lillo-Castellano et al. (2015) <sup>123</sup>	Symmetrical compression distance for arrhythmia discrimination in cloud-based big-data services	ML	Clasificación de arritmias en la nube a partir de datos de DAI procedentes de la plataforma SCOOP
Tanantong et al. (2015) <sup>124</sup>	False alarm reduction in BSN-based cardiac monitoring using signal quality and activity type information	ML	Reducción de falsas alarmas en la monitorización de las unidades de críticos
Ravens et al. (2015) <sup>125</sup>	Application of the RIMARC algorithm to a large data set of action potentials and clinical parameters for risk prediction of atrial fibrillation	ML	Clasificación preoperatoria del ritmo cardiaco
Jiang et al. (2015) <sup>126</sup>	Noninvasive reconstruction of cardiac transmembrane potentials using a kernelized extreme learning method	ML	Reconstrucción de los potenciales transmembrana cardíacos a partir de los potenciales de superficie corporal
Kiranyaz et al. (2016) <sup>127</sup>	Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks	DL	Sistema rápido y preciso de clasificación y control de ECG
Rosier et al. (2016) <sup>128</sup>	Personalized and automated remote monitoring of atrial fibrillation	ML	Sistema de filtrado de alertas en fibrilación auricular según su importancia clínica
Rad et al. (2016) <sup>129</sup>	Automatic cardiac rhythm interpretation during resuscitation	ML	Sistema de interpretación automática del ritmo cardiaco en la parada cardiaca
Shandilya et al. (2016) <sup>130</sup>	Integration of Attributes from Non-Linear Characterization of Cardiovascular Time-Series for Prediction of Defibrillation Outcomes	ML	Sistema para predecir el éxito de la desfibrilación eléctrica
Yu et al. (2016) <sup>131</sup>	A new method without reference channels used for ventricular fibrillation detection during cardiopulmonary resuscitation	DL	Sistema de detección de fibrilación ventricular que clasifica los artefactos que se producen durante las maniobras de RCP
Mert (2016) <sup>132</sup>	ECG feature extraction based on the bandwidth properties of variational mode decomposition	ML, DL	Sistema de clasificación de arritmias
Elhaj et al. (2016) <sup>133</sup>	Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals	ML, DL	Sistema de clasificación de arritmias
Zahid et al. (2016) <sup>134</sup>	Patient-derived models link re-entrant driver localization in atrial fibrillation to fibrosis spatial pattern	ML	Utilización de modelos auriculares para probar la hipótesis que los modelos de fibrilación auricular por reentrada son propios de regiones con patrones de fibrosis específicos
Pant et al. (2016) <sup>135</sup>	Efficient compressive sensing of ECG segments based on machine learning for QRS-based arrhythmia detection	ML	Telemonitorización de arritmias basado en análisis del QRS

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Ansari et al. (2016) <sup>136</sup>	Suppression of false arrhythmia alarms in the ICU: a machine learning approach	ML	Supresión de falsas alarmas por arritmias en unidades de críticos
Eerikainen et al. (2016) <sup>61</sup>	Reduction of false arrhythmia alarms using signal selection and machine learning	ML	Clasificación de alarmas por arritmias en verdaderas o falsas
Figuera et al. (2016) <sup>81</sup>	Machine Learning Techniques for the Detection of Shockable Rhythms in Automated External Defibrillators	ML	Clasificación de ritmos desfibrilables a partir de ECG y holters para mejorar los desfibriladores externos automáticos
Immanuel et al. (2016) <sup>137</sup>	T-wave morphology can distinguish healthy controls from LQTS patients	DL	Análisis automático de la onda T para clasificar pacientes con síndrome de QT largo frente a controles
Li et al. (2016) <sup>138</sup>	Integrated Machine Learning Approaches for Predicting Ischemic Stroke and Thromboembolism in Atrial Fibrillation	ML	Predicción de ictus cardioembólico y tromboembolismo pulmonar en pacientes con fibrilación auricular
Nemati et al. (2016) <sup>139</sup>	Monitoring and detecting atrial fibrillation using wearable technology	ML	Identificación de fibrilación auricular a partir de dispositivos electrónicos de pulsera
Xiong et al. (2016) <sup>140</sup>	A stacked contractive denoising auto-encoder for ECG signal denoising	DL	Desarrollo de un filtro para evitar artefactos electrocardiográficos
Yanting et al. (2016) <sup>141</sup>	Risk prediction for cardiovascular disease using ECG data in the China kadoorie biobank	ML	Clasificación de patologías cardíacas y predicción de riesgo a partir del ECG
Zhang et al. (2016) <sup>142</sup>	In silico prediction of hERG potassium channel blockage by chemical category approaches	ML	Identificación de alertas de prolongación del QT por bloqueo del gen del canal hERG para la valoración de nuevos fármacos
He et al. (2016) <sup>143</sup>	Combining Amplitude Spectrum Area with Previous Shock Information Using Neural Networks Improves Prediction Performance of Defibrillation Outcome for Subsequent Shocks in Out-Of-Hospital Cardiac Arrest Patients	DL	Predicción del éxito de desfibrilación
Lancaster et al. (2016) <sup>144</sup>	Improved Prediction of Drug-Induced Torsades de Pointes Through Simulations of Dynamics and Machine Learning Algorithms	ML	Predicción de aparición de torsión de puntas en el desarrollo de fármacos
Khan et al. (2016) <sup>145</sup>	A cardiac electrical activity model based on a cellular automata system in comparison with neural network model	ML, DL	Modelo del comportamiento autómatas celular
Alqaraawi et al. 2016 <sup>146</sup>	Heart rate variability estimation in photoplethysmography signals using Bayesian learning approach	ML	Algoritmo para estimar la frecuencia cardíaca por fotoplethysmografía para dispositivos electrónicos de muñeca
Kalidas et al. 2016 <sup>147</sup>	Cardiac arrhythmia classification using multi-modal signal analysis	ML	Clasificador de alarmas para identificar falsas alarmas a partir de monitorización de intensivos resultado del datatón: Physionet/Computing in Cardiology 2015 Challenge
Lee et al. (2016) <sup>148</sup>	Prediction of Ventricular Tachycardia One Hour before Occurrence Using Artificial Neural Networks	DL	Predictor de la aparición de taquicardia ventricular en pacientes monitorizados en unidades de críticos
Abdul-Kadir et al. (2016) <sup>149</sup>	Dynamic ECG features for atrial fibrillation recognition	ML, DL	Clasificador de fibrilación auricular

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Erem et al. (2016) <sup>150</sup>	Extensions to a manifold learning framework for time-series analysis on dynamic manifolds in bioelectric signals	ML	Explorar la estructura geométrica de la integración temporal de ECG y electroencefalogramas
Kurzendorfer et al. (2016) <sup>151</sup>	Cryo-Balloon Catheter Localization Based on a Support-Vector-Machine Approach	ML	Localización automática de catéteres de crioblación
Evangelista et al. (2017) <sup>152</sup>	Predicting adherence to use of remote health monitoring systems in a cohort of patients with chronic heart failure	ML	Identificación de pacientes con insuficiencia cardiaca candidatos a utilizar sistemas de monitorización remota
Giffard-Roisin et al. (2017) <sup>153</sup>	Noninvasive Personalization of a Cardiac Electrophysiology Model From Body Surface Potential Mapping	ML	Modelo predictivo de la respuesta a diferentes condiciones de estimulación
Li et al. (2017) <sup>154</sup>	Genetic algorithm for the optimization of features and neural networks in ECG signals classification	DL	Clasificación de señales a partir del ECG
Mjahad et al. (2017) <sup>155</sup>	Ventricular Fibrillation and Tachycardia detection from surface ECG using time-frequency representation images as input dataset for machine learning	ML, DL	Clasificador automático de arritmias ventriculares utilizando imágenes
Okutucu et al. (2017) <sup>156</sup>	Data mining experiments on the Angiotensin II-Antagonist in Paroxysmal Atrial Fibrillation (ANTIPAF-AFNET 2) trial: 'exposing the invisible'	ML	Predictores clínicos de aparición de fibrilación auricular
Acharya et al. (2017) <sup>157</sup>	A deep convolutional neural network model to classify heartbeats	DL	Clasificación de arritmias basado en la identificación de diferentes latidos cardiacos
Arvanaghi et al. (2017) <sup>158</sup>	Fusion of ECG and ABP signals based on wavelet transform for cardiac arrhythmias classification	DL	Clasificación de hasta 5 tipos de ritmo cardiaco
Tripathy et al. (2017) <sup>159</sup>	Automated detection of heart ailments from 12-lead ECG using complex wavelet sub-band bi-spectrum features	ML	Clasificación de arritmias, infarto, y bloqueo de rama a partir del ECG de 12 derivaciones
Luo et al. (2017) <sup>160</sup>	Patient-Specific Deep Architectural Model for ECG Classification	DL	Clasificador automático de la imagen tiempo-frecuencia electrocardiográfica
Li et al. (2017) <sup>161</sup>	High-Performance Personalized Heartbeat Classification Model for Long-Term ECG Signal	DL	Clasificador de arritmias cardiacas a partir de ECG
Pandit et al. (2017) <sup>162</sup>	A lightweight QRS detector for single lead ECG signals using a max-min difference algorithm	DL	Algoritmo para clasificar arritmias basado en una derivación electrocardiográfica
Rad et al. (2017) <sup>163</sup>	ECG-Based Classification of Resuscitation Cardiac Rhythms for Retrospective Data Analysis	ML, DL	Clasificación automática de los ritmos resucitables
Zhou et al. (2017) <sup>164</sup>	Premature ventricular contraction detection combining deep neural networks and rules inference	DL	Detección de extrasístoles ventriculares
Kiranyaz et al. (2017) <sup>165</sup>	Personalized Monitoring and Advance Warning System for Cardiac Arrhythmias	DL	Sistema de clasificación de arritmias cardiacas
Mustaqeem et al. (2017) <sup>166</sup>	Wrapper method for feature selection to classify cardiac arrhythmia	ML	Clasificador de arritmias cardiacas a partir de ECG

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Wang et al. (2017) <sup>167</sup>	An integrated anti-arrhythmic target network of a Chinese medicine compound, Wenxin Keli, revealed by combined machine learning and molecular pathway analysis	ML	Predicción del efecto antiarrítmico del Wenxin Keli (medicina china patentada) sobre el uso de la quinidina en modelo animal
Sadrawi et al. (2017) <sup>168</sup>	Arrhythmia Evaluation in Wearable ECG Devices	DL	Algoritmo clasificador de arritmia en dispositivos electrónicos comparado con otros algoritmos previos
Lemkaddem et al. (2017) <sup>169</sup>	An autonomous medical monitoring system: Validation on arrhythmia detection	ML	Plataforma para clasificación de arritmias
Liu et al. (2017) <sup>170</sup>	Precision Cohort Finding with Outcome-Driven Similarity Analytics: A Case Study of Patients with Atrial Fibrillation	ML	Identificación de pacientes en fibrilación auricular con bajo riesgo de ictus cardioembólico
Sharifi et al. (2017) <sup>171</sup>	Development of models for predicting Torsade de Pointes cardiac arrhythmias using perceptron neural networks	DL	Desarrollo de un modelo estructura-actividad que se pueda utilizar como una prueba temprana para detectar el potencial proarrítmico (torsión de puntas) en candidatos a fármacos
Sahoo et al. (2017) <sup>172</sup>	ECG beat classification using empirical mode decomposition and mixture of features	DL	Clasificador de seis tipos de arritmias
Lenning et al. (2017) <sup>173</sup>	Real-Time Monitoring and Analysis of Zebrafish Electrocardiogram with Anomaly Detection	ML	Estudio monitorización electrocardiográfica en pez cebra con inteligencia artificial
Clifford et al. (2017) <sup>174</sup>	AF Classification from a Short Single Lead ECG Recording: the PhysioNet/Computing in Cardiology Challenge 2017	ML, DL	Identificación de ritmo sinusal frente a fibrilación auricular o artefactos en el datatón: PhysioNet/Computing in Cardiology Challenge 2017
Yu et al. (2017) <sup>175</sup>	[Research on malignant arrhythmia detection algorithm using neural network optimized by genetic algorithm]	DL	Algoritmos para detección automática de arritmias ventriculares que se puedan aplicar a desfibriladores
Ferrer-Albero et al. (2017) <sup>176</sup>	Non-invasive localization of atrial ectopic beats by using simulated body surface P-wave integral maps	ML	Agrupación espacial y clasificación de focos auriculares ectópicos mediante el uso del mapa integral de la onda P de la superficie corporal
Li et al. (2017) <sup>177</sup>	Modeling of the hERG K <sup>+</sup> Channel Blockage Using Online Chemical Database and Modeling Environment (OCHEM)	ML	Modelo predictivo de prolongación de QT a partir del efecto de otros fármacos publicados en la literatura en relación con la inhibición de los canales de potasio hERG
Yildirim et al. (2018) <sup>63</sup>	Arrhythmia detection using deep convolutional neural network with long duration ECG signals	DL	Clasificación de arritmias cardiacas basada en el análisis de señales electrocardiográficas de larga duración
Marzec et al. (2018) <sup>178</sup>	Device-measured physical activity data for classification of patients with ventricular arrhythmia events: A pilot investigation	ML	Predicción de aparición de taquicardia ventricular a partir de sujetos con holter implantable
Iqbal et al. (2018) <sup>179</sup>	Deep Deterministic Learning for Pattern Recognition of Different Cardiac Diseases through the Internet of Medical Things	DL	Clasificación de fibrilación auricular e infarto de miocardio combinando actividades cardiacas predefinidas con conjunto de datos fusionados

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Yao et al. (2018) <sup>180</sup>	Arrhythmia Classification from Single Lead ECG by Multi-Scale Convolutional Neural Networks	DL	Clasificador de arritmias a partir de una única derivación electrocardiográfica
Nguyen et al (2018) <sup>181</sup>	Deep Feature Learning for Sudden Cardiac Arrest Detection in Automated External Defibrillators	DL, ML	Algoritmo para la detección de TV/FV a partir de la señal electrocardiográfica
Liu et al. (2018) <sup>182</sup>	Classification of Heart Diseases Based On ECG Signals Using Long Short-Term Memory	DL	Clasificación de patologías cardíacas a partir del ECG utilizando secuencias temporales
Yu et al. (2018) <sup>183</sup>	QRS Detection and Measurement Method of ECG Paper Based on Convolutional Neural Networks	DL	Detección automática del QRS a partir de imágenes
Malik et al. (2018) <sup>184</sup>	A machine learning approach to reconstruction of heart surface potentials from body surface potentials	DL	Reconstrucción de potenciales cardíacos de superficie a partir de potenciales de superficie corporales
He et al. (2018) <sup>185</sup>	LiteNet: Lightweight Neural Network for Detecting Arrhythmias at Resource-Constrained Mobile Devices	DL	Sistema de detección de arritmias para dispositivos móviles
Lee et al. (2018) <sup>186</sup>	Heart Rate Estimated from Body Movements at Six Degrees of Freedom by Convolutional Neural Networks	DL	Estimación de la frecuencia cardíaca a partir de movimiento torácico respiratorio
Yildirim (2018) <sup>187</sup>	A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification	DL	Clasificación de ritmo sinusal, extrasistolia ventricular, latido estimulado, bloqueo de rama izquierda y rama derecha.
Sengupta et al. (2018) <sup>188</sup>	Prediction of Abnormal Myocardial Relaxation From Signal Processed Surface ECG	ML	Predicción de la relajación anormal miocárdica a partir de la señal de superficie electrocardiográfica
Mustaqeem et al. (2018) <sup>189</sup>	Multiclass Classification of Cardiac Arrhythmia Using Improved Feature Selection and SVM Invariants	ML	Clasificador en 16 tipos de arritmias cardíacas a partir de análisis electrocardiográficos
Sadr et al. (2018) <sup>190</sup>	A low-complexity algorithm for detection of atrial fibrillation using an ECG	DL	Modelo de predicción de fibrilación auricular basado en el análisis de intervalos RR
Mathews et al. (2018) <sup>191</sup>	A novel application of deep learning for single-lead ECG classification	DL	Clasificación de ECG de una sola derivación
Godoy et al. (2018) <sup>192</sup>	Atrial Fibrosis Hampers Non-invasive Localization of Atrial Ectopic Foci From Multi-Electrode Signals: A 3D Simulation Study	ML	Analizar el efecto de la fibrosis auricular sobre los sistemas de mapeo de superficie en la predicción de focos ectópicos auriculares
Shaik et al. (2018) <sup>193</sup>	Protein phenotype diagnosis of autosomal dominant calmodulin mutations causing irregular heart rhythms	ML	Algoritmo para diagnosticar cambios en las propiedades fenotípicas, de estabilidad y de unión del canal Ca <sup>2+</sup> de las mutaciones causantes de síndromes arrítmicos
Dobkowska-Chudon et al. (2018) <sup>194</sup>	Detecting cerebrovascular changes in the brain caused by hypertension in atrial fibrillation group using acoustocerebrography	ML	Comparación de la señal de acustocerebrografía en pacientes con fibrilación auricular con y sin hipertensión arterial
Parvaneh et al. (2018) <sup>195</sup>	Analyzing single-lead short ECG recordings using dense convolutional neural networks and feature-based post-processing to detect atrial fibrillation	DL	Detección de fibrilación auricular en una sola derivación electrocardiográfica



Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Bollepalli et al. (2018) <sup>196</sup>	Robust Heartbeat Detection from Multimodal Data via CNN-based Generalizable Information Fusion	DL	Fusión de la información de múltiples señales (fundamentalmente ECG)
Warrick et al. (2018) <sup>197</sup>	Ensembling convolutional and long short-term memory networks for electrocardiogram arrhythmia detection	DL	Identificación de fibrilación auricular en el ECG
Faust et al. (2018) <sup>198</sup>	Automated detection of atrial fibrillation using long short-term memory network with RR interval signals	DL	Detección automática de fibrilación auricular utilizando los intervalos RR
Xu et al. (2018) <sup>199</sup>	Atrial Fibrillation Beat Identification Using the Combination of Modified Frequency Slice Wavelet Transform and Convolutional Neural Networks	DL	Identificación automática de fibrilación auricular
Amezquita-Sanchez et al. (2018) <sup>200</sup>	A Novel Wavelet Transform-Homogeneity Model for Sudden Cardiac Death Prediction Using ECG Signals	DL	Predicción de muerte súbita a partir del análisis del ECG
Rubin et al. (2018) <sup>201</sup>	Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings	DL	Identificación de fibrilación auricular a partir de una derivación electrocardiográfica
Bond et al. (2018) <sup>202</sup>	Automation bias in medicine: The influence of automated diagnoses on interpreter accuracy and uncertainty when reading electrocardiograms	ML	Interpretación automática de trazados electrocardiográficos
Attia et al. (2018) <sup>203</sup>	Noninvasive assessment of dofetilide plasma concentration using a deep learning (neural network) analysis of the surface electrocardiogram: A proof of concept study	DL	Evaluación de los cambios morfológicos electrocardiográficos asociados al uso de dofetilide
Budzianowski et al. (2018) <sup>60</sup>	Predictors of atrial fibrillation early recurrence following cryoballoon ablation of pulmonary veins using statistical assessment and machine learning algorithms	ML	Predicción de nueva aparición de fibrilación auricular tras crioablación a partir de datos bioquímicos y clínicos
Kropf et al. (2018) <sup>204</sup>	Cardiac anomaly detection based on time and frequency domain features using tree-based classifiers	ML	Método combinado de análisis de señales electrocardiográficas durante el datatón: CinC 2017
Zhao et al. (2018) <sup>205</sup>	ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation	DL	Sistema para la autenticación de señales electrocardiográficas
Blecker et al. (2018) <sup>206</sup>	Early Identification of Patients With Acute Decompensated Heart Failure	ML	Identificación de pacientes hospitalizados con insuficiencia cardiaca aguda
Cikes et al. (2018) <sup>207</sup>	Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy	ML	Identificación de pacientes respondedores a resincronización cardiaca a partir de datos clínicos y ecocardiográficos
Heam et al. (2018) <sup>208</sup>	Neural Networks for Prognostication of Patients With Heart Failure	DL	Predicción de la situación evolución clínica funcional de pacientes con insuficiencia cardiaca a partir de datos de la ergoespirometría
Kalscheur et al. (2018) <sup>88</sup>	Machine Learning Algorithm Predicts Cardiac Resynchronization Therapy Outcomes: Lessons From the COMPANION Trial	ML	Identificación de pacientes respondedores a terapia de resincronización cardiaca y predicción de sucesos cardiovasculares

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Sodmann et al. (2018) <sup>209</sup>	A convolutional neural network for ECG annotation as the basis for classification of cardiac rhythms	ML, DL	Mejorar la estimación del tipo de ritmo cardiaco a partir del ECG
Xu et al. (2018) <sup>210</sup>	Towards End-to-End ECG Classification with Raw Signal Extraction and Deep Neural Networks	DL	Clasificación de arritmias cardiacas utilizando datos brutos electrocardiográficos
Yang et al. (2018) <sup>211</sup>	Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine	ML, DL	Combinación de ambos métodos de inteligencia artificial para filtrar artefactos y clasificar arritmias automáticamente desde el ECG
Fan et al. (2018) <sup>212</sup>	Multiscaled Fusion of Deep Convolutional Neural Networks for Screening Atrial Fibrillation From Single Lead Short ECG Recordings	DL	Identificación de fibrilación auricular a partir de una derivación única electrocardiográfica
Plesinger et al. (2018) <sup>213</sup>	Parallel use of a convolutional neural network and bagged tree ensemble for the classification of Holter ECG	ML, DL	Detección automática de arritmias a partir de datos de holter tras el datatón PhysioNet/CinC Challenge 2017
Das et al. (2018) <sup>214</sup>	Unsupervised heart-rate estimation in wearables with Liquid states and a probabilistic readout	DL	Estimación de la frecuencia cardiaca en dispositivos electrónicos
Xiang et al. (2018) <sup>215</sup>	Automatic QRS complex detection using two-level convolutional neural network	DL	Detección automática de complejos QRS
Tison et al. (2018) <sup>216</sup>	Passive Detection of Atrial Fibrillation Using a Commercially Available Smartwatch	DL	Detección de fibrilación auricular a partir de datos de dispositivos de muñeca
Lyon et al. (2018) <sup>82</sup>	Distinct ECG Phenotypes Identified in Hypertrophic Cardiomyopathy Using Machine Learning Associate With Arrhythmic Risk Markers	ML	Identificación de distintos fenotipos de miocardiopatía hipertrófica familiar basados en el análisis del ECG y caracterizar las diferencias clínicas y anatómicas con Rm
Gliner et al. (2018) <sup>217</sup>	An SVM approach for identifying atrial fibrillation	ML	Algoritmo para identificar fibrilación auricular basado en el pico de la onda R
Faganel et al. (2018) <sup>218</sup>	A topological approach to delineation and arrhythmic beats detection in unprocessed long-term ECG signals	ML	Algoritmo basado en principios topológicos (teoría de Morse) para detectar latidos arrítmicos en el ECG
Clerx et al. (2018) <sup>219</sup>	Predicting changes to INa from missense mutations in human SCN5A	ML	Estudio de simulación mecanicista con resultados no consistentes en la predicción de cambios en el canal de Na secundarios a mutaciones del gen SCN5A
Lawson et al. (2018) <sup>220</sup>	Slow Recovery of Excitability Increases Ventricular Fibrillation Risk as Identified by Emulation	ML	Simulación de escenarios electrofisiológicos para investigar el papel de la variabilidad en la densidad y cinética de los canales iónicos en el comportamiento arrítmico impulsado por rotor
He et al. (2018) <sup>221</sup>	Automatic Detection of Atrial Fibrillation Based on Continuous Wavelet Transform and 2D Convolutional Neural Networks	DL	Sistema automático de detección de fibrilación auricular basado en características tiempo-frecuencia del ECG
Ebrahimzadeh et al. (2018) <sup>59</sup>	Prediction of paroxysmal Atrial Fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on HRV signal	ML	Identificación de fibrilación auricular de datos del ECG utilizando una combinación de algoritmos

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Isasi et al. (2018) <sup>222</sup>	A Machine Learning Shock Decision Algorithm for use during Piston-driven Chest Compressions	ML	Algoritmo para mejorar la RCP con decisiones sobre ritmos a desfibrilar y compresiones en el LUCAS2
Xie et al. (2018) <sup>223</sup>	Bidirectional Recurrent Neural Network And Convolutional Neural Network (BiRCNN) For ECG Beat Classification	DL	Clasificación automática de latidos cardiacos desde el ECG
Miyatani et al. (2018) <sup>224</sup>	Denosing Autoencoder-based Modification of RRI data with Premature Ventricular Contraction for Precise Heart Rate Variability Analysis	DL	Mejora en la estimación de la variabilidad de frecuencia cardiaca evitando contabilizar latidos prematuros
Khamis et al. (2018) <sup>225</sup>	Detection of Atrial Fibrillation from RR Intervals and PQRST Morphology using a Neural Network Ensemble	DL	Algoritmo para detectar fibrilación auricular establecido en el datatón: PhysioNet Computing in Cardiology Challenge 2017
Zhou et al. (2018) <sup>226</sup>	Premature Ventricular Contraction Detection from Ambulatory ECG Using Recurrent Neural Networks	DL	Detección de extrasistolia ventricular
Hernandez et al. (2018) <sup>227</sup>	Atrial Fibrillation Detection in Short Single Lead ECG Recordings Using Wavelet Transform and Artificial Neural Networks	DL	Detección automática de fibrilación auricular
Anwar et al. (2018) <sup>228</sup>	Arrhythmia Classification of ECG Signals Using Hybrid Features	ML	Clasificación automática de arritmias cardiacas
Rizwan et al. (2018) <sup>229</sup>	AF detection from ECG recordings using feature selection, sparse coding, and ensemble learning	ML	Clasificación automática de fibrilación auricular resultante del datatón: 2017 PhysioNet Challenge
Yang et al. (2018) <sup>230</sup>	Localization of Origins of Premature Ventricular Contraction by Means of Convolutional Neural Network From 12-Lead ECG	DL	Identificación de extrasistolia ventricular a partir de ECG de 12 derivaciones
Xia et al. (2018) <sup>231</sup>	Detecting atrial fibrillation by deep convolutional neural networks	DL	Detección de fibrilación auricular
Kamaleswaran et al. (2018) <sup>232</sup>	A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using single lead electrocardiograms of variable length	DL	Clasificador de arritmias a partir de una derivación electrocardiográfica
Deng et al. (2018) <sup>233</sup>	Extracting cardiac dynamics within ECG signal for human identification and cardiovascular diseases classification	DL	Sistema dinámico de clasificación de patologías cardiacas a partir del ECG
Zhong et al. (2018) <sup>234</sup>	A deep learning approach for fetal QRS complex detection	ML, DL	Sistema de detección de complejos QRS fetales
Hajimolahoseini et al. (2018) <sup>235</sup>	Inflection point analysis: A machine learning approach for extraction of IEGM active intervals during atrial fibrillation	ML	Algoritmo de detección de fibrilación auricular mediante el análisis de intervalos intracardiacos
Xiong et al. (2018) <sup>236</sup>	ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network	DL	Sistema para clasificar arritmia desarrollado a partir del datatón: 2017 PhysioNet/Computing in Cardiology (CinC) Challenge
Mena et al. (2018) <sup>237</sup>	Mobile Personal Health Monitoring for Automated Classification of Electrocardiogram Signals in Elderly	ML	Sistema de monitorización móvil del ECG diseñado para personas mayores

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Teijeiro et al. 2018 <sup>238</sup>	Abductive reasoning as a basis to reproduce expert criteria in ECG atrial fibrillation identification	ML, DL	Identificación automática de fibrilación auricular a partir del datatón: 2017 PhysioNet/CinC Challenge
Poh et al. (2018) <sup>239</sup>	Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms	DL	Identificación de fibrilación auricular a partir de ondas de pulso en pletismografía
Oh et al. (2018) <sup>240</sup>	Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats	DL	Detección de patrones electrocardiográficos diferentes con RR variables
Kumar et al. (2018) <sup>241</sup>	Heart rate monitoring and therapeutic devices: A wavelet transform based approach for the modeling and classification of congestive heart failure	ML	Monitorización de la frecuencia cardíaca a partir de la detección del QRS del electrocardiograma
Savalia et al. (2018) <sup>242</sup>	Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks	DL	Identificación de diferentes arritmias a partir del ECG
McGillivray et al. (2018) <sup>243</sup>	Machine learning methods for locating re-entrant drivers from electrograms in a model of atrial fibrillation	ML	Modelo celular para ubicar los rotorec en fibrilación auricular de forma automática usando mediciones indirectas de ECG
G et al. (2018) <sup>244</sup>	Classification of ECG beats using deep belief network and active learning	DL	Modelo para identificar arritmias a partir de QRS aislados del ECG
Ma et al. (2018) <sup>245</sup>	A novel real-time computational framework for detecting catheters and rigid guidewires in cardiac catheterization procedures	ML	Sistema para distinguir catéteres de guía de otros artefactos similares en la radioscopia durante procedimientos de electrofisiología
Zhu et al. (2018) <sup>246</sup>	A lightweight piecewise linear synthesis method for standard 12-lead ECG signals based on adaptive region segmentation	DL	Sistema para reconstrucción en un ECG de 12 derivaciones a partir de 3 derivaciones (I, II, V <sub>2</sub> )
Wu et al. (2018) <sup>247</sup>	A Comparison of 1-D and 2-D Deep Convolutional Neural Networks in ECG Classification	DL	Clasificación del electrocardiograma entre normal y anormal
Nagaraj et al. (2018) <sup>248</sup>	The revised Cerebral Recovery Index improves predictions of neurological outcome after cardiac arrest	ML	Predicción del daño por encefalopatía anóxica tras la muerte súbita recuperada a partir del análisis de electroencefalogramas
Nanayakkara et al. (2018) <sup>62</sup>	Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study	ML	Predicción de mortalidad en pacientes ingresados con muerte súbita resucitada
Akca et al. (2018) <sup>249</sup>	Is There Any Difference in Risk Factors between Male and Female Patients in New-Onset Atrial Fibrillation after Coronary Artery Bypass Grafting?	ML	Algoritmo de predicción de entrada en fibrilación auricular tras cirugía de pontaje aorto coronario
Kwon et al. (2018) <sup>250</sup>	An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest	DL	Predicción de parada cardíaca intrahospitalaria
Sbrollini et al. (2019) <sup>251</sup>	Serial electrocardiography to detect newly emerging or aggravating cardiac pathology: a deep-learning approach	DL	Detección de patologías cardíacas en fase temprana o avanzada a partir de electrocardiografía seriada

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Erdenebayar et al. (2019) <sup>252</sup>	Automatic Prediction of Atrial Fibrillation Based on Convolutional Neural Network Using a Short-term Normal Electrocardiogram Signal	DL	Predicción automática de fibrilación auricular basada en redes neuronales convolucionales aplicadas a electrocardiogramas de corto plazo
Biswas et al. (2019) <sup>253</sup>	CorNET: Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment	DL	Estimación del pulso cardíaco basado en PPG a través de deep learning con dispositivos <i>wireless</i> en escenarios ambulatorios
Hannun et al. (2019) <sup>254</sup>	Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network	DL	Detección y clasificación de arritmias en escenarios ambulatorios
Bhattacharya et al. (2019) <sup>255</sup>	Identifying Ventricular Arrhythmias and Their Predictors by Applying Machine Learning Methods to Electronic Health Records in Patients With Hypertrophic Cardiomyopathy (HCM-VAR-Risk Model)	ML	Identificación y predicción de arritmias ventriculares en pacientes de cardiomiopatía hipertrófica basada en registros clínicos electrónicos
Attia et al. (2019) <sup>256</sup>	Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram	DL	Detección de disfunción de la contractilidad cardíaca a través de deep learning aplicado a señal electrocardiográfica
Gradolewski et al. (2019) <sup>257</sup>	A Wavelet Transform-Based Neural Network Denoising Algorithm for Mobile Phonocardiography	DL	Eliminación de ruido en señales fonocardiográficas
Karaoğuz et al. (2019) <sup>258</sup>	The quality of ECG data acquisition, and diagnostic performance of a novel adhesive patch for ambulatory cardiac rhythm monitoring in arrhythmia detection	ML	Inteligencia artificial para la evaluación de dispositivos Holter
Lozoya et al. (2019) <sup>259</sup>	Model-Based Feature Augmentation for Cardiac Ablation Target Learning From Images	DL	Modelo para la generación automática de datos para la mejora de algoritmos deep learning de clasificación de pacientes candidatos a ablación cardíaca por radiofrecuencia
Oliveira et al. (2019) <sup>260</sup>	Geometrical features for premature ventricular contraction recognition with analytic hierarchy process based machine learning algorithms selection	ML	Machine learning para la detección de características geométricas de contracción ventricular prematura
Coult et al. (2019) <sup>261</sup>	Ventricular Fibrillation Waveform Analysis During Chest Compressions to Predict Survival From Cardiac Arrest	ML	Análisis de formas de onda de fibrilación ventricular durante compresiones torácicas para la predicción de supervivencia tras parada cardíaca
Gordon et al. (2019) <sup>262</sup>	PVC Detection Using a Convolutional Autoencoder and Random Forest Classifier	DL	Detección de contracciones ventriculares prematuras con redes neuronales convolucionales
Sutton et al. (2019) <sup>263</sup>	PhysOnline: An Open Source Machine Learning Pipeline for Real-Time Analysis of Streaming Physiological Waveform	ML	PhysOnline: plataforma de machine learning código abierto para el análisis en tiempo real de electrocardiogramas
Cazeau et al. (2019) <sup>264</sup>	Statistical ranking of electromechanical dyssynchrony parameters for CRT	ML	Clasificación estadística de parámetros de desincronización electromecánica para aplicación de terapias de resincronización cardíaca

*Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)*

Alqudah et al. (2019) <sup>265</sup>	Developing of robust and high accurate ECG beat classification by combining Gaussian mixtures and wavelets features	ML	Clasificación de señales electrocardiográficas
Xie et al. (2019) <sup>266</sup>	A compound heterozygosity of Tecrl gene confirmed in a catecholaminergic polymorphic ventricular tachycardia family	ML	Identificación de genes responsables de taquicardia ventricular

**Anexo 5 - Tabla 2. Publicaciones inteligencia artificial en cardiopatía isquémica**

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Referencia	Título	Modalidad	Observación
Liu et al. (2014) <sup>267</sup>	Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection	ML	Predicción de sucesos adversos en unidades de dolor torácico
Kumar et al. (2014) <sup>268</sup>	Classification of atherosclerotic and non-atherosclerotic individuals using multiclass state vector machine	ML	Clasificación entre individuos con o sin enfermedad coronaria
Alshurafa et al. (2014) <sup>269</sup>	Remote health monitoring: predicting outcome success based on contextual features for cardiovascular disease	ML	Determinación del riesgo cardiovascular a partir del análisis de factores de riesgo y otros parámetros clínicos
Zellweger et al. (2014) <sup>270</sup>	A new memetic pattern based algorithm to diagnose/exclude coronary artery disease	ML	Algoritmo para excluir enfermedad coronaria
Guidi et al. (2014) <sup>271</sup>	A machine learning system to improve heart failure patient assistance	ML, DL	Clasificación de gravedad de insuficiencia cardíaca
Liu et al. (2014) <sup>272</sup>	Risk stratification for prediction of adverse coronary events in emergency department chest pain patients with a machine learning score compared with the TIMI score	ML	Predicción de síndrome coronario agudo en las unidades de dolor torácico
Nakajima et al. (2015) <sup>273</sup>	Diagnostic Performance of Artificial Neural Network for Detecting Ischemia in Myocardial Perfusion Imaging	DL	Diagnóstico automático de cardiopatía isquémica
Chen et al. (2015) <sup>274</sup>	Using Blood Indexes to Predict Overweight Statuses: An Extreme Learning Machine-Based Approach	ML	Este estudio explora una nueva técnica que utiliza mediciones de sangre y bioquímicas para reconocer la condición de sobrepeso; y por ende riesgo cardiovascular
Panahiazar et al. (2015) <sup>275</sup>	Using EHRs for Heart Failure Therapy Recommendation Using Multidimensional Patient Similarity Analytics	ML	Individualización del tratamiento médico en insuficiencia cardíaca
Panahiazar et al. (2015) <sup>276</sup>	Using EHRs and Machine Learning for Heart Failure Survival Analysis	ML	Predicción de mortalidad en pacientes con insuficiencia cardíaca
Torii et al. (2015) <sup>277</sup>	Risk factor detection for heart disease by applying text analytics in electronic medical records	NLP	Detección de patología cardiovascular identificando factores de riesgo cardiovascular en el análisis de textos
Arsanjani et al. (2015) <sup>278</sup>	Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population	ML	Predicción de revascularización coronaria a partir de datos clínicos y de SPECT
Drenos et al. (2015) <sup>279</sup>	Networks in Coronary Heart Disease Genetics As a Step towards Systems Epidemiology	ML	Predicción de enfermedad coronaria a partir de múltiples datos que incluyen polimorfismos genéticos
Roberts et al. (2015) <sup>280</sup>	The role of fine-grained annotations in supervised recognition of risk factors for heart disease from EHRs	NLP	Identificación automática de factores de riesgo cardiovascular a partir de textos
Poddar et al. (2015) <sup>281</sup>	Automated diagnosis of coronary artery diseased patients by heart rate variability analysis using linear and non-linear methods	ML	Diagnóstico automático de enfermedad coronaria a partir del análisis de la variabilidad de frecuencia cardíaca



Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Sladojevic et al. (2015) <sup>282</sup>	Data Mining Approach for in-Hospital Treatment Outcome in Patients with Acute Coronary Syndrome	ML	Predicción de riesgo en el síndrome coronario agudo sometido a revascularización percutánea
Yang et al. (2015) <sup>283</sup>	A hybrid model for automatic identification of risk factors for heart disease	NLP	Identificación automática de factores de riesgo cardiovascular a partir de textos consecuencia del datatón i2b2/UTHealth NLP Challenge
Jonnagaddala et al. (2015) <sup>284</sup>	Identification and Progression of Heart Disease Risk Factors in Diabetic Patients from Longitudinal Electronic Health Records	NLP	Identificación automática de factores de riesgo cardiovascular a partir de textos
Chen et al. (2015) <sup>285</sup>	An automatic system to identify heart disease risk factors in clinical texts over time	NLP	Identificación automática de factores de riesgo cardiovascular a partir de textos
Isma'eel et al. (2016) <sup>286</sup>	Artificial neural network modeling enhances risk stratification and can reduce downstream testing for patients with suspected acute coronary syndromes, negative cardiac biomarkers, and normal ECGs	DL	Uso de redes neuronales artificiales para mejorar la estratificación de riesgo y la predicción de isquemia miocárdica
Tylman et al. (2016) <sup>287</sup>	Real-time prediction of acute cardiovascular events using hardware-implemented Bayesian networks	ML	Sistema de apoyo en la estimación del estado general de un paciente y riesgo de eventos cardiovasculares
Churpek et al. (2016) <sup>288</sup>	Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards	ML	Algoritmo de detección para identificar pacientes críticamente enfermos en plantas de hospitalización
Wiharto et al. (2016) <sup>289</sup>	Intelligence System for Diagnosis Level of Coronary Heart Disease with K-Star Algorithm	ML	Sistema para predecir enfermedad coronaria
Hoyt et al. (2016) <sup>290</sup>	Digital Family History Data Mining with Neural Networks: A Pilot Study	DL	Sistema basado en redes neuronales para predecir riesgo de padecer diabetes, hipertensión o enfermedad coronaria
Araki et al. (2016) <sup>291</sup>	A new method for IVUS-based coronary artery disease risk stratification: A link between coronary & carotid ultrasound plaque burdens	ML	Evaluación de riesgo coronario mediante la combinación de la escala de grises de placas ateroscleróticas por IVUS y gorsor íntima-media carotídea
Araki et al. (2016) <sup>292</sup>	PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: A link between carotid and coronary grayscale plaque morphology	ML	Sistema de evaluación de riesgo utilizando el análisis de morfología de placas de ateroma
Celutkienė et al. (2016) <sup>293</sup>	Combination of single quantitative parameters into multiparametric model for ischemia detection is not superior to visual assessment during dobutamine stress echocardiography	ML	Modelo para detectar isquemia miocárdica durante la ecocardiografía de estrés
Motwani et al. (2016) <sup>294</sup>	Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis	ML	Sistema de predicción de mortalidad en pacientes con sospecha de cardiopatía isquémica sometidos a TC coronario

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Afarideh et al. (2016) <sup>295</sup>	Complex association of serum alanine aminotransferase with the risk of future cardiovascular disease in type 2 diabetes	DL	Evaluación de la asociación prospectiva entre los niveles séricos ALT y enfermedad cardiovascular en diabéticos tipo 2 usando redes neuronales
Liu et al. (2016) <sup>296</sup>	Beatquency domain and machine learning improve prediction of cardiovascular death after acute coronary syndrome	ML	Evaluación de la frecuencia y dominios de frecuencia en la predicción de riesgo tras un síndrome coronario agudo
Ross et al. (2016) <sup>297</sup>	The use of machine learning for the identification of peripheral artery disease and future mortality risk	ML	Algoritmo para predecir mortalidad a partir de pacientes sometidos a angiografía coronaria
Lopez-de-Andres et al. (2016) <sup>298</sup>	Predictors of in-hospital mortality following major lower extremity amputations in type 2 diabetic patients using artificial neural networks	DL	Predicción de mortalidad en pacientes diabéticos sometidos amputados
Hu et al. (2016) <sup>299</sup>	Utilizing Chinese Admission Records for MACE Prediction of Acute Coronary Syndrome	ML	Predicción de sucesos adversos cardiovasculares tras síndrome coronario agudo
Narain et al. (2016) <sup>300</sup>	Cardiovascular risk prediction: a comparative study of Framingham and quantum neural network based approach	DL	Predicción del riesgo cardiovascular comparado con el score clásico de Framingham
Spertus et al. (2016) <sup>301</sup>	Assessing Hospital Performance After Percutaneous Coronary Intervention Using Big Data	ML	Evaluación mortalidad en pacientes sometidos a revascularización percutánea
Mirian et al. (2016) <sup>302</sup>	A Hybrid ANN-GA Model to Prediction of Bivariate Binary Responses: Application to Joint Prediction of Occurrence of Heart Block and Death in Patients with Myocardial Infarction	DL	Predicción de muerte o bloqueo AV en pacientes con infarto de miocardio
Orphanou et al. (2016) <sup>303</sup>	DBN-Extended: A Dynamic Bayesian Network Model Extended With Temporal Abstractions for Coronary Heart Disease Prognosis	DL	Predicción de enfermedad coronaria
Berikol et al. (2016) <sup>304</sup>	Diagnosis of Acute Coronary Syndrome with a Support Vector Machine	ML	Clasificación de síndrome coronario agudo en unidades de dolor torácico en urgencias
Verma et al. (2016) <sup>305</sup>	A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data	ML	Predicción de enfermedad coronaria a partir de datos clínicos
McRae et al. (2016) <sup>306</sup>	Programmable Bio-nanochip Platform: A Point-of-Care Biosensor System with the Capacity To Learn	ML	“Point of care” nanochip en combinación con algoritmos computacionales para el diagnóstico de infarto agudo de miocardio y otras enfermedades
Karimi et al. (2017) <sup>307</sup>	A Combination of Constitutive Damage Model and Artificial Neural Networks to Characterize the Mechanical Properties of the Healthy and Atherosclerotic Human Coronary Arteries	DL	Predicción de remodelado de las fibras de colágeno de las arterias coronarias
Schafer et al. (2017) <sup>308</sup>	Titin-truncating variants affect heart function in disease cohorts and the general population	ML	Asociación de variantes del gen de titina con variables clínicas y de imagen en el desarrollo de miocardiopatía dilatada en humanos y modelos animales

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Arabasadi et al. (2017) <sup>309</sup>	Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm	DL	Modelo para diagnosticar enfermedad coronaria
Podda et al. (2017) <sup>310</sup>	Prediction of high on-treatment platelet reactivity in clopidogrel-treated patients with acute coronary syndromes	ML, DL	Predicción de reactividad plaquetaria en pacientes que reciben clopidogrel tras síndrome coronario agudo
Alghamdi et al. (2017) <sup>311</sup>	Predicting diabetes mellitus using SMOTE and ensemble machine learning approach: The Henry Ford Exercise Testing (FIT) project	ML	Predicción de diabetes en pacientes sin enfermedad coronaria o insuficiencia cardíaca conocidas sometidos a prueba de esfuerzo
Gao et al. (2017) <sup>312</sup>	Changes and classification in myocardial contractile function in the left ventricle following acute myocardial infarction	ML	Identificación de incremento de la tensión en la pared ventricular izquierda como signo indirecto de aumento en la demanda contráctil de los segmentos adyacentes al infarto
Shah et al. (2017) <sup>313</sup>	Association of Multiorgan Computed Tomographic Phenomap With Adverse Cardiovascular Health Outcomes: The Framingham Heart Study	ML	Riesgo de enfermedad cardíaca (infarto, ictus o muerte cardiovascular) en participantes del estudio Framingham sometidos a TC tóracoabdominal
Oguz et al. (2017) <sup>314</sup>	Genotype-driven identification of a molecular network predictive of advanced coronary calcium in ClinSeq® and Framingham Heart Study cohorts	ML, DL	Predictor de encontrar calcio coronario
Hu et al. (2017) <sup>315</sup>	Acute Coronary Syndrome Risk Prediction Based on GRACE Risk Score	ML	Estimación de riesgo en el síndrome coronario agudo comparado con el score GRACE
Karimi et al. (2017) <sup>316</sup>	A combination of experimental and numerical methods to investigate the role of strain rate on the mechanical properties and collagen fiber orientations of the healthy and atherosclerotic human coronary arteries	DL	Predicción de la alteración del ángulo de las fibras de colágeno en arterias coronarias ateroscleróticas
Lee et al. (2017) <sup>317</sup>	Physiome approach for the analysis of vascular flow reserve in the heart and brain	ML	Algoritmo automático para estimar la reserva de flujo coronario a partir de datos existentes de simulación fisiómica
Weng et al. (2017) <sup>64</sup>	Can machine-learning improve cardiovascular risk prediction using routine clinical data?	ML	Predicción del primer evento cardiovascular a 10 años, estimado el riesgo cardiovascular
Forssen et al. (2017) <sup>318</sup>	Evaluation of Machine Learning Methods to Predict Coronary Artery Disease Using Metabolomic Data	ML	Predicción de enfermedad coronaria a partir de datos metabolómicos
Takeuchi et al. (2017) <sup>319</sup>	Novel Risk Assessment Tool for Immunoglobulin Resistance in Kawasaki Disease: Application Using a Random Forest Classifier	ML	Algoritmo para predecir la resistencia del tratamiento con inmunoglobulina en la enfermedad coronaria de Kawasaki en niños
Buchan et al. (2017) <sup>320</sup>	Automatic prediction of coronary artery disease from clinical narratives	NLP	Predicción de enfermedad coronaria a partir del análisis de texto de historias clínicas
Wallert et al. (2017) <sup>67</sup>	Predicting two-year survival versus non-survival after first myocardial infarction using machine learning and Swedish national register data	ML	Predicción de mortalidad tras infarto agudo de miocardio con datos procedentes del registro SWEDEHEART

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Ambale-Venkatesh et al. (2017) <sup>321</sup>	Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis	ML	Predicción de sucesos cardiovasculares en comparación con scores de riesgo cardiovascular clásicos
Basu et al. (2017) <sup>322</sup>	Development and validation of Risk Equations for Complications Of type 2 Diabetes (RECODE) using individual participant data from randomised trials	ML	Predicción de sucesos cardiovasculares en diabéticos a partir de datos demográficos, clínicos, comorbilidades, medicación y biomarcadores
Shouval et al. (2017) <sup>66</sup>	Machine learning for prediction of 30-day mortality after ST elevation myocardial infraction: An Acute Coronary Syndrome Israeli Survey data mining study	ML	Predicción de mortalidad a 30 días en pacientes con infarto agudo de miocardio con elevación del segmento ST y comparación con scores clásicos
Kim et al. (2017) <sup>323</sup>	Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis	DL	Predicción de riesgo cardiovascular comparado con otros scores clásicos
El-Askary et al. (2017) <sup>324</sup>	Remote sensing observation of annual dust cycles and possible causality of Kawasaki disease outbreaks in Japan	DL	Modelo predictivo de brotes para la enfermedad de Kawasaki con el objetivo final de pronosticar e intervenir de forma temprana utilizando un sistema de alerta temprana
Huang et al. (2017) <sup>65</sup>	MACE prediction of acute coronary syndrome via boosted resampling classification using electronic medical records	ML	Predicción de sucesos cardiovasculares adversos en pacientes con síndrome coronario agudo
Isma'eel et al. (2018) <sup>325</sup>	Artificial neural network-based model enhances risk stratification and reduces non-invasive cardiac stress imaging compared to Diamond-Forrester and Morise risk assessment models: A prospective study	DL	Predicción de enfermedad coronaria en sujetos sometidos a test de isquemia
Tao et al. (2018) <sup>326</sup>	Magnetocardiography based Ischemic Heart Disease Detection and Localization using Machine Learning Methods	ML	Algoritmo de localización automática de isquemia miocárdica a partir del análisis de la onda T
Hae et al. (2018) <sup>327</sup>	Machine learning assessment of myocardial ischemia using angiography: Development and retrospective validation	ML	Sistema para predecir territorios isquémicos
Shi et al. (2018) <sup>328</sup>	Vulnerable Plaque Recognition Based on Attention Model with Deep Convolutional Neural Network	DL	Identificación de placas vulnerables con datos provenientes del datatón 2017 CCCV-IVOCT Challenge
Pieszko et al. (2018) <sup>329</sup>	Machine-learned models using hematological inflammation markers in the prediction of short-term acute coronary syndrome outcomes	ML	Algoritmo de predicción de sucesos adversos en el síndrome coronario agudo utilizando marcadores inflamatorios
van Rosendaal et al. (2018) <sup>330</sup>	Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification; insights from the CONFIRM registry	ML	Predicción del riesgo cardiovascular con la información del árbol coronario de 16 segmentos derivada de TC y comparación con escalas convencionales
Duggal et al. (2018) <sup>331</sup>	Survival outcomes post percutaneous coronary intervention: Why the hype about stent type? Lessons from a healthcare system in India	ML	Algoritmo para predecir sucesos cardiovasculares tras el intervencionismo percutáneo coronario

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Alizadehsani et al. (2018) <sup>332</sup>	Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries	ML	Algoritmo de identificación de enfermedad coronaria que especifica arteria coronaria afectada
Xiao et al. (2018) <sup>333</sup>	A Deep Learning Approach to Examine Ischemic ST Changes in Ambulatory ECG Recordings	DL	Identificador de cambios en el ST para detectar isquemia precoz
Rose et al. (2018) <sup>334</sup>	Double robust estimation for multiple unordered treatments and clustered observations: Evaluating drug-eluting coronary artery stents	ML	Evaluación de stents farmacológicos a través de datos heterogéneos
Dozio et al. (2018) <sup>335</sup>	Plasma fatty acid profile as biomarker of coronary artery disease: a pilot study using fourth generation artificial neural networks	DL	Papel diagnóstico de los ácidos grasos en el diagnóstico de la enfermedad coronaria
Steele et al. (2018) <sup>336</sup>	Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease	ML	Predicción de muerte en la enfermedad coronaria
Zellweger et al. (2018) <sup>337</sup>	A new non-invasive diagnostic tool in coronary artery disease: artificial intelligence as an essential element of predictive, preventive, and personalized medicine	ML	Estimación del riesgo coronario en comparación con el score de Framingham
Body (2018) <sup>338</sup>	Acute coronary syndromes diagnosis, version 2.0: Tomorrow's approach to diagnosing acute coronary syndromes?	ML	Identificación de síndrome coronario agudo en el dolor torácico en urgencias
Yahyaie et al. (2018) <sup>339</sup>	Use of Internet of Things to Provide a New Model for Remote Heart Attack Prediction	DL	Modelo predictivo para sucesos cardiacos adversos siguiendo el concepto "Internet of Things" es decir dispositivos con acceso a internet que permite estratificación inmediata
Brown et al. (2018) <sup>340</sup>	Does Twitter language reliably predict heart disease? A commentary on Eichstaedt et al. (2015a)	ML	Análisis de las expresiones utilizadas en Twitter (categorizadas como buen o mal lenguaje) y su asociación con mortalidad isquémica
McAllister et al. (2018) <sup>341</sup>	Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets	DL	Clasificación por imagen del tipo de comida para control dietético en pacientes diabéticos o con riesgo cardiovascular
Zhang et al. (2018) <sup>342</sup>	Predicting Locations of High-Risk Plaques in Coronary Arteries in Patients Receiving Statin Therapy	ML	Predicción de localización y tipo de placa aterosclerótica coronaria en pacientes de alto riesgo con estatinas
Huang et al. (2018) <sup>343</sup>	A Regularized Deep Learning Approach for Clinical Risk Prediction of Acute Coronary Syndrome Using Electronic Health Records	DL	Estratificación clínica en pacientes con síndrome coronario agudo
Tan et al. (2018) <sup>344</sup>	Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals	DL	Detección automática de signos de isquemia miocárdica a partir de señales electrocardiográficas

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Margulis et al. (2018) <sup>345</sup>	Combining Desorption Electrospray Ionization Mass Spectrometry Imaging and Machine Learning for Molecular Recognition of Myocardial Infarction	ML	Modelo para diferenciar infarto de miocardio por espectrometría de masas en modelo animal murino
Yang et al. (2018) <sup>346</sup>	Multi-dimensional proprio-proximus machine learning for assessment of myocardial infarction	ML	Clasificación de grado de infarto en regiones del ventrículo izquierdo basado en información extraída de resonancia magnética
Kigka et al. (2018) <sup>347</sup>	A Machine Learning Approach for the Prediction of the Progression of Cardiovascular Disease based on Clinical and Non-Invasive Imaging Data	ML	Predicción de enfermedad coronaria a partir de datos clínicos y de imagen
Singh et al. (2018) <sup>348</sup>	Effective heart disease prediction system using data mining techniques	DL	Sistema de predicción de enfermedad coronaria a partir de datos como edad, sexo, presión arterial, colesterol y obesidad.
Haro Alonso et al. (2018) <sup>349</sup>	Prediction of cardiac death after adenosine myocardial perfusion SPECT based on machine learning	ML	Predicción de muerte cardiovascular a partir de datos clínicos y de SPECT
Naushad et al. (2018) <sup>350</sup>	Machine learning algorithm-based risk prediction model of coronary artery disease	ML	Modelo predictivo de enfermedad coronaria y porcentaje de estenosis
Stuckey et al. (2018) <sup>351</sup>	Cardiac Phase Space Tomography: A novel method of assessing coronary artery disease utilizing machine learning	ML	Determinación del rendimiento diagnóstico del análisis del TC en el espacio espacial en pacientes con dolor torácico remitidos para angiografía coronaria
Good et al. (2018) <sup>352</sup>	Temporal Performance of Laplacian Eigenmaps and 3D Conduction Velocity in Detecting Ischemic Stress	ML	Detección de isquemia aguda por el análisis automático electrocardiográfico en modelo animal
Chen et al. (2018) <sup>353</sup>	A Meta-Prediction of Metylenetetrahydrofolate-Reductase Polymorphisms and Air Pollution Increased the Risk of Ischemic Heart Diseases Worldwide	ML	Análisis entre los polimorfismos de la metilenotetrahidrofolato reductasa y el riesgo de enfermedad coronaria
Öman et al. (2019) <sup>354</sup>	3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke	DL	Detección de lesiones isquémicas en angiografías por tomografía computarizada
Heo et al. (2019) <sup>355</sup>	Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke	ML	Predicción de evolución de pacientes infarto agudo basado en modelos machine learning
Vallée et al. (2019) <sup>356</sup>	Coronary heart disease diagnosis by artificial neural networks including aortic pulse wave velocity index and clinical parameters	DL	Clasificación de riesgo de enfermedad coronaria con redes neuronales artificiales a partir de datos clínicos y velocidades de pulso aórtico
Balanescu et al. (2019) <sup>357</sup>	A Cancer Paradox: Machine-Learning Backed Propensity-Score Analysis of Coronary Angiography Findings in Cardio-Oncology	ML	Evaluación del riesgo de enfermedad cardiovascular en pacientes oncológicos
Zabihollahy et al. (2019) <sup>358</sup>	Convolutional neural network-based approach for segmentation of left ventricle myocardial scar from 3D late gadolinium enhancement MR images	DL	Segmentación de lesiones en el miocardio del ventrículo izquierdo en imágenes por resonancia magnética con realce tardío de gadolinio

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Yu et al. (2019) <sup>359</sup>	The best predictor of ischemic coronary stenosis: subtended myocardial volume, machine learning-based FFRCT, or high-risk plaque features?	ML	Comparativa de predictores de estenosis coronaria con isquemia, con FFRCT basada en machine learning, entre ellos
Vernon et al. (2019) <sup>360</sup>	Utilizing state-of-the-art "omics" technology and bioinformatics to identify new biological mechanisms and biomarkers for coronary artery disease	ML	Utilización del estado del arte en tecnologías ómicas, bioinformática y machine learning para la identificación de nuevos mecanismos y biomarcadores de enfermedad coronaria
Chen et al. (2019) <sup>361</sup>	Prediction of cardiovascular outcomes with machine learning techniques: application to the Cardiovascular Outcomes in Renal Atherosclerotic Lesions (CORAL) study	ML	Predicción de diagnósticos. Aplicación al estudio CORAL (Cardiovascular Outcomes in Renal Atherosclerotic Lesions)
Li et al. (2019) <sup>362</sup>	Unhealthy Behaviors, Prevention Measures, and Neighborhood Cardiovascular Health: A Machine Learning Approach	ML	Identificación de predictores de enfermedad coronaria e infarto a partir de datos poblacionales, geográficos y de hábitos saludables
Freiman et al. (2019) <sup>363</sup>	Unsupervised abnormality detection through mixed structure regularization (MSR) in deep sparse autoencoders	DL	Detección automática de lesiones
Goto et al. (2019) <sup>364</sup>	Artificial intelligence to predict needs for urgent revascularization from 12-lead electrocardiography in emergency patients	ML	Inteligencia artificial para la predicción de revascularizaciones urgentes a partir de electrocardiografía de 12 señales en pacientes en unidades de emergencia
Pieszko et al. (2019) <sup>365</sup>	Predicting Long-Term Mortality after Acute Coronary Syndrome Using Machine Learning Techniques and Hematological Markers	ML	Predicción de mortalidad a largo plazo tras síndrome coronario agudo
Cho et al. (2019) <sup>366</sup>	Angiography-Based Machine Learning for Predicting Fractional Flow Reserve in Intermediate Coronary Artery Lesions	ML	Cálculo de la reserva fraccional de flujo en coronarias con lesiones intermedias basado en angiografía
Duan et al. (2019) <sup>367</sup>	Utilizing dynamic treatment information for MACE prediction of acute coronary syndrome	DL	Predicción de síndrome coronario agudo a partir de registros clínicos electrónicos
Bazylev et al. (2019) <sup>368</sup>	[Predictors of early occlusion of coronary bypass grafts in patients with extremely low ejection fraction]	ML	Identificación de predictores de oclusión de bypasses coronarios en pacientes con fracción de eyección severamente deprimida
Adela et al. (2019) <sup>369</sup>	Serum protein signature of coronary artery disease in type 2 diabetes mellitus	ML	Identificación de biomarcadores proteínicos de enfermedad coronaria en pacientes de diabetes tipo 2
Sharma et al. (2019) <sup>370</sup>	Patient Phenotypes, Cardiovascular Risk, and Ezetimibe Treatment in Patients After Acute Coronary Syndromes (from IMPROVE-IT)	ML	Fenotipado, estratificación de riesgo y tratamiento en pacientes de síndrome coronario agudo

**Anexo 5 - Tabla 3. Publicaciones inteligencia artificial en insuficiencia cardiaca**



Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Referencia	Título	Modalidad	Observación
Myers et al. (2014) <sup>371</sup>	A neural network approach to predicting outcomes in heart failure using cardiopulmonary exercise testing	DL	Predicción de mortalidad en pacientes con insuficiencia cardíaca crónica a partir de datos de la ergoespirometría
Schmitz et al. (2014) <sup>372</sup>	Identification of genetic markers for treatment success in heart failure patients: insight from cardiac resynchronization therapy	ML	Identificación de respondedores a la terapia de resincronización a partir de datos clínicos y variantes genéticas
Ooi et al. (2014) <sup>373</sup>	Robust aortic valve non-opening detection for different cardiac conditions	ML	Detección automática del grado de apertura valvular aórtico en la optimización de asistencias ventriculares y recuperación miocárdica
Bondar et al. (2014) <sup>374</sup>	Comparison of whole blood and peripheral blood mononuclear cell gene expression for evaluation of the perioperative inflammatory response in patients with advanced heart failure	ML	Predicción de fallo multiorgánico en insuficiencia cardíaca a partir del análisis de genes compartidos por células mononucleares periféricas y sangre total
Summers et al. (2014) <sup>375</sup>	Functionality of empirical model-based predictive analytics for the early detection of hemodynamic instability	ML	Simulador computacional de inestabilizaciones hemodinámicas
Guidi et al. (2015) <sup>376</sup>	A multi-layer monitoring system for clinical management of Congestive Heart Failure	ML	Sistema de monitorización remota para pacientes con insuficiencia cardíaca
Nilsson et al. (2015) <sup>377</sup>	The International Heart Transplant Survival Algorithm (IHTSA): a new model to improve organ sharing and survival	ML	Predicción de supervivencia en pacientes trasplantados
Meystre et al. (2015) <sup>378</sup>	Heart Failure Medications Detection and Prescription Status Classification in Clinical Narrative Documents	ML	Identificación de tratamiento médico en pacientes con insuficiencia cardíaca
Xu et al. (2015) <sup>379</sup>	Multi-objective optimization of pulsatile ventricular assist device hemocompatibility based on neural networks and a genetic algorithm	DL	Optimización de la hemocompatibilidad de las asistencias ventriculares
Alonso-Betanzos et al. (2015) <sup>380</sup>	Exploring Guidelines for Classification of Major Heart Failure Subtypes by Using Machine Learning	ML	Clasificación del tipo de insuficiencia cardíaca (con FE reducida o preservada) determinando el punto de corte ideal de la FE
Mohammadzadeh et al. (2015) <sup>381</sup>	Chronic Heart Failure Follow-up Management Based on Agent Technology	ML	Monitorización de insuficiencia cardíaca
Liu et al. (2015) <sup>382</sup>	Blood pressure and heart rate from the arterial blood pressure waveform can reliably estimate cardiac output in a conscious sheep model of multiple hemorrhages and resuscitation using computer machine learning approaches	ML	Estimación del gasto cardíaco en un modelo animal
Steventon et al. (2016) <sup>383</sup>	Effect of telehealth on hospital utilisation and mortality in routine clinical practice: a matched control cohort study in an early adopter site	ML	Sistema de monitorización en domicilio en pacientes con insuficiencia cardíaca

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Sideris et al. (2016) <sup>384</sup>	A flexible data-driven comorbidity feature extraction framework	ML	Sistema no supervisado para predecir evolución en una cohorte extensa que incluye pacientes con insuficiencia cardíaca
Hou et al. (2016) <sup>385</sup>	[Investigation into Feasibility of Congestive Heart Failure Diagnosis Based on Analysis of Very Short-term Heart Rate Variability]	DL	Análisis de la variabilidad de la frecuencia cardíaca en el diagnóstico de insuficiencia cardíaca
Blecker et al. (2016) <sup>386</sup>	Comparison of Approaches for Heart Failure Case Identification From Electronic Health Record Data	ML	Desarrollo de un algoritmo que utilizando datos clínicos identifica pacientes con insuficiencia cardíaca
Karanasiou et al. (2016) <sup>387</sup>	Predicting adherence of patients with HF through machine learning techniques	ML	Sistema de predicción de adherencia terapéutica en pacientes con insuficiencia cardíaca
Altan et al. (2016) <sup>388</sup>	A new approach to early diagnosis of congestive heart failure disease by using Hilbert-Huang transform	DL	Clasificación de pacientes con insuficiencia cardíaca o enfermedad coronaria a partir de datos del ECG
Mortazavi et al. (2016) <sup>84</sup>	Analysis of Machine Learning Techniques for Heart Failure Readmissions	ML	Predicción de reingresos hospitalarios en insuficiencia cardíaca
Farhan et al. (2016) <sup>389</sup>	A Predictive Model for Medical Events Based on Contextual Embedding of Temporal Sequences	ML	Modelo predictivo combinando diferentes variables (diagnóstica, prescripciones, test laboratorios) aplicado a la insuficiencia cardíaca
Argerich et al. (2016) <sup>390</sup>	Evaluation of periodic breathing in respiratory flow signal of elderly patients using SVM and linear discriminant analysis	ML	Identificación de insuficiencia cardíaca mediante el análisis de la amplitud del patrón respiratorio
Ng et al. (2016) <sup>391</sup>	Early Detection of Heart Failure Using Electronic Health Records: Practical Implications for Time Before Diagnosis, Data Diversity, Data Quantity, and Data Density	ML	Predicción de insuficiencia cardíaca en atención primaria
Tsai et al. (2016) <sup>392</sup>	Length of Hospital Stay Prediction at the Admission Stage for Cardiology Patients Using Artificial Neural Network	DL	Predicción de la estancia hospitalaria en tres patologías: insuficiencia cardíaca, infarto de miocardio y aterosclerosis coronaria
Masetic et al. (2016) <sup>393</sup>	Congestive heart failure detection using random forest classifier	ML, DL	Diagnóstico de insuficiencia cardíaca a partir de datos electrocardiográficos
Buzaev et al. (2016) <sup>394</sup>	Artificial intelligence: Neural network model as the multidisciplinary team member in clinical decision support to avoid medical mistakes	DL	Predicción de éxito de revascularización (quirúrgico vs. Percutáneo) en pacientes con enfermedad coronaria
Choi et al. (2017) <sup>395</sup>	Using recurrent neural network models for early detection of heart failure onset	DL	Sistema para mejora en el diagnóstico inicial de insuficiencia cardíaca
Frizell et al. (2017) <sup>85</sup>	Prediction of 30-Day All-Cause Readmissions in Patients Hospitalized for Heart Failure: Comparison of Machine Learning and Other Statistical Approaches	ML	Predicción de rehospitalizaciones en insuficiencia cardíaca

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Shameer et al. (2017) <sup>68</sup>	Predictive Modeling of Hospital Readmission Rates Using Electronic Medical Record-Wide Machine Learning: A Case-Study Using Mount Sinai Heart Failure Cohort	ML	Predicción de reingresos hospitalarios en insuficiencia cardiaca
Katz et al. (2017) <sup>396</sup>	Phenomapping for the Identification of Hypertensive Patients with the Myocardial Substrate for Heart Failure with Preserved Ejection Fraction	ML	Algoritmo de identificación de pacientes hipertensos que puedan desarrollar insuficiencia cardiaca con función ventricular conservada
Zhang et al. (2017) <sup>397</sup>	Automatic Methods to Extract New York Heart Association Classification from Clinical Notes	NLP	Identificación de la clase funcional a partir de notas en la historia electrónica
Sanchez-Martinez et al. (2017) <sup>398</sup>	Characterization of myocardial motion patterns by unsupervised multiple kernel learning	ML	Caracterización automática de los patrones de respuesta al estrés en la insuficiencia cardiaca con fracción de eyección preservada
Meystre et al. (2017) <sup>399</sup>	Congestive heart failure information extraction framework for automated treatment performance measures assessment	NLP	Optimización del tratamiento médico de la insuficiencia cardiaca utilizando sistemas de reconocimiento automático de texto
Gleeson et al. (2017) <sup>400</sup>	ECG-derived spatial QRS-T angle is associated with ICD implantation, mortality and heart failure admissions in patients with LV systolic dysfunction	ML	Algoritmo de predicción pronóstica en pacientes con insuficiencia cardiaca con fracción de eyección entre 31-40% evaluando la relación entre el ángulo espacial QRS-T
Xing et al. (2017) <sup>401</sup>	Machine-Learning-Assisted Approach for Discovering Novel Inhibitors Targeting Bromodomain-Containing Protein 4	ML	Predicción de la probabilidad que un compuesto sea un inhibidor de la proteína 4 que contiene bromodominio (BRD4) implicada en la patogénesis de la insuficiencia cardiaca
Iborra-Egea et al. (2017) <sup>402</sup>	Mechanisms of action of sacubitril/valsartan on cardiac remodeling: a systems biology approach	ML	Sistema <i>in silico</i> para conocer los mecanismos de acción del sacubitrilo-valsartán
Oskouie et al. (2017) <sup>403</sup>	Differences in Repolarization Heterogeneity Among Heart Failure With Preserved Ejection Fraction Phenotypic Subgroups	ML	Diferenciación entre diferentes grupos de insuficiencia cardiaca con fracción de eyección preservada en función de la heterogeneidad de la repolarización electrocardiográfica
Steinhoff et al. (2017) <sup>404</sup>	Cardiac Function Improvement and Bone Marrow Response -: Outcome Analysis of the Randomized PERFECT Phase III Clinical Trial of Intramyocardial CD133+ Application After Myocardial Infarction	ML	Identificación de respondedores a terapia celular tras infarto de miocardio y mala función ventricular
Al-Mallah et al. (2017) <sup>405</sup>	Using Machine Learning to Define the Association between Cardiorespiratory Fitness and All-Cause Mortality (from the Henry Ford Exercise Testing Project)	ML	Predicción de mortalidad a 10 años en una cohorte de pacientes que fueron sometidos a prueba de esfuerzo
Myers et al. (2017) <sup>406</sup>	Machine Learning Improves Risk Stratification After Acute Coronary Syndrome	ML	Predicción mortalidad en el síndrome coronario agudo sin elevación del segmento ST

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

van der Ster et al. (2017) <sup>407</sup>	Support Vector Machine Based Monitoring of Cardio-Cerebrovascular Reserve during Simulated Hemorrhage	ML	Algoritmo para predecir parámetros hemodinámicos que facilitarían la detección del shock hipovolémico
Wenhu et al. (2017) <sup>408</sup>	A CHF detection method based on deep learning with RR intervals	DL	Detección de insuficiencia cardíaca basada en el análisis de la variabilidad de frecuencia cardíaca
Mustaqeem et al. (2017) <sup>409</sup>	A statistical analysis based recommender model for heart disease patients	ML	Modelo predictivo de enfermedad coronaria y recomendaciones medicamentosas
Sakr et al. (2017) <sup>410</sup>	Comparison of machine learning techniques to predict all-cause mortality using fitness data: the Henry ford exercise testing (FIT) project	ML	Predicción de mortalidad utilizando datos provenientes de ergometría
Au-Yeung et al. (2018) <sup>411</sup>	Development and validation of warning system of ventricular tachyarrhythmia in patients with heart failure with heart rate variability data	ML	Algoritmo para predecir arritmias ventriculares utilizando la variabilidad de la frecuencia cardíaca en pacientes con DAI procedentes del estudio SCD-HeFT
Wang et al. (2018) <sup>412</sup>	CHF Detection with LSTM Neural Network	DL	Clasificación de insuficiencia cardíaca a partir de la variabilidad de la frecuencia cardíaca
Zhang et al. (2018) <sup>413</sup>	Congestive Heart Failure Detection Via Short-Time Electrocardiographic Monitoring For Fast Reference Advice In Urgent Medical Conditions	DL, ML	Clasificación de insuficiencia cardíaca utilizando intervalos RR, variabilidad frecuencia cardíaca y datos clínicos
McKinley et al. (2018) <sup>414</sup>	Impact of a Pharmacist-Led Intervention on 30-Day Readmission and Assessment of Factors Predictive of Readmission in African American Men With Heart Failure	ML	Sistema de predicción de reingreso por insuficiencia cardíaca en afro-americanos sometidos a un sistema de adherencia e intervención farmacológica
Xiao et al. (2018) <sup>415</sup>	Readmission prediction via deep contextual embedding of clinical concepts	DL	Modelo de predicción de reingresos hospitalarios en pacientes con insuficiencia cardíaca
Wang et al. (2018) <sup>416</sup>	Mortality prediction system for heart failure with orthogonal relief and dynamic radius means	ML	Sistema de predicción de mortalidad en insuficiencia cardíaca
Mahajan et al. (2018) <sup>417</sup>	Predicting Risk of 30-Day Readmissions Using Two Emerging Machine Learning Methods	ML	Predicción de reingreso en paciente con insuficiencia cardíaca
Rasmy et al. (2018) <sup>418</sup>	A study of generalizability of recurrent neural network-based predictive models for heart failure onset risk using a large and heterogeneous EHR data set	DL	Predicción de insuficiencia cardíaca
Lindholm et al. (2018) <sup>419</sup>	Bioimpedance and New-Onset Heart Failure: A Longitudinal Study of >500 000 Individuals From the General Population	ML	Establecer nuevos factores de riesgo de insuficiencia cardíaca para un diagnóstico precoz y tratamiento preventivo
Kwon et al. (2018) <sup>420</sup>	RetainVis: Visual Analytics with Interpretable and Interactive Recurrent Neural Networks on Electronic Medical Records	DL	Sistema interactivo para la predicción de riesgo basado en códigos médicos individuales usado en insuficiencia cardíaca

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Kakarmath et al. (2018) <sup>421</sup>	Validating a Machine Learning Algorithm to Predict 30-Day Re-Admissions in Patients With Heart Failure: Protocol for a Prospective Cohort Study	ML	Predicción rehospitalización en pacientes con insuficiencia cardiaca
Tabassian et al. (2018) <sup>422</sup>	Diagnosis of Heart Failure With Preserved Ejection Fraction: Machine Learning of Spatiotemporal Variations in Left Ventricular Deformation	ML	Análisis del tiempo, amplitud y deformación ventricular izquierda durante el ciclo cardiaco como criterios diagnósticos de insuficiencia cardiaca con función sistólica preservada
Bose et al. (2018) <sup>423</sup>	Using Unsupervised Machine Learning to Identify Subgroups Among Home Health Patients With Heart Failure Using Telehealth	ML	Identificación de usuarios ideales para monitorización domiciliaria con insuficiencia cardiaca
Michalik et al. (2018) <sup>424</sup>	An interactive assistant for patients with cardiac implantable electronic devices: A study protocol of the LUCY trial	ML	Diseño de un estudio multicéntrico para encontrar un modelo predictivo que identifique beneficiarios de un seguimiento de dispositivos implantables en pacientes con insuficiencia cardiaca
Seah et al. (2018) <sup>425</sup>	Chest Radiographs in Congestive Heart Failure: Visualizing Neural Network Learning	DL	Diagnóstico de insuficiencia cardiaca a partir de la clasificación de radiografías de tórax convencionales
Kashihara (2018) <sup>426</sup>	Nonlinear System Identification Based on Convolutional Neural Networks for Multiple Drug Interactions	DL	Predicción de la respuesta hemodinámica a agentes inotrópicos y vasodilatadores
Etemadi et al. (2018) <sup>427</sup>	Wearable ballistocardiogram and seismocardiogram systems for health and performance	ML	Estimación automática del gasto cardiaco, contractilidad y presión arterial con dispositivos móviles
Dogan et al. (2018) <sup>428</sup>	Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study	ML	Predicción de enfermedad coronaria utilizando información del estudio Framingham que incluye datos genéticos
Inan et al. (2018) <sup>69</sup>	Novel Wearable Seismocardiography and Machine Learning Algorithms Can Assess Clinical Status of Heart Failure Patients	ML	Clasificación en estados de descompensación de insuficiencia cardiaca utilizando dispositivos móviles y la respuesta de la frecuencia cardiaca al ejercicio submáximo
Medved et al. (2018) <sup>70</sup>	Improving prediction of heart transplantation outcome using deep learning techniques	DL	Predicción de mortalidad en el trasplante cardiaco y comparación con scores clásicos
Yoon et al. (2018) <sup>429</sup>	Personalized survival predictions via Trees of Predictors: An application to cardiac transplantation	ML	Algoritmo de predicción pronóstica en pacientes trasplantados
Samad et al. (2018) <sup>430</sup>	Predicting deterioration of ventricular function in patients with repaired tetralogy of Fallot using machine learning	ML	Predicción de deterioro ventricular derecho en pacientes con tetralogía de Fallot
Guaricci et al. 2018 <sup>431</sup>	CarDiac MagnEtic Resonance for Primary Prevention Implantable CardioVerter DebrillAtor ThERapy international registry: Design and rationale of the DERIVATE study	ML	Diseño de estudio que pretende mejorar la selección de pacientes con insuficiencia cardiaca que se beneficiarán del implante de DAI
Sanchez-Martinez et al. (2018) <sup>71</sup>	Machine Learning Analysis of Left Ventricular Function to Characterize Heart Failure With Preserved Ejection Fraction	ML	Clasificación de insuficiencia cardiaca con FE preservada frente a sujetos sanos a partir del análisis automático de la función del ventrículo izquierdo en reposo y ejercicio

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Nirschl et al. (2018) <sup>432</sup>	A deep-learning classifier identifies patients with clinical heart failure using whole-slide images of H&E tissue	DL	Clasificador para detectar insuficiencia cardiaca a partir de imágenes de biopsias endomiocárdicas teñidas con hematoxilina-eosina
Ahmad et al. (2018) <sup>433</sup>	Machine Learning Methods Improve Prognostication, Identify Clinically Distinct Phenotypes, and Detect Heterogeneity in Response to Therapy in a Large Cohort of Heart Failure Patients	ML	Identificación de distintos fenotipos de pacientes con insuficiencia cardiaca y la heterogeneidad en la respuesta al tratamiento
Golas et al. (2018) <sup>434</sup>	A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data	ML	Modelo predictivo de reingreso hospitalario a 30 días en pacientes con insuficiencia cardiaca
Zhang et al. (2018) <sup>435</sup>	Discovering and identifying New York heart association classification from electronic health records	NLP	Identificación de la clase funcional en pacientes sometidos a terapia de resincronización cardiaca
Puyol et al. (2018) <sup>436</sup>	Regional Multi-view Learning for Cardiac Motion Analysis: Application to Identification of Dilated Cardiomyopathy Patients	ML	Clasificación automática de pacientes con miocardiopatía dilatada
Sahl et al. (2019) <sup>437</sup>	Multiscale characterization of heart failure	ML	Caracterización multiescala de pacientes de insuficiencia cardiaca
Awan et al. (2019) <sup>438</sup>	Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics	ML	Predicción de riesgo de readmisión o muerte en pacientes de insuficiencia cardiaca
Przewlocka-Kosmala et al. (2019) <sup>439</sup>	Contribution of Cardiovascular Reserve to Prognostic Categories of Heart Failure With Preserved Ejection Fraction: A Classification Based on Machine Learning	ML	Clasificación de pacientes de insuficiencia cardiaca con fracción de eyección preservada
Maragatham et al. (2019) <sup>440</sup>	LSTM Model for Prediction of Heart Failure in Big Data	DL	Redes neuronales LSTM (long short term memory) y big data para la predicción de insuficiencia cardiaca
Al'Aref et al. (2019) <sup>441</sup>	Determinants of In-Hospital Mortality After Percutaneous Coronary Intervention: A Machine Learning Approach	ML	Identificación de predictores de mortalidad intrahospitalaria tras intervención coronaria percutánea
Chung et al. (2019) <sup>442</sup>	Unsupervised classification of multi-omics data during cardiac remodeling using deep learning	DL	Clasificación no supervisada de datos multi-ómicos durante el proceso de remodelación cardiaca
Wang et al. (2019) <sup>443</sup>	Detection of Congestive Heart Failure Based on LSTM-Based Deep Network via Short-Term RR Intervals	DL	Detección de pacientes de insuficiencia cardiaca congestiva
Li et al. (2019) <sup>444</sup>	Time Series Characteristics of Serum Branched-Chain Amino Acids for Early Diagnosis of Chronic Heart Failure	ML	Identificación de biomarcadores para el diagnóstico temprano de insuficiencia cardiaca crónica

**Anexo 5 - Tabla 4. Publicaciones inteligencia artificial en imagen cardiaca**

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Referencia	Título	Modalidad	Observación
Sheet et al. (2014) <sup>445</sup>	Joint learning of ultrasonic backscattering statistical physics and signal confidence primal for characterizing atherosclerotic plaques using intravascular ultrasound	ML	Caracterización tisular de placas de ateroma a partir de imágenes de IVUS
Sun et al. (2014) <sup>446</sup>	A computer-aided diagnostic algorithm improves the accuracy of transesophageal echocardiography for left atrial thrombi: a single-center prospective study	DL	Identificación de trombo en orejuela izquierda a partir del ecocardiograma transesofágico
Takx et al. (2014) <sup>447</sup>	Automated coronary artery calcification scoring in non-gated chest CT: agreement and reliability	ML	Cuantificación automática del score de calcio coronario en pacientes con cáncer sometidos a TC
Afshin et al. (2014) <sup>448</sup>	Regional assessment of cardiac left ventricular myocardial function via MRI statistical features	ML	Segmentación automática del ventrículo izquierdo a partir de imágenes de RM
Sheet et al. (2014) <sup>449</sup>	Hunting for necrosis in the shadows of intravascular ultrasound	ML	Clasificación automática de imágenes de placa de ateroma proveniente de IVUS
Yamak et al. (2014) <sup>450</sup>	Non-calcified coronary atherosclerotic plaque characterization by dual energy computed tomography	ML, DL	Caracterización de la placa de ateroma a partir del análisis de los píxeles procedentes de TC
Wei et al. (2014) <sup>451</sup>	Computerized detection of noncalcified plaques in coronary CT angiography: evaluation of topological soft gradient prescreening method and luminal analysis	ML	Detección de placas coronarias vulnerables a partir de imágenes del TC
Sankaran et al. (2014) <sup>452</sup>	Real-time sensitivity analysis of blood flow simulations to lumen segmentation uncertainty	ML	Estimación en tiempo real de simulaciones del flujo coronario
Ralovich et al. (2014) <sup>453</sup>	6DoF catheter detection, application to intracardiac echocardiography	ML	Detección de catéteres intracardiacos para mejorar la fusión de imágenes
Otaki et al. (2015) <sup>454</sup>	Relationship of epicardial fat volume from noncontrast CT with impaired myocardial flow reserve by positron emission tomography	ML	Relación entre volumen de grasa epicárdica medido por TC y reserva de flujo miocárdico alterada cuantificada por PET
Xiong et al. (2015) <sup>455</sup>	Myocardial perfusion analysis in cardiac computed tomography angiographic images at rest	ML	Determinación de la perfusión miocárdica a partir de imágenes de TC
Gopalakrishnan et al. (2015) <sup>456</sup>	cMRI-BED: A novel informatics framework for cardiac MRI biomarker extraction and discovery applied to pediatric cardiomyopathy classification	ML	Caracterización automática miocárdica a partir de RM
Zha et al. (2015) <sup>457</sup>	3-D Stent Detection in Intravascular OCT Using a Bayesian Network and Graph Search	ML	Análisis e identificación automática de stents a partir de imágenes de OCT
Sankaran et al. (2015) <sup>458</sup>	Fast Computation of Hemodynamic Sensitivity to Lumen Segmentation Uncertainty	ML	Segmentación automática coronaria
Kang et al. (2015) <sup>459</sup>	Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography	ML	Clasificación de lesiones coronarias a partir de TC



Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Born et al. (2015) <sup>460</sup>	Stent Maps--Comparative Visualization for the Prediction of Adverse Events of Transcatheter Aortic Valve Implantations	ML	Predicción de leak paravalvular tras el implante percutáneo de prótesis aórtica
Dey et al. (2015) <sup>461</sup>	Relationship Between Quantitative Adverse Plaque Features From Coronary Computed Tomography Angiography and Downstream Impaired Myocardial Flow Reserve by 13N-Ammonia Positron Emission Tomography: A Pilot Study	ML	Estimación cuantitativa de la placa de ateroma y la disfunción vascular coronaria a partir del TC y PET
Macedo et al. (2015) <sup>462</sup>	A bifurcation identifier for IV-OCT using orthogonal least squares and supervised machine learning	ML	Clasificación de las imágenes de OCT para identificar bifurcaciones coronarias
Emad et al. (2015) <sup>463</sup>	Automatic localization of the left ventricle in cardiac MRI images using deep learning	DL	Localización automática del ventrículo izquierdo en imágenes de eje corto por RM
Knackstedt et al. (2015) <sup>75</sup>	Fully Automated Versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: The FAST-EFs Multicenter Study	ML	Análisis automatizado de las imágenes ecocardiográficas y estimación reproducible de fracción de eyección y <i>strain</i> longitudinal
Shalev et al. (2016) <sup>464</sup>	Classification of calcium in intravascular OCT images for the purpose of intervention planning	DL	Método para la clasificación automática de calcio en imágenes de OCT
Nam et al. (2016) <sup>465</sup>	Automated detection of vessel lumen and stent struts in intravascular optical coherence tomography to evaluate stent apposition and neointimal coverage	DL	Algoritmo que utiliza imágenes de intensidad en la OCT para la detección automática del stent
Moghaddasi et al. (2016) <sup>466</sup>	Automatic assessment of mitral regurgitation severity based on extensive textural features on 2D echocardiography videos	ML	Patrones para clasificar el grado de insuficiencia mitral
Yu et al. (2016) <sup>467</sup>	The Potential of Computational Fluid Dynamics Simulation on Serial Monitoring of Hemodynamic Change in Type B Aortic Dissection	DL	Modelos de simulación de dinámica de fluidos computacional en la monitorización del cambio hemodinámico de la disección aórtica tipo B
Leni et al. (2016) <sup>468</sup>	Development of a 4D numerical chest phantom with customizable breathing	DL	Aplicación que simula la posición y volumen de pulmones, corazón y esófago en cada ciclo respiratorio
Ma et al. (2016) <sup>469</sup>	An SPCNN-GVF-based approach for the automatic segmentation of left ventricle in cardiac cine MR images	DL	Modelo autocorrector para la segmentación automática del ventrículo izquierdo
Zhou et al. (2016) <sup>470</sup>	Watertight modeling and segmentation of bifurcated Coronary arteries for blood flow simulation using CT imaging	DL	Segmentación y construcción de la geometría de las arterias coronarias a partir de imágenes de TC para aplicar directamente en la simulación del flujo coronario sanguíneo
Ciampi et al. (2016) <sup>471</sup>	Computer-aided detection of intracoronary stent in intravascular ultrasound sequences	ML, DL	Sistema automático para la detección de stents intracoronarios a partir de secuencias de imágenes de IVUS
Narula et al. (2016) <sup>74</sup>	Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography	ML	Diferenciación automática entre miocardiopatía hipertrófica familiar y corazón de atleta utilizando datos ecocardiográficos

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Norlem et al. (2016) <sup>472</sup>	Automatic pericardium segmentation and quantification of epicardial fat from computed tomography angiography	ML	Segmentación del pericardio y cuantificación de grasa epicárdica
Nasr-Esfahani et al. (2016) <sup>473</sup>	Vessel extraction in X-ray angiograms using deep learning	DL	Reconstrucción árbol coronario a partir de angiografía
Alzubiedi et al. (2016) <sup>474</sup>	Pharmacogenetic-guided Warfarin Dosing Algorithm in African-Americans	ML, DL	Algoritmo para dosificación de warfarina
Rodrigues et al. (2016) <sup>475</sup>	A novel approach for the automated segmentation and volume quantification of cardiac fats on computed tomography	ML, DL	Cuantificación y diferenciación de la grasa epicárdica y mediastínica a partir de imágenes de TC
Avendi et al. (2016) <sup>93</sup>	A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI	DL	Segmentación automática del ventrículo izquierdo a partir de RM utilizando datos del datatón MICCAI 2009
Itu et al. (2016) <sup>476</sup>	A machine-learning approach for computation of fractional flow reserve from coronary computed tomography	ML	Determinación automática de la FFR a partir de imágenes de TC
Wong et al. (2016) <sup>477</sup>	Regional infarction identification from cardiac CT images: a computer-aided biomechanical approach	ML	Identificación automática de infarto de miocardio a partir de imágenes de TC
Wolterink et al. (2016) <sup>98</sup>	Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks	DL	Cuantificación automática del score de calcio coronario a partir de imágenes de TC
Lungu et al. (2016) <sup>478</sup>	Diagnosis of pulmonary hypertension from magnetic resonance imaging-based computational models and decision tree analysis	ML	Clasificador de hipertensión pulmonar a partir de datos de RM
Hagenah et al. (2016) <sup>479</sup>	Prediction of individual prosthesis size for valve-sparing aortic root reconstruction based on geometric features	ML	Predicción preoperatoria del tamaño de las prótesis quirúrgicas en posición aórtica a partir de datos ecocardiográficos
Liang et al. (2017) <sup>480</sup>	Machine learning-based 3-D geometry reconstruction and modeling of aortic valve deformation using 3-D computed tomography images	ML, DL	Método computacional para reconstruir anatómicamente las geometrías en 3D de la válvula aórtica a partir de imágenes de TC
Su et al. (2017) <sup>481</sup>	An artificial neural network method for lumen and media adventitia border detection in IVUS	DL	Algoritmos de aprendizaje para la detección de los límites de la luz y márgenes vasculares en imágenes de IVUS
Dawes et al. (2017) <sup>482</sup>	Machine Learning of Three-dimensional Right Ventricular Motion Enables Outcome Prediction in Pulmonary Hypertension: A Cardiac MR Imaging Study	ML, DL	Modelo de supervivencia que utiliza el movimiento cardiaco 3D en pacientes con hipertensión pulmonar
Avendi et al. (2017) <sup>483</sup>	Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach	DL	Segmentación automática del ventrículo derecho a partir de imágenes de RM
Yu et al. (2017) <sup>484</sup>	Segmentation of Fetal Left Ventricle in Echocardiographic Sequences Based on Dynamic Convolutional Neural Networks	DL	Segmentación automática ecocardiográfica del ventrículo izquierdo
Liao et al. (2017) <sup>485</sup>	Estimation of the Volume of the Left Ventricle From MRI Images Using Deep Neural Networks	DL	Segmentación del ventrículo izquierdo y cálculo de volumen con datos del datatón: Second Annual Data Science Bowl

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Betancur et al. (2017) <sup>486</sup>	Automatic Valve Plane Localization in Myocardial Perfusion SPECT/CT by Machine Learning: Anatomic and Clinical Validation	ML	Definición del plano valvular mitral durante la segmentación por SPECT utilizando datos de TC
Freiman et al. (2017) <sup>487</sup>	Improving CCTA-based lesions' hemodynamic significance assessment by accounting for partial volume modeling in automatic coronary lumen segmentation	ML, DL	Algoritmo automático de segmentación de la luz coronaria a partir de TC utilizando datos del datatón MICCAI 2012
Wang et al. (2017) <sup>488</sup>	Detecting Cardiovascular Disease from Mammograms With Deep Learning	DL	Detección de calcio intracoronario a partir de imágenes de mamografía para identificar enfermedad coronaria
Rodrigues et al. (2017) <sup>489</sup>	Machine learning in the prediction of cardiac epicardial and mediastinal fat volumes	ML	Cuantificación automática grasa epicárdica y mediastínica a partir de TC
Lian et al. (2017) <sup>490</sup>	An automatic segmentation method of a parameter-adaptive PCNN for medical images	DL	Segmentación automática que integran en uno varios pasos utilizados habitualmente
Tan et al. (2017) <sup>491</sup>	Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences	DL	Segmentación ventricular izquierda en eje corto a partir de imágenes de RM
Larroza et al. (2017) <sup>492</sup>	Differentiation between acute and chronic myocardial infarction by means of texture analysis of late gadolinium enhancement and cine cardiac magnetic resonance imaging	ML	Diferenciación entre infarto de miocardio agudo y crónico a partir de RM
Mazo et al. (2017) <sup>493</sup>	Classification of cardiovascular tissues using LBP based descriptors and a cascade SVM	ML	Clasificación automática de tejidos cardiovasculares basada en su composición
Molaei et al. (2017) <sup>494</sup>	Deep Convolutional Neural Networks for left ventricle segmentation	DL	Segmentación del ventrículo izquierdo en RM
Yang et al. (2017) <sup>495</sup>	Convolutional Neural Network for the Detection of End-Diastole and End-Systole Frames in Free-Breathing Cardiac Magnetic Resonance Imaging	DL	Identificación automática tanto de la diástole como la sístole en imágenes de RM de respiración libre
Duguay et al. (2017) <sup>496</sup>	Coronary Computed Tomographic Angiography-Derived Fractional Flow Reserve Based on Machine Learning for Risk Stratification of Non-Culprit Coronary Narrowings in Patients with Acute Coronary Syndrome	ML	Evaluación pronóstica del FFR determinado por TC en pacientes con síndrome coronario agudo
Nakajima et al. (2017) <sup>497</sup>	Diagnostic accuracy of an artificial neural network compared with statistical quantitation of myocardial perfusion images: a Japanese multicenter study	DL	Diagnóstico a partir de imágenes de gated-SPECT
Mengdi et al. (2017) <sup>498</sup>	Fibroatheroma identification in Intravascular Optical Coherence Tomography images using deep features	DL	Identificación de fibroateroma a partir de imágenes de OCT
Yu et al. (2017) <sup>499</sup>	Determination of Fetal Left Ventricular Volume Based on Two-Dimensional Echocardiography	DL	Cuantificación de volumen ventricular izquierdo fetal por ecocardiografía
Jeganathan et al. (2017) <sup>500</sup>	Artificial intelligence in mitral valve analysis	ML	Evaluación automática de la válvula mitral a partir de imágenes de ecocardiografía transesofágica

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Ngo et al. (2017) <sup>94</sup>	Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance	DL	Segmentación automática del ventrículo izquierdo a partir de imágenes de RM
Abdolmanafi et al. (2017) <sup>99</sup>	Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography	ML, DL	Clasificación automática de las capas arteriales coronarias a partir de imágenes de OCT provenientes de niños con Kawasaki
Tabassian et al. (2017) <sup>501</sup>	Machine learning of the spatio-temporal characteristics of echocardiographic deformation curves for infarct classification	ML	Análisis automático de los perfiles temporales de las curvas de deformación segmentaria del ventrículo izquierdo y sus interrelaciones para poder identificar cambios en su mecánica y función
Yang et al. (2017) <sup>502</sup>	3D Motion Modeling and Reconstruction of Left Ventricle Wall in Cardiac MRI	DL	Modelado de movimiento 3D y análisis de la pared del ventrículo izquierdo en imágenes de RM
Wolterink et al. (2017) <sup>503</sup>	Generative Adversarial Networks for Noise Reduction in Low-Dose CT	DL	Optimización de imágenes TC adquiridas con dosis bajas de radiación
Xue et al. (2017) <sup>504</sup>	Direct Multitype Cardiac Indices Estimation via Joint Representation and Regression Learning	DL	Estimación del índice cardiaco a partir de imágenes de RM
Salem Omar et al. (2017) <sup>89</sup>	Artificial Intelligence-Based Assessment of Left Ventricular Filling Pressures From 2-Dimensional Cardiac Ultrasound Images	ML, DL	Estimación automática de las presiones de llenado del ventrículo izquierdo a partir de imágenes ecocardiográficas
Zahnd et al. (2017) <sup>505</sup>	Contour segmentation of the intima, media, and adventitia layers in intracoronary OCT images: application to fully automatic detection of healthy wall regions	ML	Segmentación y clasificación automática de arterias coronarias sanas o enfermas a partir de imágenes de OCT
Solecki et al. (2017) <sup>506</sup>	What is the optimal anatomic location for coronary artery pressure measurement at CT-derived FFR?	ML	Distancia óptima tras la estenosis coronaria donde estimar la FFR en TC coronario
Banchhor et al. (2017) <sup>507</sup>	Wall-based measurement features provides an improved IVUS coronary artery risk assessment when fused with plaque texture-based features during machine learning paradigm	ML	Predicción del riesgo sumando las características de la pared y la textura de la placa medidas por IVUS
Tesche et al. (2017) <sup>508</sup>	Coronary Computed Tomographic Angiography-Derived Fractional Flow Reserve for Therapeutic Decision Making	ML	Predicción de FFR a partir de TC comparada con estudio invasivo angiográfico
Yong et al. (2017) <sup>509</sup>	Linear-regression convolutional neural network for fully automated coronary lumen segmentation in intravascular optical coherence tomography	DL	Segmentación coronaria a partir de imágenes de OCT
Dangi et al. (2018) <sup>510</sup>	Cine Cardiac MRI Slice Misalignment Correction Towards Full 3D Left Ventricle Segmentation	DL	Segmentación del ventrículo izquierdo
Wardziak et al. (2018) <sup>511</sup>	Coronary CTA enhanced with CTA based FFR analysis provides higher diagnostic value than invasive coronary angiography in patients with intermediate coronary stenosis	ML	Valoración de la predicción de estenosis coronaria mediante análisis de FFR en el TC coronario

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Zhang et al. (2018) <sup>97</sup>	Fully Automated Echocardiogram Interpretation in Clinical Practice	DL	Segmentación cardiaca automática de imágenes ecocardiográficas para cuantificar volúmenes, masa ventricular izquierda, FEVI, y <i>strain</i> longitudinal
Abdolmanafi et al. (2018) <sup>512</sup>	Characterization of coronary artery pathological formations from OCT imaging using deep learning	DL	Modelo para la caracterización del tejido coronario en la enfermedad de Kawasaki procedente de imágenes de OCT
Mazo et al. (2018) <sup>513</sup>	Transfer learning for classification of cardiovascular tissues in histological images	DL	Reconocimiento y clasificación automática de tejidos y órganos
Yao et al. (2018) <sup>514</sup>	An Adaptive Seismocardiography (SCG)-ECG Multimodal Framework for Cardiac Gating Using Artificial Neural Networks	DL	Sistema que utiliza ECG y sismocardiografía para mejorar adquisición y reducir la exposición a la radiación en el TC
van Hamersvelt et al. (2018) <sup>515</sup>	Deep learning analysis of left ventricular myocardium in CT angiographic intermediate-degree coronary stenosis improves the diagnostic accuracy for identification of functionally significant stenosis	DL	Análisis miocárdico del ventrículo izquierdo a partir del TC de pacientes con enfermedad coronaria para identificar lesiones funcionalmente significativas
Yang et al. (2018) <sup>516</sup>	Multiview Sequential Learning and Dilated Residual Learning for a Fully Automatic Delineation of the Left Atrium and Pulmonary Veins from Late Gadolinium-Enhanced Cardiac MRI Images	DL	Método para segmentación de la aurícula izquierda y venas pulmonares a partir de imágenes de RM con gadolinio
Moccia et al. (2018) <sup>517</sup>	Development and testing of a deep learning-based strategy for scar segmentation on CMR-LGE images	DL	Segmentación de cicatrices en ventrículo izquierdo a partir de imágenes de RM
Masuda et al. (2018) <sup>518</sup>	Machine-learning integration of CT histogram analysis to evaluate the composition of atherosclerotic plaques: Validation with IB-IVUS	ML	Caracterización de la placa coronaria con análisis de histograma de la angiografía coronaria por TC en comparación con el método de corte convencional
Kolluru et al. (2018) <sup>519</sup>	Deep neural networks for A-line-based plaque classification in coronary intravascular optical coherence tomography images	DL	Clasificador de tipos de placas de ateroma coronarias a partir de OCT
Zreik et al. (2018) <sup>520</sup>	A Recurrent CNN for Automatic Detection and Classification of Coronary Artery Plaque and Stenosis in Coronary CT Angiography	DL	Clasificación de tipo de placa coronaria y gravedad a partir de imágenes de TC
Tesche et al. (2018) <sup>76</sup>	Coronary CT Angiography-derived Fractional Flow Reserve: Machine Learning Algorithm versus Computational Fluid Dynamics Modeling	ML	Comparación de la determinación de reserva de flujo fraccional derivada de la angiografía por TC o angiografía
Toth et al. (2018) <sup>521</sup>	3D/2D model-to-image registration by imitation learning for cardiac procedures	DL	Creación de modelos anatómicos a partir de TC para el diagnóstico, planificación y orientación de procedimientos terapéuticos
Yu et al. (2018) <sup>522</sup>	CT morphological index provides incremental value to machine learning based CT-FFR for predicting hemodynamically significant coronary stenosis	ML	Algoritmo para diferenciar funcionalmente las lesiones coronarias observadas en TC con referencia al FFR
Slomka et al. (2018) <sup>523</sup>	Rationale and design of the REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT (REFINE SPECT)	ML, DL	Diseño de un registro multicéntrico con datos clínicos e imágenes de SPECT para analizar con inteligencia artificial

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Jin et al. (2018) <sup>524</sup>	Left Atrial Appendage Segmentation Using Fully Convolutional Neural Networks and Modified Three-Dimensional Conditional Random Fields	DL	Segmentación automática de la orejuela izquierda
Zheng et al. (2018) <sup>96</sup>	3-D Consistent and Robust Segmentation of Cardiac Images by Deep Learning With Spatial Propagation	DL	Segmentación cardíaca desde la base al ápex del corazón procedente de RM
Hu et al. (2018) <sup>525</sup>	Diagnostic performance of machine-learning-based computed fractional flow reserve (FFR) derived from coronary computed tomography angiography for the assessment of myocardial ischemia verified by invasive FFR	DL	Comparativa diagnóstica de FFR (fracción de reserva de flujo) determinada por TC vs. Invasiva
Wong et al. (2018) <sup>526</sup>	Building medical image classifiers with very limited data using segmentation networks	DL	Segmentación automática y clasificación de diferentes estructuras (tórax, aorta, arteria pulmonar)
Cano-Espinosa et al. (2018) <sup>527</sup>	Automated Agatston Score Computation in non-ECG Gated CT Scans Using Deep Learning	DL	Cuantificación automática del score de Agatston
Parikh et al. (2018) <sup>528</sup>	Decision Tree Based Classification of Abdominal Aortic Aneurysms Using Geometry Quantification Measures	DL	Cuantificación automática de aneurismas de aorta abdominal
Gessert et al. (2018) <sup>529</sup>	Automatic Plaque Detection in IVOCT Pullbacks Using Convolutional Neural Networks	DL	Detección automática de placas de ateroma a partir de imágenes de OCT
Dorner et al. (2018) <sup>530</sup>	Heart Chamber Segmentation from CT Using Convolutional Neural Networks	DL	Segmentación cámaras cardíacas a partir de imágenes de TC
Garcia et al. (2018) <sup>531</sup>	Diagnostic performance of an artificial intelligence-driven cardiac-structured reporting system for myocardial perfusion SPECT imaging	ML, DL	Sistema automático de análisis de los 17 segmentos cardíacos del SPECT comparado con la cuantificación manual
Bai et al. (2018) <sup>532</sup>	Automated cardiovascular magnetic resonance image analysis with fully convolutional networks	DL	Análisis automático funcional a partir de RM
Xu et al. (2018) <sup>533</sup>	Direct delineation of myocardial infarction without contrast agents using a joint motion feature learning architecture	DL	Desarrollo de un sistema de segmentación del miocardio y caracterización tisular
Sulas et al. (2018) <sup>534</sup>	Automatic Recognition of Complete Atrioventricular Activity in Fetal Pulsed-Wave Doppler Signals	DL	Identificación de las características más importantes de la ecocardiografía-Doppler fetal para el diagnóstico temprano de cardiopatías fetales
Duchateau et al. (2018) <sup>535</sup>	Model-Based Generation of Large Databases of Cardiac Images: Synthesis of Pathological Cine MR Sequences From Real Healthy Cases	ML, DL	Diferenciación automática de secuencias cardíacas patológicas de secuencias reales sanas
Ouzir et al. (2018) <sup>536</sup>	Motion Estimation in Echocardiography Using Sparse Representation and Dictionary Learning	ML, DL	Estimación automática de la movilidad cardíaca a partir de imágenes ecocardiográficas 2D
Luo et al. (2018) <sup>537</sup>	Multi-Views Fusion CNN for Left Ventricular Volumes Estimation on Cardiac MR Images	DL	Estimación automática de los volúmenes del ventrículo izquierdo

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Mannil et al. (2018) <sup>538</sup>	Texture Analysis and Machine Learning for Detecting Myocardial Infarction in Noncontrast Low-Dose Computed Tomography: Unveiling the Invisible	ML, DL	Detección de infarto de miocardio en imágenes de T de baja dosis de radiación sin contraste
Dey et al. (2018) <sup>539</sup>	Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study	ML	Predicción de FFR a partir de imágenes de TC
Larroza et al. (2018) <sup>540</sup>	Texture analysis of cardiac cine magnetic resonance imaging to detect nonviable segments in patients with chronic myocardial infarction	ML	Identificación automática de segmentos miocárdicos no viables a partir de RM
Kolluru et al. (2018) <sup>541</sup>	Voxel-based plaque classification in coronary intravascular optical coherence tomography images using decision trees	ML	Clasificación de tipos de placas por voxels obtenidos de OCT de arterias coronarias de cadáveres
Schlemper et al. (2018) <sup>101</sup>	A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction	DL	Reconstrucción de imágenes a partir de RM
Betancur et al. (2018) <sup>83</sup>	Prognostic Value of Combined Clinical and Myocardial Perfusion Imaging Data Using Machine Learning	ML	Predictor pronóstico combinando datos clínicos y de SPECT
Zreik et al. (2018) <sup>102</sup>	Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis	DL	Identificación de isquemia miocárdica a partir de TC de pacientes con enfermedad coronaria
Nakajima et al. (2018) <sup>542</sup>	Artificial neural network retrained to detect myocardial ischemia using a Japanese multicenter database	DL	Detección automática de isquemia miocárdica usando SPECT
Dormer et al. (2018) <sup>543</sup>	Convolutional Neural Networks for the Detection of Diseased Hearts Using CT Images and Left Atrium Patches	DL	Modelo para identificar enfermedad estructural cardíaca a partir de TC convencionales
Shao et al. (2018) <sup>544</sup>	Texture analysis of magnetic resonance T1 mapping with dilated cardiomyopathy: A machine learning approach	ML	Diagnóstico de miocardiopatía dilatada a partir de parámetros del T1 en resonancia magnética
Dong et al. (2018) <sup>545</sup>	A Combined Fully Convolutional Networks and Deformable Model for Automatic Left Ventricle Segmentation Based on 3D Echocardiography	DL	Segmentación del ventrículo izquierdo a partir de imágenes 3D ecocardiográficas
Betancur et al. (2018) <sup>546</sup>	Deep Learning Analysis of Upright-Supine High-Efficiency SPECT Myocardial Perfusion Imaging for Prediction of Obstructive Coronary Artery Disease: A Multicenter Study	DL	Predicción de enfermedad coronaria tras análisis combinado de SPECT en posiciones semivertical y supina
Shibutani et al. (2018) <sup>547</sup>	Accuracy of an artificial neural network for detecting a regional abnormality in myocardial perfusion SPECT	DL	Cuantificación automática de la extensión y gravedad de la isquemia miocárdica a partir de imágenes de SPECT
Narang et al. (2018) <sup>548</sup>	Machine learning based automated dynamic quantification of left heart chamber volumes	ML	Algoritmo automático para el análisis volumétrico de datos ecocardiográficos 3D en la medición precisa de volúmenes de ventrículo y aurícula izquierdos
Hasegawa et al. (2018) <sup>549</sup>	[Automated Classification of Calcification and Stent on Computed Tomography Coronary Angiography Using Deep Learning]	DL	Reconocer automáticamente calcificaciones o stents en las arterias coronarias y diferenciarlas de arterias normales a partir de imágenes de TC

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Abdolmanafi et al. (2018) <sup>550</sup>	Intra-slice motion correction of intravascular OCT images using deep features	DL	Corrección del movimiento para el análisis automático de las variaciones de volumen y formas de imágenes en OCT
Du et al. (2018) <sup>551</sup>	Direct Segmentation-based Full Quantification for Left Ventricle via Deep Multi-task Regression Learning Network	DL	Segmentación del ventrículo izquierdo, pero en lugar de usar imágenes crudas, se crea un contorno segmentado para estimar los índices cardíacos
Mastrodicasa et al. (2018) <sup>552</sup>	Artificial intelligence machine learning-based coronary CT fractional flow reserve (CT-FFRML): Impact of iterative and filtered back projection reconstruction techniques	DL	Algoritmo de reconstrucción a partir de imágenes de TC para estimar el FFR
Jun et al. (2018) <sup>553</sup>	Automated detection of vulnerable plaque in intravascular ultrasound images	ML, DL	Clasificación de placa vulnerable a partir de imágenes de IVUS
Huang et al. (2018) <sup>554</sup>	Coronary Artery Segmentation by Deep Learning Neural Networks on Computed Tomographic Coronary Angiographic Images	DL	Segmentación automática de la luz arterial coronaria a partir de imágenes de TC
Que et al. (2018) <sup>555</sup>	CardioXNet: Automated Detection for Cardiomegaly Based on Deep Learning	DL	Sistema de identificación automática de cardiomegalia en la radiografía de tórax convencional
Nasr-Esfahani et al. (2018) <sup>556</sup>	Left Ventricle Segmentation in Cardiac MR Images Using Fully Convolutional Network	DL	Segmentación automática del ventrículo a partir de imágenes de resonancia magnética
Omar et al. (2018) <sup>557</sup>	Automated Myocardial Wall Motion Classification using Handcrafted Features vs a Deep CNN-based mapping	ML, DL	Clasificación automática de la motilidad cardíaca a partir de ecocardiografía
Leng et al. (2018) <sup>558</sup>	Computational Platform Based on Deep Learning for Segmenting Ventricular Endocardium in Long-axis Cardiac MR Imaging	DL	Segmentación automática endocárdica de ventrículos izquierdo y derecho
Ostvik et al. (2018) <sup>559</sup>	Real-Time Standard View Classification in Transthoracic Echocardiography Using Convolutional Neural Networks	DL	Automatización en la secuencia de trabajo y adquisición de planos en ecocardiografía
Xue et al. (2018) <sup>560</sup>	Extraction of Aortic Knuckle Contour in Chest Radiographs Using Deep Learning	DL	Segmentación automática del botón aórtico
Kang et al. (2018) <sup>561</sup>	Cycle Consistent Adversarial Denoising Network for Multiphase Coronary CT Angiography	DL	Optimización de la adquisición de imágenes de TC obtenidas con baja radiación
Lossau et al. (2018) <sup>562</sup>	Motion artifact recognition and quantification in coronary CT angiography using convolutional neural networks	DL	Optimización del movimiento en la adquisición de imágenes de TC
Zhang et al. (2018) <sup>563</sup>	Automatic Assessment of Full Left Ventricular Coverage in Cardiac Cine Magnetic Resonance Imaging with Fisher Discriminative 3D CNN	DL	Segmentación ventricular izquierda de base a ápex a partir de imágenes de RM
Huang et al. (2018) <sup>564</sup>	Enhancing the prediction of acute kidney injury risk after percutaneous coronary intervention using machine learning techniques: A retrospective cohort study	ML	Modelo predictivo de daño renal tras administración de contraste en el intervencionismo percutáneo



Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Nous et al. (2018) <sup>565</sup>	Comparison of the Diagnostic Performance of Coronary Computed Tomography Angiography-Derived Fractional Flow Reserve in Patients With Versus Without Diabetes Mellitus (from the MACHINE Consortium)	ML	Estimación de FFR a partir de imágenes de TC comparativo entre diabéticos y no diabéticos
Kwon et al. (2018) <sup>566</sup>	Deep learning for predicting in-hospital mortality among heart disease patients based on echocardiography	DL	Modelo predictivo de mortalidad intrahospitalaria a partir de datos ecocardiográficos durante la hospitalización
von Knebel Doeberitz et al. (2018) <sup>567</sup>	Coronary CT angiography-derived plaque quantification with artificial intelligence CT fractional flow reserve for the identification of lesion-specific ischemia	DL	Estimación de FFR a partir de imágenes de TC
Han et al. (2018) <sup>568</sup>	Incremental role of resting myocardial computed tomography perfusion for predicting physiologically significant coronary artery disease: A machine learning approach	ML	Predicción de isquemia miocárdica a partir de imágenes de miocardio y coronarias por TC
Isgum et al. (2018) <sup>569</sup>	Automatic determination of cardiovascular risk by CT attenuation correction maps in Rb-82 PET/CT	ML	Cuantificación automática del calcio score coronario a partir de PET/TC
Baessler et al. (2018) <sup>570</sup>	Subacute and Chronic Left Ventricular Myocardial Scar: Accuracy of Texture Analysis on Nonenhanced Cine MR Images	ML	Clasificación e identificación de tejido cardíaco como isquémico/necrótico a partir de imágenes de RM
Oktay et al. (2018) <sup>571</sup>	Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation	DL	Sistema de segmentación que incorpora las propiedades anatómicas globales de la anatomía subyacente en diferentes patologías cardíacas
Xue et al. (2018) <sup>572</sup>	Full left ventricle quantification via deep multitask relationships learning	DL	Cuantificación del ventrículo izquierdo; que tiene en cuenta dos áreas (cavidad y miocardio), espesores regionales de pared, tres dimensiones y fases diastólica y sistólica
Wojnarski et al. (2018) <sup>573</sup>	Machine-learning phenotypic classification of bicuspid aortopathy	ML	Clasificación de diferentes tipos de aortopatía en pacientes con válvula aórtica bicúspide a partir de imágenes de TC
Tan et al. (2018) <sup>574</sup>	Fully automated segmentation of the left ventricle in cine cardiac MRI using neural network regression	DL	Segmentación automática del ventrículo izquierdo a partir de RM
Winther et al. (2018) <sup>575</sup>	v-net: Deep Learning for Generalized Biventricular Mass and Function Parameters Using Multicenter Cardiac MRI Data	DL	Segmentación automática de ventrículos izquierdo y derecho
Lessmann et al. (2018) <sup>576</sup>	Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks With Dilated Convolutions	DL	Detección automática de calcificaciones en arteria coronaria, aorta torácica, y válvula aórtica a partir de TC utilizando baja radiación
Betancur et al. (2018) <sup>577</sup>	Deep Learning for Prediction of Obstructive Disease From Fast Myocardial Perfusion SPECT: A Multicenter Study	DL	Predicción de enfermedad coronaria obstructiva a partir de imágenes de perfusión de SPECT
Nakanishi et al. (2018) <sup>578</sup>	Automated estimation of image quality for coronary computed tomographic angiography using machine learning	ML	Evaluación automatizada de imágenes coronarias procedentes de TC y comparación frente al análisis visual

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Gernaat et al. (2018) <sup>579</sup>	Automatic quantification of calcifications in the coronary arteries and thoracic aorta on radiotherapy planning CT scans of Western and Asian breast cancer patients	DL	Cuantificación automática de la calcificación de arterias coronarias y aorta torácica a partir de TC para programar radioterapia en cáncer de mama
Lancaster et al. (2018) <sup>90</sup>	Phenotypic Clustering of Left Ventricular Diastolic Function Parameters: Patterns and Prognostic Relevance	ML	Clasificación no supervisada de variables ecocardiográficas para evaluar la disfunción ventricular izquierda diastólica
Baessler et al. (2018) <sup>580</sup>	Texture analysis and machine learning of non-contrast T1-weighted MR images in patients with hypertrophic cardiomyopathy-Preliminary results	ML	Clasificación de la textura del miocardio buscando áreas de fibrosis en miocardiopatía hipertrófica a partir de imágenes de RM sin contraste
Fu et al. (2018) <sup>581</sup>	Segmentation of histological images and fibrosis identification with a convolutional neural network	DL	Segmentación de imágenes histológicas, particularmente aquellas con tinción tricrómica de Masson
Coenen et al. (2018) <sup>582</sup>	Diagnostic Accuracy of a Machine-Learning Approach to Coronary Computed Tomographic Angiography-Based Fractional Flow Reserve: Result From the MACHINE Consortium	ML	Estimación de FFR a partir de TC coronario y comparación frente a estimaciones angiográficas convencionales
Cruz-Aceves et al. (2018) <sup>583</sup>	A Novel Multiscale Gaussian-Matched Filter Using Neural Networks for the Segmentation of X-Ray Coronary Angiograms	DL	Segmentación automática de las arterias coronarias a partir de imágenes angiográficas de fluoroscopia
Vigneault et al. (2018) <sup>584</sup>	Omega-Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks	DL	Segmentación automática simultánea de píxeles en pacientes con miocardiopatía hipertrófica
Commandeur et al. (2018) <sup>585</sup>	Deep Learning for Quantification of Epicardial and Thoracic Adipose Tissue From Non-Contrast CT	DL	Cuantificación automática de la grasa epicárdica y torácica a partir de imágenes de TC sin contraste
Pavoni et al. (2018) <sup>586</sup>	Convolutional neural network-based image enhancement for x-ray percutaneous coronary intervention	DL	Optimización de las imágenes angiográficas procedentes de fluoroscopia con baja radiación
Al et al. (2018) <sup>587</sup>	Automatic aortic valve landmark localization in coronary CT angiography using colonial walk	ML	Optimización automática de la localización valvular aórtica para implante de TAVI a partir de imágenes TC angiográfico
Qin et al. (2018) <sup>588</sup>	Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction	DL	Reconstrucción automática de imágenes de RM de alta calidad
Chen et al. (2018) <sup>589</sup>	Correlated Regression Feature Learning for Automated Right Ventricle Segmentation	DL	Segmentación automática del ventrículo derecho a partir de imágenes de RM
Wang et al. (2018) <sup>590</sup>	Neural network fusion: a novel CT-MR Aortic Aneurysm image segmentation method	DL	Segmentación de imagen multimodal; RM y TC
Green et al. (2018) <sup>591</sup>	3-D Neural denoising for low-dose Coronary CT Angiography (CCTA)	DL	Reconstrucción automática de la anatomía coronaria por TC reduciendo las dosis de radiación
Strange et al. (2018) <sup>592</sup>	The National Echocardiography Database Australia (NEDA): Rationale and methodology	ML	Base de datos nacional ecocardiográfica con el objetivo de realizar análisis con inteligencia artificial
Yang et al. (2018) <sup>593</sup>	Neural multi-atlas label fusion: Application to cardiac MR images	DL	Segmentación multiatlas y multifusión a partir de imágenes de RM

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Zotti et al. (2018) <sup>594</sup>	Convolutional Neural Network with Shape Prior Applied to Cardiac MRI Segmentation	DL	Segmentación de imágenes de eje corto de RM
Fahmy et al. (2018) <sup>595</sup>	Automated Cardiac MR Scar Quantification in Hypertrophic Cardiomyopathy Using Deep Convolutional Neural Networks	DL	Cuantificación automática del tejido fibroso en la miocardiopatía hipertrófica procedente de imágenes de RM
Samad et al. (2018) <sup>73</sup>	Predicting Survival From Large Echocardiography and Electronic Health Record Datasets	ML	Predicción de mortalidad en pacientes sometidos a ecocardiografía
Bratt et al. (2019) <sup>596</sup>	Machine learning derived segmentation of phase velocity encoded cardiovascular magnetic resonance for fully automated aortic flow quantification	DL	Cuantificación automatizada del flujo aórtico a través de la segmentación de resonancias magnéticas con machine learning
Togo et al. (2019) <sup>597</sup>	Cardiac sarcoidosis classification with deep convolutional neural network-based features using polar maps	DL	Clasificación de pacientes con o sin sarcoidosis cardiaca a través de redes neuronales y tomografía computarizada
Curiale et al. (2019) <sup>598</sup>	Automatic quantification of the LV function and mass: A deep learning approach for cardiovascular MRI	DL	Cuantificación de masa y función del ventrículo izquierdo a partir de resonancia magnética
Biswas et al. (2019) <sup>599</sup>	Dynamic MRI using model-based deep learning and STORM priors: MoDL-STORM	DL	Resonancia magnética cardiaca dinámica a través de modelos deep learning
Xiong et al. (2019) <sup>600</sup>	Fully Automatic Left Atrium Segmentation From Late Gadolinium Enhanced Magnetic Resonance Imaging Using a Dual Fully Convolutional Neural Network	DL	Segmentación completa de la aurícula izquierda a través de redes neuronales convolucionales aplicadas a resonancias magnéticas con realce de gadolinio
Bahrami et al. (2019) <sup>601</sup>	Automated selection of myocardial inversion time with a convolutional neural network: Spatial temporal ensemble myocardium inversion network (STEMI-NET)	DL	Selección automática del tiempo de inversión miocárdica en resonancias magnéticas
Goldfarb et al. (2019) <sup>602</sup>	Water-fat separation and parameter mapping in cardiac MRI via deep learning with a convolutional neural network	DL	Mapeado de tejido adiposo en imagen por resonancia magnética cardiaca a través de redes neuronales convolucionales
Du et al. (2019) <sup>603</sup>	Cardiac-DeepIED: Automatic Pixel-Level Deep Segmentation for Cardiac Bi-Ventricle Using Improved End-to-End Encoder-Decoder Network	DL	Segmentación de ventrículos derecho e izquierdo en imagen por resonancia
Khened et al. (2019) <sup>604</sup>	Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers	DL	Segmentación y diagnóstico automático de imágenes por resonancia magnética cardiaca a través de redes neuronales convolucionales residuales multiescala
Wang et al. (2019) <sup>605</sup>	A learning-based automatic segmentation and quantification method on left ventricle in gated myocardial perfusion SPECT imaging: A feasibility study	DL	Segmentación automática del ventrículo izquierdo en imagen de perfusión miocárdica SPECT
Duan et al. (2019) <sup>606</sup>	Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach	DL	Segmentación 3D biventricular de imágenes por resonancia magnética cardiaca

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Priya et al. (2019) <sup>607</sup>	Adaptive Fruitfly Based Modified Region Growing Algorithm for Cardiac Fat Segmentation Using Optimal Neural Network	DL	Segmentación de tejido adiposo cardiaco
Hauptmann et al. (2019) <sup>608</sup>	Real-time cardiovascular MR with spatio-temporal artifact suppression using deep learning-proof of concept in congenital heart disease	DL	Redes neuronales convolucionales para la reconstrucción de resonancias magnéticas cardiacas
Liu et al. (2019) <sup>609</sup>	Multi-sequence myocardium segmentation with cross-constrained shape and neural network-based initialization	DL	Segmentación multisequencia del miocardio
Zhao et al. (2019) <sup>610</sup>	Deep feature regression (DFR) for 3D vessel segmentation	DL	Segmentación 3D de vasos coronarios
De Vos et al. (2019) <sup>611</sup>	A deep learning framework for unsupervised affine and deformable image registration	DL	Registro de secuencias de imágenes por resonancia magnética cardiaca y tomografías de tórax
Wang et al. (2019) <sup>612</sup>	Diagnostic accuracy of a deep learning approach to calculate FFR from coronary CT angiography	DL	Precisión diagnóstica de los enfoques basados en deep learning para el cálculo de la reserva fraccional de flujo a partir de tomografía coronaria
Murphy et al. (2019) <sup>613</sup>	Noise-robust bioimpedance approach for cardiac output measurement	DL	Monitorización de parámetros cardiacos de pacientes de insuficiencia cardiaca congestiva
va et al. (2019) <sup>614</sup>	Application of speCtraL computed tomogrAphy to impRove specIficity of cardiac compuTed tomographY (CLARITY study): rationale and design	DL	Aplicación de tomografía espectral para mejorar la especificidad en la detección de estenosis de la tomografía computarizada cardiaca
Wu et al. (2019) <sup>615</sup>	Automated anatomical labeling of coronary arteries via bidirectional tree LSTMs	DL	Reconstructor automático del árbol coronario a partir de imágenes de TC
Wolterink et al. (2019) <sup>616</sup>	Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier	DL	Predicción de la dirección y diámetro de una arteria coronaria a partir de imágenes de TC
Tao et al. (2019) <sup>617</sup>	Deep Learning-based Method for Fully Automatic Quantification of Left Ventricle Function from Cine MR Images: A Multivendor, Multicenter Study	DL	Cuantificación automática de la función ventricular izquierda a partir de imágenes de resonancia magnética y evaluación multimarca y multicéntrica
Leclerc et al. (2019) <sup>618</sup>	Deep Learning for Segmentation using an Open Large-Scale Dataset in 2D Echocardiography	DL	Segmentación de ecocardiogramas
Volpato et al. (2019) <sup>619</sup>	Automated, machine learning-based, 3D echocardiographic quantification of left ventricular mass	ML	Cuantificación automatizada de masa del ventrículo izquierdo basada en ecocardiografía 3D

**Anexo 5 - Tabla 5. Publicaciones inteligencia artificial en otras aplicaciones en cardiología**

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Referencia	Título	Modalidad	Observación
Goldstein et al. (2014) <sup>620</sup>	Near-term prediction of sudden cardiac death in older hemodialysis patients using electronic health records	ML	Predictor de muerte súbita en pacientes ancianos en hemodiálisis
Ahmad et al. (2014) <sup>621</sup>	Machine learning classification of cell-specific cardiac enhancers uncovers developmental subnetworks regulating progenitor cell division and cell fate specification	ML	Clasificación de las células implicadas en el desarrollo cardiaco en modelo animal
A. Wahab et al. (2014) <sup>622</sup>	Comparing time to adverse drug reaction signals in a spontaneous reporting database and a claims database: a case study of rofecoxib-induced myocardial infarction and rosiglitazone-induced heart failure signals in Australia	DL	Método para evaluar la seguridad cardiovascular en la postcomercialización de fármacos, concretamente rofecoxib y rosiglitazona
Liu et al. (2014) <sup>623</sup>	Improvement of adequate use of warfarin for the elderly using decision tree-based approaches	ML	Mejora en la efectividad del uso de antagonistas de la vitamina K en pacientes ancianos
Perry et al. (2014) <sup>624</sup>	Supervised embedding of textual predictors with applications in clinical diagnostics for pediatric cardiology	NLP, ML	Clasificación de pacientes pediátricos con cardiopatía
Gharehbaghi et (2014) <sup>625</sup>	Detection of systolic ejection click using time growing neural network	DL	Clasificación de los ruidos cardiacos
Elgendi et al. (2014) <sup>626</sup>	Spectral analysis of the heart sounds in children with and without pulmonary artery hypertension	ML	Diagnóstico de hipertensión pulmonar en niños a partir del análisis de ruidos cardiacos a través de fonendoscopios
Elgendi et al. (2014) <sup>627</sup>	Time-domain analysis of heart sound intensity in children with and without pulmonary artery hypertension: a pilot study using a digital stethoscope	ML	Diagnóstico de hipertensión pulmonar en niños a partir del análisis temporales de ruidos cardiacos a través de fonendoscopios
Nouei et al. (2014) <sup>628</sup>	Developing a genetic fuzzy system for risk assessment of mortality after cardiac surgery	ML, DL	Sistema predictivo de mortalidad tras cirugía cardiaca
Gorospe et al. (2014) <sup>629</sup>	Automated grouping of action potentials of human embryonic stem cell-derived cardiomyocytes	ML	Clasificación fenotípica de cardiomiocitos derivados de células humanas embrionarias
LaFaro et al. (2015) <sup>630</sup>	Neural Network Prediction of ICU Length of Stay Following Cardiac Surgery Based on Pre-Incision Variables	DL	Modelo de red neuronal para predecir la duración de la estancia en UCI tras cirugía cardiaca
Economou et al. (2015) <sup>631</sup>	Exploiting expert systems in cardiology: a comparative study	DL	Sistema de clasificación de diferentes patologías cardiovasculares: enfermedad coronaria, hipertensión arterial, fibrilación auricular, insuficiencia cardiaca y diabetes
Han et al. (2015) <sup>632</sup>	The relationship between left ventricle myocardial performance index of healthy women and geographical factors	DL	Relación entre los factores geográficos y el índice de rendimiento del miocardio del ventrículo izquierdo
Huang et al. (2015) <sup>633</sup>	Post-operative bleeding risk stratification in cardiac pulmonary bypass patients using artificial neural network	DL	Predicción de sangrado en pacientes sometidos a pontaje aortocoronario
Nahato et al. (2015) <sup>634</sup>	Knowledge mining from clinical datasets using rough sets and backpropagation neural network	ML	Clasificador de enfermedad cardiaca en general a partir de grandes bases de datos

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Wise et al. (2015) <sup>635</sup>	Prediction of in-hospital mortality after ruptured abdominal aortic aneurysm repair using an artificial neural network	DL	Predicción de muerte hospitalaria tras reparación quirúrgica de aneurismas abdominales
Lee et al. (2015) <sup>636</sup>	Machine learning plus optical flow: a simple and sensitive method to detect cardioactive drugs	ML	Detección de efectos cardiotoxicos de fármacos
Mendes et al. (2015) <sup>637</sup>	Predicting reintubation, prolonged mechanical ventilation and death in post-coronary artery bypass graft surgery: a comparison between artificial neural networks and logistic regression models	ML, DL	Predicción de complicaciones tras cirugía de revascularización coronaria
Liu et al. (2015) <sup>638</sup>	Manifold ranking based scoring system with its application to cardiac arrest prediction: A retrospective study in emergency department patients	ML	Predicción de muerte súbita en pacientes en los servicios de emergencias
Luo et al. (2015) <sup>639</sup>	Using Computational Approaches to Improve Risk-Stratified Patient Management: Rationale and Methods	ML	Diseño de estudio para analizar un modelo predictivo de resultado en salud y costes clasificando el riesgo del paciente con diferentes enfermedades crónicas, incluyendo cardiacas
Mohamadipanah et al. (2015) <sup>640</sup>	Predictive Model Reference Adaptive Controller to Compensate Heart Motion in Minimally Invasive CABG Surgery	ML	Modelo para corregir el movimiento cardiaco durante la cirugía mínimamente invasiva
Gharehbaghi et al. (2015) <sup>641</sup>	A Novel Method for Screening Children with Isolated Bicuspid Aortic Valve	ML, DL	Diagnóstico de válvula aórtica bicúspide en niños a partir del análisis de ruidos cardiacos
Elgendi et al. (2016) <sup>642</sup>	The unique heart sound signature of children with pulmonary artery hypertension	ML	Análisis de patrones en la auscultación para identificar niños con hipertensión pulmonar
Bashir et al. (2016) <sup>643</sup>	IntelliHealth: A medical decision support application using a novel weighted multi-layer classifier ensemble framework	ML	Sistema de clasificación validado en diferentes patologías cardiacas
Monsalve-Torra et al. (2016) <sup>644</sup>	Using machine learning methods for predicting inhospital mortality in patients undergoing open repair of abdominal aortic aneurysm	ML	Sistema para predecir mortalidad hospitalaria tras cirugía de aorta abdominal
Sengupta et al. (2016) <sup>77</sup>	Cognitive Machine-Learning Algorithm for Cardiac Imaging: A Pilot Study for Differentiating Constrictive Pericarditis From Restrictive Cardiomyopathy	ML	Algoritmo de diferenciación diagnóstica entre pericarditis constrictiva y miocardiopatía restrictiva
Wang et al. (2016) <sup>645</sup>	ADMET Evaluation in Drug Discovery. 16. Predicting hERG Blockers by Combining Multiple Pharmacophores and Machine Learning Approaches	ML	Modelo predictivo para evaluar el riesgo de cardiotoxicidad en las primeras etapas de diseño de fármacos bloqueadores hERG
Antink et al. (2016) <sup>646</sup>	Reducing false alarms in the ICU by quantifying self-similarity of multimodal biosignals	ML	Sistema de control de alarmas para unidades de críticos
Hu et al. (2016) <sup>647</sup>	Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model	DL	Sistema de predicción de alarmas en la monitorización de parada cardiaca en pacientes en unidades críticos
Seyres et al. (2016) <sup>648</sup>	Identification and in silico modeling of enhancers reveals new features of the cardiac differentiation network	ML	Utilización de inteligencia artificial para identificar factores de transcripción responsables en el desarrollo cardiaco en <i>Drosophila</i>

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Hansen et al. (2016) <sup>649</sup>	Identifying Drug-Drug Interactions by Data Mining: A Pilot Study of Warfarin-Associated Drug Interactions	ML	Identificación de interacciones farmacológicas de los antagonistas de vitamina K
Thompson et al. (2016) <sup>650</sup>	Predictive models for mortality after ruptured aortic aneurysm repair do not predict futility and are not useful for clinical decision making	DL	Estudio neutro con un modelo que no mejoraba la predicción de muerte tras cirugía de rotura de aorta abdominal
Waardenberg et al. (2016) <sup>651</sup>	Prediction and validation of protein-protein interactors from genome-wide DNA-binding data using a knowledge-based machine-learning approach	ML	Predicción de las interacciones de la proteína NKX2-5 basadas en datos del genoma completo
Ruiz-Fernandez et al. (2016) <sup>652</sup>	Aid decision algorithms to estimate the risk in congenital heart surgery	ML	Clasificación de riesgo en la cirugía de cardiopatías congénitas
Wolf et al. (2016) <sup>653</sup>	Rationale and methodology of a collaborative learning project in congenital cardiac care	DL	Sistema en red, 5 hospitales,, para mejorar los resultados en cardiopatías congénitas
Tang et al. (2017) <sup>654</sup>	A non-invasive approach to investigation of ventricular blood pressure using cardiac sound features	DL	Determinación de la presión arterial a partir del análisis de ruidos cardiacos
Kario (2017) <sup>655</sup>	Perfect 24-h management of hypertension: clinical relevance and perspectives	ML	Sistema para control ambulatorio de la presión arterial
Flores et al. (2017) <sup>656</sup>	Predicting the physiological response of Tivela stultorum hearts with digoxin from cardiac parameters using artificial neural networks	DL	Predicción de la concentración de digoxina necesaria para obtener la máxima cardioactividad
Schiltz et al. (2017) <sup>657</sup>	Identifying Specific Combinations of Multimorbidity that Contribute to Health Care Resource Utilization: An Analytic Approach	ML	Identificación de enfermedades crónicas, limitaciones funcionales y síndromes geriátricos asociados con costes sanitarios y utilización de recursos hospitalarios
Allyn et al. (2017) <sup>658</sup>	A Comparison of a Machine Learning Model with EuroSCORE II in Predicting Mortality after Elective Cardiac Surgery: A Decision Curve Analysis	ML	ML es más preciso en la predicción de mortalidad que el EuroSCORE II en pacientes sometidos a cirugía cardiaca electiva
Wise et al. (2017) <sup>659</sup>	Prediction of Prolonged Ventilation after Coronary Artery Bypass Grafting: Data from an Artificial Neural Network	DL	Identificación de factores preoperatorios asociados a ventilación prolongada e identificación temprana de pacientes que la puedan requerir tras cirugía de pontaje coronario
Pereira et al. (2017) <sup>660</sup>	Automated detection of coarctation of aorta in neonates from two-dimensional echocardiograms	DL	Detección de coartación de aorta utilizando imágenes ecocardiográficas
Li et al. (2017) <sup>661</sup>	An artificial neural network prediction model of congenital heart disease based on risk factors: A hospital-based case-control study	DL	Predicción de cardiopatías congénitas en fetos de mujeres embarazadas
Zhang et al. (2017) <sup>662</sup>	[Research on prediction method of left ventricular blood pressure based on external heart sounds]	DL	Modelo predictivo en perros para determinar la presión ventricular izquierda a partir de los ruidos cardiacos
Gharehbaghi et al. (2017) <sup>663</sup>	A Decision Support System for Cardiac Disease Diagnosis Based on Machine Learning Methods	ML	Sistema para screening de cardiopatías congénitas en niños en atención primaria basado en el análisis de los ruidos cardiacos



Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Radchenko et al. (2017) <sup>664</sup>	Computer-aided estimation of the hERG-mediated cardiotoxicity risk of potential drug components	DL	Modelo predictivo de cardiotoxicidad de futuros fármacos a través del canal de potasio hERG
Eslamizadeh et al. (2017) <sup>665</sup>	Heart murmur detection based on wavelet transformation and a synergy between artificial neural network and modified neighbor annealing methods	DL	Clasificador de soplos cardiacos
Sudha (2017) <sup>666</sup>	Evolutionary and Neural Computing Based Decision Support System for Disease Diagnosis from Clinical Data Sets in Medical Practice	DL	Identificador de enfermedad cardiaca
Awad et al. (2017) <sup>667</sup>	Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach	ML	Predictor de mortalidad en pacientes ingresados en unidades de críticos
Xiao et al. (2017) <sup>668</sup>	Estimation of aortic systolic blood pressure from radial systolic and diastolic blood pressures alone using artificial neural networks	DL	Estimación de la presión sistólica en aorta a partir de las presiones sistólicas y diastólicas de arteria radial
Indja et al. (2017) <sup>669</sup>	Neural network imaging to characterize brain injury in cardiac procedures: the emerging utility of connectomics	DL	Caracterización de la lesión cerebral isquémica postquirúrgica
Liang et al. (2017) <sup>670</sup>	A machine learning approach to investigate the relationship between shape features and numerically predicted risk of ascending aortic aneurysm	ML	Predicción automática de rotura de aneurismas de aorta ascendente
Abdollahpur et al. (2017) <sup>671</sup>	Detection of pathological heart sounds	ML, DL	Clasificación automática de ruidos cardiacos a partir del datatón PhysioNet/CinC Challenge 2016
No listado autores (2017) <sup>672</sup>	[Quantitative structure-activity relationship model for prediction of cardiotoxicity of chemical components in traditional Chinese medicines]	ML	Predicción de la posible cardiotoxicidad de los componentes químicos de las medicinas tradicionales chinas
Gharehbaghi et al. (2017) <sup>673</sup>	Intelligent Phonocardiography for Screening Ventricular Septal Defect Using Time Growing Neural Network	DL	Clasificación entre comunicación interventricular o insuficiencia valvular aurículo-ventricular a partir de la auscultación cardiaca
Silva et al. (2017) <sup>674</sup>	Combined Thoracic Ultrasound Assessment during a Successful Weaning Trial Predicts Postextubation Distress	ML	Predicción de estrés respiratorio postextubación a partir de datos ecocardiográficos respiratorios, cardiacos y de diafragma
Ebert et al. (2017) <sup>675</sup>	Automatic detection of hemorrhagic pericardial effusion on PMCT using deep learning - a feasibility study	DL	Detección y segmentación de hemopericardio postmortem a partir de TC
Kay et al. (2017) <sup>676</sup>	DropConnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds	DL	Clasificación de ruidos cardiacos en normales o anormales
Kublanov et al. (2017) <sup>677</sup>	Comparison of Machine Learning Methods for the Arterial Hypertension Diagnostics	ML	Clasificación en hipertensión arterial o no a partir del análisis de la variabilidad de la frecuencia cardiaca
Attallah et al. (2017) <sup>678</sup>	Using multiple classifiers for predicting the risk of endovascular aortic aneurysm repair re-intervention through hybrid feature selection	ML, DL	Predicción del riesgo de reintervención en aneurismas de aorta tratados endovascularmente

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Liang et al. (2017) <sup>679</sup>	A deep learning approach to estimate chemically-treated collagenous tissue nonlinear anisotropic stress-strain responses from microscopy images	DL	Estimación de las propiedades elásticas del tejido colágeno directamente de imágenes de microscopía para la fabricación de bioprótesis cardíacas
Conant et al. (2017) <sup>680</sup>	Kinase inhibitor screening using artificial neural networks and engineered cardiac biowires	DL	Predicción de cardiotoxicidad de diferentes inhibidores de quinasas utilizados en cáncer
Vivekanandan et al. (2017) <sup>681</sup>	Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease	DL	Predicción de enfermedad cardíaca a partir de múltiples datos
Lee et al. (2017) <sup>682</sup>	Machine Learning of Human Pluripotent Stem Cell-Derived Engineered Cardiac Tissue Contractility for Automated Drug Classification	ML	Algoritmo de clasificación para predecir la acción mecanicista de una droga cardioactiva desconocida
Bergau et al. (2018) <sup>683</sup>	Machine-Learning Prediction of Drug-Induced Cardiac Arrhythmia: Analysis of Gene Expression and Clustering	ML	Sistema para evaluar la seguridad de medicamentos realizando análisis de expresión de genes
Amiriparian et al. (2018) <sup>684</sup>	Deep Unsupervised Representation Learning for Abnormal Heart Sound Classification	DL	Clasificador para identificar ruidos cardíacos anormales
Dey et al. (2018) <sup>685</sup>	InstaBP: Cuff-less Blood Pressure Monitoring on Smartphone using Single PPG Sensor	ML	Desarrollo de un modelo de predicción para la monitorización de la presión arterial por un solo sensor basado en particiones demográficas y fisiológicas
Thompson et al. (2018) <sup>686</sup>	Artificial Intelligence-Assisted Auscultation of Heart Murmurs: Validation by Virtual Clinical Trial	ML	Algoritmo clasificador de soplos cardíacos
Cai et al. (2018) <sup>687</sup>	In Silico Pharmacoepidemiologic Evaluation of Drug-Induced Cardiovascular Complications Using Combined Classifiers	ML	Desarrollo de modelos <i>in silico</i> para la identificación sistemática de complicaciones cardiovasculares inducidas por fármacos
Pan et al. (2018) <sup>688</sup>	Identifying Patients with Atrioventricular Septal Defect in Down Syndrome Populations by Using Self-Normalizing Neural Networks and Feature Selection	DL	Identificación de canal aurículo-ventricular en el Síndrome de Down
Elgendi et al. (2018) <sup>689</sup>	The Voice of the Heart: Vowel-Like Sound in Pulmonary Artery Hypertension	ML	Análisis de patrones de auscultación para el diagnóstico no invasivo de hipertensión pulmonar
Burghardt et al. (2018) <sup>690</sup>	Neural/Bayes network predictor for inheritable cardiac disease pathogenicity and phenotype	DL	Predictor de cardiopatías familiares y fenotipo a partir del análisis genético
Zhu et al. (2018) <sup>691</sup>	[Left ventricle segmentation in echocardiography based on adaptive mean shift]	ML	Segmentación automática del ventrículo izquierdo en ecocardiografía
Maharlou et al. (2018) <sup>692</sup>	Predicting Length of Stay in Intensive Care Units after Cardiac Surgery: Comparison of Artificial Neural Networks and Adaptive Neuro-fuzzy System	DL	Modelo para predecir estancias en unidades de críticos tras cirugía cardíaca

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Liem et al (2018) <sup>693</sup>	Phrase mining of textual data to analyze extracellular matrix protein patterns across cardiovascular disease	ML	Asociación entre los patrones de proteínas de matriz extracelular y enfermedad cardiovascular a partir de extracción de datos textuales en la literatura
Juhola et al. (2018) <sup>694</sup>	Detection of genetic cardiac diseases by Ca <sup>2+</sup> transient profiles using machine learning methods	ML	Identificación de anomalías en los canales de Ca <sup>2+</sup> para el diagnóstico de cardiopatías familiares
McCarthy et al. (2018) <sup>695</sup>	A clinical and proteomics approach to predict the presence of obstructive peripheral arterial disease: From the Catheter Sampled Blood Archive in Cardiovascular Diseases (CASABLANCA) Study	ML	Sistema para el diagnóstico de enfermedad vascular periférica
Hatib et al. (2018) <sup>696</sup>	Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis	ML	Análisis de las formas de onda de presión arterial para predecir hipotensión
Parreco et al. (2018) <sup>697</sup>	Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement	ML	Identificación de pacientes con ventilación prolongada y necesidad de traqueostomía
Giri et al. (2018) <sup>698</sup>	Increased Plasma Nitrite and von Willebrand Factor Indicates Early Diagnosis of Vascular Diseases in Chemotherapy Treated Cancer Patients	ML	Identificar biomarcadores en plasma de carditoxicidad
Kan et al. (2018) <sup>699</sup>	Handmade trileaflet valve design and validation for patch-valved conduit reconstruction using generalized regression machine learning model	DL	Determinación de los parámetros, incluidos el ancho, longitud y estructura curva, de conductos valvulados
Rösler et al. (2018) <sup>700</sup>	Development and Application of a System Based on Artificial Intelligence for Transcatheter Aortic Prosthesis Selection	ML	Sistema experto para la optimización en la selección de la TAVI a implantar
Luo et al. (2018) <sup>701</sup>	Integrating hypertension phenotype and genotype with hybrid non-negative matrix factorization	ML	Estratificación en pacientes hipertensos a partir de datos fenotípicos y genotípicos
Lee et al. (2018) <sup>103</sup>	Derivation and Validation of Machine Learning Approaches to Predict Acute Kidney Injury after Cardiac Surgery	ML	Predicción de fallo renal tras cirugía cardíaca
Luis Ahumadal et al. (2018) <sup>702</sup>	Prediction of One-Year Transplant-Free Survival after Norwood Procedure Based on the Pre-Operative Data	DL	Predicción de mala evolución en niños con ventrículo único hipoplásico sometidos a cirugía de Norwood
Lu et al. (2018) <sup>703</sup>	A Novel Deep Learning based Neural Network for Heartbeat Detection in Ballistocardiograph	DL	Detección automática de latidos cardíacos durante la balistocardiografía (registro de los movimientos producidos por el impacto de la sangre en el corazón y grandes vasos)
Chen et al. (2018) <sup>704</sup>	Artificial Neural Network: A Method for Prediction of Surgery-Related Pressure Injury in Cardiovascular Surgical Patients	DL	Predicción moderada de lesión por presión tras cirugía cardíaca
Aviles-Rivero et al. (2018) <sup>705</sup>	Sliding to predict: vision-based beating heart motion estimation by modeling temporal interactions	ML	Corrección de la movilidad cardíaca en la estimación estructural para emplear en cirugía robótica
Kang et al. (2018) <sup>706</sup>	Cardiac Auscultation Using Smartphones: Pilot Study	DL	Utilización de teléfonos móviles para auscultación cardíaca

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Llucia-Valldeperas et al. (2018) <sup>707</sup>	Unravelling the effects of mechanical physiological conditioning on cardiac adipose tissue-derived progenitor cells in vitro and in silico	DL	Análisis del proteoma de células cardiacas derivadas de tejido adiposo
Matam et al. (2018) <sup>708</sup>	Machine learning based framework to predict cardiac arrests in a paediatric intensive care unit : Prediction of cardiac arrests	ML	Patrones (frecuencias cardiaca y respiratoria, presión arterial, saturación) predictores de parada cardiaca en pacientes pediátricos ingresados en unidades de críticos
Jordanski et al. (2018) <sup>709</sup>	Machine Learning Approach for Predicting Wall Shear Distribution for Abdominal Aortic Aneurysm and Carotid Bifurcation Models	ML, DL	Algoritmo de cálculo de la distribución de la tensión de cizallamiento de la pared arterial, como mecanismo esencial en el desarrollo de aterosclerosis
Diaz et al. (2018) <sup>710</sup>	Modeling the control of the central nervous system over the cardiovascular system using support vector machines	ML	Modelos hemodinámicos de predicción del sistema cardiovascular en respuesta a estímulos del sistema nervioso central
Ferng et al. (2018) <sup>711</sup>	Study of continuous blood pressure estimation based on pulse transit time, heart rate and photoplethysmography-derived hemodynamic covariates	ML	Estimación continua de la presión arterial
Wacker et al. (2018) <sup>712</sup>	Performance of Machine Learning Algorithms for Qualitative and Quantitative Prediction Drug Blockade of hERG1 channel	ML	Modelo predictivo de cardiotoxicidad para evaluar fármacos con capacidad de bloquear el canal de potasio hERG
Siramshetty et al. (2018) <sup>713</sup>	The Catch-22 of Predicting hERG Blockade Using Publicly Accessible Bioactivity Data	ML	Clasificador de cardiotoxicidad por fármacos sobre el canal hERG con datos de bioactividad extraídos del dominio público
Lu et al. (2018) <sup>714</sup>	Research on Improved Depth Belief Network-Based Prediction of Cardiovascular Diseases	DL	Clasificador de patología cardiaca a partir de datos de pacientes ingresados en unidades de cuidados intensivos
Vervier et al. 2018 <sup>715</sup>	TiSAn: estimating tissue-specific effects of coding and non-coding variants	ML	Discriminación de las variantes relevantes para un tejido de aquellas que no influyen en su función
Bozkurt et al. (2018) <sup>716</sup>	A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection	DL	Clasificación de patología cardiaca en pacientes pediátricos por el análisis de ruidos cardiacos
Messner et al. (2018) <sup>717</sup>	Heart Sound Segmentation-An Event Detection Approach Using Deep Recurrent Neural Networks	DL	Clasificación de patología cardiaca a partir del análisis de primer y segundo ruidos cardiacos con datos procedentes del datatón 2016 PhysioNet/CinC Challenge
Mejia et al. (2018) <sup>718</sup>	Predictive performance of six mortality risk scores and the development of a novel model in a prospective cohort of patients undergoing valve surgery secondary to rheumatic fever	ML	Predicción de riesgo tras cirugía valvular en pacientes con cardiopatía reumática y comparación con scores de riesgo clásicos
Marateb et al. (2018) <sup>719</sup>	Prediction of dyslipidemia using gene mutations, family history of diseases and anthropometric indicators in children and adolescents: The CASPIAN-III study	ML	Diagnóstico de dislipemia usando datos genéticos, familiares y antropométricos en niños y adolescentes
Jalali et al. (2018) <sup>720</sup>	Prediction of Periventricular Leukomalacia in Neonates after Cardiac Surgery Using Machine Learning Algorithms	ML	Predicción de leucomalacia periventricular en neonatos sometidos a cirugía cardiaca

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Cai et al. (2019) <sup>721</sup>	Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity	DL	Predicción de cardiotoxicidad inducida por fármacos a través de deep learning
Diller et al. (2019) <sup>722</sup>	Utility of machine learning algorithms in assessing patients with a systemic right ventricle	ML	Evaluación de pacientes con ventrículo derecho sistémico a partir de imagen ecocardiográfica transtorácica
Harford et al. (2019) <sup>723</sup>	A machine learning based model for Out of Hospital cardiac arrest outcome classification and sensitivity analysis	ML	Predicción de paro cardíaco extrahospitalario con modelos machine learning
Blomberg et al. (2019) <sup>724</sup>	Machine learning as a supportive tool to recognize cardiac arrest in emergency calls	ML	Reconocimiento de pacientes en parada cardíaca en llamadas telefónicas de emergencia a través
Melero-Alegria et al. (2019) <sup>9</sup>	SALMANTICOR study. Rationale and design of a population-based study to identify structural heart disease abnormalities: a spatial and machine learning analysis	ML	Diseño de estudio poblacional para la identificación de enfermedades cardíacas estructurales.
Casaclang-Verzosa et al. (2019) <sup>725</sup>	Network Tomography for Understanding Phenotypic Presentations in Aortic Stenosis	ML	Fenotipado de ventrículos izquierdos en respuesta a la aparición de estenosis aórtica
Bello et al. (2019) <sup>726</sup>	Deep learning cardiac motion analysis for human survival prediction	DL	Análisis del movimiento cardíaco para la predicción de supervivencia
Huo et al. (2019) <sup>727</sup>	A machine learning model to classify aortic dissection patients in the early diagnosis phase	ML	Clasificación de pacientes con riesgo de disección de aorta en fases de diagnóstico
Diller et al. (2019) <sup>728</sup>	Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients	ML	Evaluación pronóstica y terapéutica en pacientes adultos con patologías cardíacas congénitas
Zhao et al. (2019) <sup>729</sup>	Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction	DL	Predicción de eventos cardiovasculares a partir del análisis de registros clínicos electrónicos y datos genéticos
Daghistani et al. (2019) <sup>730</sup>	Predictors of in-hospital length of stay among cardiac patients: A machine learning approach	DL	Identificación de predictores de tiempos de estancia intrahospitalaria de pacientes cardíacos
Sparapani et al. (2019) <sup>731</sup>	Detection of Left Ventricular Hypertrophy Using Bayesian Additive Regression Trees: The MESA	ML	Detección de hipertrofia del ventrículo izquierdo
Pulido et al. (2019) <sup>732</sup>	Blood Pressure Classification Using the Method of the Modular Neural Networks	ML	Clasificación de presiones arteriales
Ibrahim et al. (2019) <sup>733</sup>	A clinical, proteomics, and artificial intelligence-driven model to predict acute kidney injury in patients undergoing coronary angiography	ML	Predicción de daño renal agudo en pacientes sometidos a intervención coronaria percutánea
va et al. (2019) <sup>734</sup>	Prognostic value of CT myocardial perfusion imaging and CT-derived fractional flow reserve for major adverse cardiac events in patients with coronary artery disease	ML	Evaluación de valores pronósticos de imagen de perfusión miocárdica por tomografía y reserva fraccional de flujo a partir de tomografía para la predicción de eventos cardíacos mayores en pacientes con enfermedad coronaria

Anexo 5. Análisis de publicaciones sobre inteligencia artificial en cardiología (fuente PubMed)

Ruiz et al. (2019) <sup>735</sup>	Early prediction of critical events for infants with single-ventricle physiology in critical care using routinely collected data	ML	Predicción temprana de eventos críticos en neonatos con ventrículo único
Ris et al. (2019) <sup>736</sup>	Inflammatory biomarkers in infective endocarditis: machine learning to predict mortality	ML	Identificación de biomarcadores de endocarditis
Hsich et al. (2019) <sup>737</sup>	Variables of importance in the Scientific Registry of Transplant Recipients database predictive of heart transplant waitlist mortality	ML	Identificación de características para la predicción de mortalidad en listas de espera a trasplante cardíaco
Emre et al. (2019) <sup>738</sup>	The analysis of the effects of acute rheumatic fever in childhood on cardiac disease with data mining	ML	Análisis de efectos de la fiebre reumática aguda en pacientes infantiles con patologías cardíacas
Sun et al. (2019) <sup>739</sup>	[Radiomics strategy based on cardiac magnetic resonance imaging cine sequence for assessing the severity of mitral value regurgitation]	ML	Evaluación de severidad de insuficiencias mitrales a partir de resonancia magnética cardíaca

## Anexo 5 - Bibliografía (673 referencias)

1. Dilsizian SE, Siegel EL. Artificial intelligence in medicine and cardiac imaging: harnessing big data and advanced computing to provide personalized medical diagnosis and treatment. *Curr Cardiol Rep.* 2014;16(1):441.
2. Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development.* 1959;3(3):210-229.
3. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature.* 2015;521(7553):436-444.
4. Poplin R, Varadarajan AV, Blumer K, et al. Prediction of cardiovascular risk factors from retinal fundus photographs via deep learning. *Nature Biomedical Engineering.* 2018;2:158-164.
5. Medrano IH, Guijarro JT, Belda C, et al. Savana: Re-using Electronic Health Records with Artificial Intelligence. *International Journal of Interactive Multimedia and Artificial Intelligence.* 2018;4(7):8-12.
6. González-Mansilla A, Martínez-Legazpi P, Prieto A, et al. Valve area and the risk of overestimating aortic stenosis. *Heart.* 2019;105(12):911-919.
7. Yoo JS, Kim JB, Jung SH, et al. Mitral durability after robotic mitral valve repair: analysis of 200 consecutive mitral regurgitation repairs. *J Thorac Cardiovasc Surg.* 2014;148(6):2773-2779.
8. Miller RA, Pople HE, Jr., Myers JD. Internist-1, an experimental computer-based diagnostic consultant for general internal medicine. *N Engl J Med.* 1982;307(8):468-476.
9. Melero-Alegria JI, Cascon M, Romero A, et al. SALMANTICOR study. Rationale and design of a population-based study to identify structural heart disease abnormalities: a spatial and machine learning analysis. *BMJ open.* 2019;9(2):e024605.
10. Rajkomar A, Dean J, Kohane I. Machine Learning in Medicine. *N Engl J Med.* 2019;380(14):1347-1358.
11. Murphy KP. Machine learning: a probabilistic perspective. In: MIT press; 2012.
12. Barreiro-Pérez M, Tundidor-Sanz E, Martín-García A, et al. First Magnetic Resonance Managed by a Cardiology Department in the Spanish Public Healthcare System. Experience and Difficulties of an Innovative Model. *Rev Esp Cardiol (Engl Ed).* 2018;71(5):365-372.
13. Baumgartner CF KL, Pollefeys M, Konukoglu E. An exploration of 2 D and 3 D deep learning techniques for cardiac MR image segmentation. Disponible en: <https://arxiv.org/abs/1709.04496>. In.
14. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Med Image Anal.* 2017;42:60-88.
15. Char DS, Shah NH, Magnus D. Implementing Machine Learning in Health Care - Addressing Ethical Challenges. *N Engl J Med.* 2018;378(11):981-983.
16. Gijsberts CM, Groenewegen KA, Hofer IE, et al. Race/Ethnic Differences in the Associations of the Framingham Risk Factors with Carotid IMT and Cardiovascular Events. *PLoS One.* 2015;10(7):e0132321.
17. Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: Addressing ethical challenges. *PLoS Med.* 2018;15(11):e1002689.
18. Sturm I, Lapuschkin S, Samek W, Muller KR. Interpretable deep neural networks for single-trial EEG classification. *J Neurosci Methods.* 2016;274:141-145.
19. Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. Boletín Oficial del Estado. 2018. Disponible en: <https://www.boe.es/eli/es/lo/2018/12/05/3/con>. In.
20. Mesko B, Hetenyi G, Gyorffy Z. Will artificial intelligence solve the human resource crisis in healthcare? *BMC Health Serv Res.* 2018;18(1):545.
21. Sanchez PL, Gimeno F, Ancillo P, et al. Role of the paclitaxel-eluting stent and tirofiban in patients with ST-elevation myocardial infarction undergoing

- postfibrinolysis angioplasty: the GRACIA-3 randomized clinical trial. *Circ Cardiovasc Interv.* 2010;3(4):297-307.
22. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research.* 2011;12:2825-2830.
  23. <https://eli5.readthedocs.io/en/latest/index.html>.
  24. Rodriguez JD, Perez A, Lozano JA. Sensitivity analysis of kappa-fold cross validation in prediction error estimation. *IEEE Trans Pattern Anal Mach Intell.* 2010;32(3):569-575.
  25. T. Hastie RT, J. Friedman. The Elements of Statistical Learning. In: Springer; 2009.
  26. Kuntz RE, Baim DS. DEFINING CORONARY RESTENOSIS - NEWER CLINICAL AND ANGIOGRAPHIC PARADIGMS. *Circulation.* 1993;88(3):1310-1323.
  27. Dorado-Diaz PI, Sampedro-Gomez J, Vicente-Palacios V, Sanchez PL. Applications of artificial intelligence in cardiology: the future is already here. *Rev Esp Cardiol.* 2019.
  28. Breiman L. Random forests. *Machine Learning.* 2001;45(1):5-32.
  29. L. Breiman AC, Annals of Statistics 1998. In.
  30. Varma S, Simon R. Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics.* 2006;7:91.
  31. Bergstra J, Bengio Y. Random search for hyper-parameter optimization. *J Mach Learn Res.* 2012;13:281-305.
  32. Singh M, Gersh BJ, McClelland RL, et al. Clinical and angiographic predictors of restenosis after percutaneous coronary intervention: insights from the Prevention of Restenosis With Tranilast and Its Outcomes (PRESTO) trial. *Circulation.* 2004;109(22):2727-2731.
  33. Stolker JM, Kennedy KF, Lindsey JB, et al. Predicting restenosis of drug-eluting stents placed in real-world clinical practice: derivation and validation of a risk model from the EVENT registry. *Circ Cardiovasc Interv.* 2010;3(4):327-334.
  34. Moons KG, Altman DG, Reitsma JB, et al. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. *Ann Intern Med.* 2015;162(1):W1-73.
  35. Airola A, Pahikkala T, Waegeman W, De Baets B, Salakoski T. An experimental comparison of cross-validation techniques for estimating the area under the ROC curve. *Computational Statistics & Data Analysis.* 2011;55(4):1828-1844.
  36. Bouckaert RR. Choosing between two learning algorithms based on calibrated tests. *Proceedings of the 20th International Conference on Machine Learning (ICML-03).* 2003:51-58.
  37. Saito T, Rehmsmeier M. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS One.* 2015;10(3):e0118432.
  38. Davis J, Goadrich M. The relationship between precision-recall and ROC curves. *Proceedings of the 23rd International Conference on Machine Learning (ICML-06).* 2006:233-240.
  39. Revascularization SWGfEEGfM, Revascularization ERfEEGfM, Committee SG. Comments on the 2018 ESC/EACTS Guidelines for Myocardial Revascularization. *Rev Esp Cardiol (Engl Ed).* 2019;72(1):16-20.
  40. Laham RJ, Carrozza JP, Berger C, Cohen DJ, Kuntz RE, Baim DS. Long-term (4- to 6-year) outcome of Palmaz-Schatz stenting: paucity of late clinical stent-related problems. *J Am Coll Cardiol.* 1996;28(4):820-826.
  41. Torrado J, Buckley L, Durán A, et al. Restenosis, Stent Thrombosis, and Bleeding Complications: Navigating Between Scylla and Charybdis. *J Am Coll Cardiol.* 2018;71(15):1676-1695.
  42. Califf RM, Fortin DF, Frid DJ, et al. Restenosis after coronary angioplasty: an overview. *J Am Coll Cardiol.* 1991;17(6 Suppl B):2B-13B.



43. Byrne RA, Joner M, Kastrati A. Stent thrombosis and restenosis: what have we learned and where are we going? The Andreas Gruntzig Lecture ESC 2014. *Eur Heart J*. 2015;36(47):3320-3331.
44. Chen MS, John JM, Chew DP, Lee DS, Ellis SG, Bhatt DL. Bare metal stent restenosis is not a benign clinical entity. *Am Heart J*. 2006;151(6):1260-1264.
45. Dangas GD, Claessen BE, Caixeta A, Sanidas EA, Mintz GS, Mehran R. In-stent restenosis in the drug-eluting stent era. *J Am Coll Cardiol*. 2010;56(23):1897-1907.
46. Alfonso F, Cuesta J. The Therapeutic Dilemma of Recurrent In-Stent Restenosis. *Circ Cardiovasc Interv*. 2018;11(8):e007109.
47. Kokkinidis DG, Waldo SW, Armstrong EJ. Treatment of coronary artery in-stent restenosis. *Expert Rev Cardiovasc Ther*. 2017;15(3):191-202.
48. Kasaoka S, Tobis JM, Akiyama T, et al. Angiographic and intravascular ultrasound predictors of in-stent restenosis. *J Am Coll Cardiol*. 1998;32(6):1630-1635.
49. Bauters C, Hubert E, Prat A, et al. Predictors of restenosis after coronary stent implantation. *J Am Coll Cardiol*. 1998;31(6):1291-1298.
50. Kastrati A, Schömig A, Elezi S, et al. Predictive factors of restenosis after coronary stent placement. *J Am Coll Cardiol*. 1997;30(6):1428-1436.
51. Hoffmann R, Mintz GS, Mehran R, et al. Intravascular ultrasound predictors of angiographic restenosis in lesions treated with Palmaz-Schatz stents. *J Am Coll Cardiol*. 1998;31(1):43-49.
52. Stolker JM, Kennedy KF, Lindsey JB, et al. Predicting Restenosis of Drug-Eluting Stents Placed in Real-World Clinical Practice. *Circulation: Cardiovascular Interventions*. 2010;3(4):327-334.
53. Feinberg J, Nielsen EE, Greenhalgh J, et al. Drug-eluting stents versus bare-metal stents for acute coronary syndrome. *Cochrane Database Syst Rev*. 2017;8:CD012481.
54. Mindrescu C, Brener SJ, Guerchicoff A, et al. Impact of scheduled angiographic follow-up in patients treated with primary percutaneous coronary intervention for ST-segment elevation myocardial infarction. *J Interv Cardiol*. 2013;26(4):319-324.
55. Westcott RJ, Tcheng JE. Artificial Intelligence and Machine Learning in Cardiology. *JACC Cardiovasc Interv*. 2019;12(14):1312-1314.
56. Cui S, Li K, Ang L, et al. Plasma Phospholipids and Sphingolipids Identify Stent Restenosis After Percutaneous Coronary Intervention. *JACC Cardiovasc Interv*. 2017;10(13):1307-1316.
57. Al'Aref SJ AK, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *Eur Heart J*. 2019, 40:1975-1986. In.
58. de la Torre Hernández JM, Edelman ER. From Nonclinical Research to Clinical Trials and Patient-registries: Challenges and Opportunities in Biomedical Research. *Rev Esp Cardiol (Engl Ed)*. 2017;70(12):1121-1133.
59. Ebrahimzadeh E, Kalantari M, Joulani M, Shahraki RS, Fayaz F, Ahmadi F. Prediction of paroxysmal Atrial Fibrillation: A machine learning based approach using combined feature vector and mixture of expert classification on HRV signal. *Comput Methods Programs Biomed*. 2018;165:53-67.
60. Budzianowski J, Hiczkievicz J, Burchardt P, et al. Predictors of atrial fibrillation early recurrence following cryoballoon ablation of pulmonary veins using statistical assessment and machine learning algorithms. *Heart Vessels*. 2018.
61. Eerikainen LM, Vanschoren J, Rooijackers MJ, Vullings R, Aarts RM. Reduction of false arrhythmia alarms using signal selection and machine learning. *Physiol Meas*. 2016;37(8):1204-1216.

62. Nanayakkara S, Fogarty S, Tremeer M, et al. Characterising risk of in-hospital mortality following cardiac arrest using machine learning: A retrospective international registry study. *PLoS Med.* 2018;15(11):e1002709.
63. Yildirim O, Plawiak P, Tan RS, Acharya UR. Arrhythmia detection using deep convolutional neural network with long duration ECG signals. *Comput Biol Med.* 2018;102:411-420.
64. Weng SF, Reys J, Kai J, Garibaldi JM, Qureshi N. Can machine-learning improve cardiovascular risk prediction using routine clinical data? *PLoS One.* 2017;12(4):e0174944.
65. Huang Z, Chan TM, Dong W. MACE prediction of acute coronary syndrome via boosted resampling classification using electronic medical records. *J Biomed Inform.* 2017;66:161-170.
66. Shouval R, Hadanny A, Shlomo N, et al. Machine learning for prediction of 30-day mortality after ST elevation myocardial infarction: An Acute Coronary Syndrome Israeli Survey data mining study. *Int J Cardiol.* 2017;246:7-13.
67. Wallert J, Tomasoni M, Madison G, Held C. Predicting two-year survival versus non-survival after first myocardial infarction using machine learning and Swedish national register data. *BMC Med Inform Decis Mak.* 2017;17(1):99.
68. Shameer K, Johnson KW, Yahya A, et al. Predictive Modeling of Hospital Readmission Rates Using Electronic Medical Record-Wide Machine Learning: A Case-Study Using Mount Sinai Heart Failure Cohort. *Pac Symp Biocomput.* 2017;22:276-287.
69. Inan OT, Baran Pouyan M, Javaid AQ, et al. Novel Wearable Seismocardiography and Machine Learning Algorithms Can Assess Clinical Status of Heart Failure Patients. *Circ Heart Fail.* 2018;11(1):e004313.
70. Medved D, Ohlsson M, Hoglund P, Andersson B, Nugues P, Nilsson J. Improving prediction of heart transplantation outcome using deep learning techniques. *Sci Rep.* 2018;8(1):3613.
71. Sanchez-Martinez S, Duchateau N, Erdei T, et al. Machine Learning Analysis of Left Ventricular Function to Characterize Heart Failure With Preserved Ejection Fraction. *Circ Cardiovasc Imaging.* 2018;11(4):e007138.
72. Ahmad T, Lund LH, Rao P, et al. Machine Learning Methods Improve Prognostication, Identify Clinically Distinct Phenotypes, and Detect Heterogeneity in Response to Therapy in a Large Cohort of Heart Failure Patients. *J Am Heart Assoc.* 2018;7(8).
73. Samad MD, Ulloa A, Wehner GJ, et al. Predicting Survival From Large Echocardiography and Electronic Health Record Datasets: Optimization With Machine Learning. *JACC Cardiovasc Imaging.* 2018.
74. Narula S, Shameer K, Salem Omar AM, Dudley JT, Sengupta PP. Machine-Learning Algorithms to Automate Morphological and Functional Assessments in 2D Echocardiography. *J Am Coll Cardiol.* 2016;68(21):2287-2295.
75. Knackstedt C, Bekkers SC, Schummers G, et al. Fully Automated Versus Standard Tracking of Left Ventricular Ejection Fraction and Longitudinal Strain: The FAST-EFs Multicenter Study. *J Am Coll Cardiol.* 2015;66(13):1456-1466.
76. Tesche C, De Cecco CN, Baumann S, et al. Coronary CT Angiography-derived Fractional Flow Reserve: Machine Learning Algorithm versus Computational Fluid Dynamics Modeling. *Radiology.* 2018;288(1):64-72.
77. Sengupta PP, Huang YM, Bansal M, et al. Cognitive Machine-Learning Algorithm for Cardiac Imaging: A Pilot Study for Differentiating Constrictive Pericarditis From Restrictive Cardiomyopathy. *Circ Cardiovasc Imaging.* 2016;9(6).
78. Bernard O, Lalande A, Zotti C, et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? *IEEE Trans Med Imaging.* 2018;37(11):2514-2525.

79. Xiong Z, Liu T, Tse G, et al. A Machine Learning Aided Systematic Review and Meta-Analysis of the Relative Risk of Atrial Fibrillation in Patients With Diabetes Mellitus. *Front Physiol.* 2018;9:835.
80. Li Q, Rajagopalan C, Clifford GD. Ventricular fibrillation and tachycardia classification using a machine learning approach. *IEEE Trans Biomed Eng.* 2014;61(6):1607-1613.
81. Figuera C, Irusta U, Morgado E, et al. Machine Learning Techniques for the Detection of Shockable Rhythms in Automated External Defibrillators. *PLoS One.* 2016;11(7):e0159654.
82. Lyon A, Ariga R, Mincholé A, et al. Distinct ECG Phenotypes Identified in Hypertrophic Cardiomyopathy Using Machine Learning Associate With Arrhythmic Risk Markers. *Front Physiol.* 2018;9:213.
83. Betancur J, Otaki Y, Motwani M, et al. Prognostic Value of Combined Clinical and Myocardial Perfusion Imaging Data Using Machine Learning. *JACC Cardiovasc Imaging.* 2018;11(7):1000-1009.
84. Mortazavi BJ, Downing NS, Bucholz EM, et al. Analysis of Machine Learning Techniques for Heart Failure Readmissions. *Circ Cardiovasc Qual Outcomes.* 2016;9(6):629-640.
85. Frizzell JD, Liang L, Schulte PJ, et al. Prediction of 30-Day All-Cause Readmissions in Patients Hospitalized for Heart Failure: Comparison of Machine Learning and Other Statistical Approaches. *JAMA Cardiol.* 2017;2(2):204-209.
86. Johnson KW, Torres Soto J, Glicksberg BS, et al. Artificial Intelligence in Cardiology. *J Am Coll Cardiol.* 2018;71(23):2668-2679.
87. Fallahzadeh R, Pedram M, Ghasemzadeh H. SmartSock: a wearable platform for context-aware assessment of ankle edema. *Conf Proc IEEE Eng Med Biol Soc.* 2016;2016:6302-6306.
88. Kalscheur MM, Kipp RT, Tattersall MC, et al. Machine Learning Algorithm Predicts Cardiac Resynchronization Therapy Outcomes: Lessons From the COMPANION Trial. *Circ Arrhythm Electrophysiol.* 2018;11(1):e005499.
89. Salem Omar AM, Shameer K, Narula S, et al. Artificial Intelligence-Based Assessment of Left Ventricular Filling Pressures From 2-Dimensional Cardiac Ultrasound Images. *JACC Cardiovasc Imaging.* 2018;11(3):509-510.
90. Lancaster MC, Salem Omar AM, Narula S, Kulkarni H, Narula J, Sengupta PP. Phenotypic Clustering of Left Ventricular Diastolic Function Parameters: Patterns and Prognostic Relevance. *JACC Cardiovasc Imaging.* 2018.
91. Carneiro G, Nascimento JC. Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data. *IEEE Trans Pattern Anal Mach Intell.* 2013;35(11):2592-2607.
92. Slomka PJ, Dey D, Sitek A, Motwani M, Berman DS, Germano G. Cardiac imaging: working towards fully-automated machine analysis & interpretation. *Expert Rev Med Devices.* 2017;14(3):197-212.
93. Avendi MR, Kheradvar A, Jafarkhani H. A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI. *Med Image Anal.* 2016;30:108-119.
94. Ngo TA, Lu Z, Carneiro G. Combining deep learning and level set for the automated segmentation of the left ventricle of the heart from cardiac cine magnetic resonance. *Med Image Anal.* 2017;35:159-171.
95. Dou Q, Yu L, Chen H, et al. 3D deeply supervised network for automated segmentation of volumetric medical images. *Med Image Anal.* 2017;41:40-54.
96. Zheng Q, Delingette H, Duchateau N, Ayache N. 3-D Consistent and Robust Segmentation of Cardiac Images by Deep Learning With Spatial Propagation. *IEEE Trans Med Imaging.* 2018;37(9):2137-2148.
97. Zhang J, Gajjala S, Agrawal P, et al. Fully Automated Echocardiogram Interpretation in Clinical Practice. *Circulation.* 2018;138(16):1623-1635.

98. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Isgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. *Med Image Anal.* 2016;34:123-136.
99. Abdolmanafi A, Duong L, Dahdah N, Cheriet F. Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography. *Biomed Opt Express.* 2017;8(2):1203-1220.
100. Lekadir K, Galimzianova A, Betriu A, et al. A Convolutional Neural Network for Automatic Characterization of Plaque Composition in Carotid Ultrasound. *IEEE J Biomed Health Inform.* 2017;21(1):48-55.
101. Schlemper J, Caballero J, Hajnal JV, Price AN, Rueckert D. A Deep Cascade of Convolutional Neural Networks for Dynamic MR Image Reconstruction. *IEEE Trans Med Imaging.* 2018;37(2):491-503.
102. Zreik M, Lessmann N, van Hamersvelt RW, et al. Deep learning analysis of the myocardium in coronary CT angiography for identification of patients with functionally significant coronary artery stenosis. *Med Image Anal.* 2018;44:72-85.
103. Lee HC, Yoon HK, Nam K, et al. Derivation and Validation of Machine Learning Approaches to Predict Acute Kidney Injury after Cardiac Surgery. *J Clin Med.* 2018;7(10).
104. Das MK, Ari S. ECG Beats Classification Using Mixture of Features. *Int Sch Res Notices.* 2014;2014:178436.
105. Mahajan D, Dong Y, Saxon LA, et al. Performance of an automatic arrhythmia classification algorithm: comparison to the ALTITUDE electrophysiologist panel adjudications. *Pacing Clin Electrophysiol.* 2014;37(7):889-899.
106. Caravaca J, Soria-Olivas E, Bataller M, et al. Application of machine learning techniques to analyse the effects of physical exercise in ventricular fibrillation. *Comput Biol Med.* 2014;45:1-7.
107. Daqrouq K, Alkhateeb A, Ajour MN, Morfeq A. Neural network and wavelet average framing percentage energy for atrial fibrillation classification. *Comput Methods Programs Biomed.* 2014;113(3):919-926.
108. Ebrahimzadeh E, Pooyan M, Bijar A. A novel approach to predict sudden cardiac death (SCD) using nonlinear and time-frequency analyses from HRV signals. *PLoS One.* 2014;9(2):e81896.
109. Behar J, Johnson A, Clifford GD, Oster J. A comparison of single channel fetal ECG extraction methods. *Ann Biomed Eng.* 2014;42(6):1340-1353.
110. Li Q, Rajagopalan C, Clifford GD. A machine learning approach to multi-level ECG signal quality classification. *Comput Methods Programs Biomed.* 2014;117(3):435-447.
111. Sunkaria RK, Kumar V, Saxena SC, Singhal AM. An ANN-based HRV classifier for cardiac health prognosis. *Int J Electron Healthc.* 2014;7(4):315-330.
112. Prakosa A, Sermesant M, Allain P, et al. Cardiac electrophysiological activation pattern estimation from images using a patient-specific database of synthetic image sequences. *IEEE Trans Biomed Eng.* 2014;61(2):235-245.
113. Beudel M, Tjepkema-Cloostermans MC, Boersma JH, van Putten MJ. Small-world characteristics of EEG patterns in post-anoxic encephalopathy. *Front Neurol.* 2014;5:97.
114. Nguyen LL, Su S, Nguyen HT. Neural network approach for non-invasive detection of hyperglycemia using electrocardiographic signals. *Conf Proc IEEE Eng Med Biol Soc.* 2014;2014:4475-4478.
115. Wander JD, Morris D. A combined segmenting and non-segmenting approach to signal quality estimation for ambulatory photoplethysmography. *Physiol Meas.* 2014;35(12):2543-2561.
116. Juhola M, Joutsijoki H, Varpa K, et al. On computation of calcium cycling anomalies in cardiomyocytes data. *Conf Proc IEEE Eng Med Biol Soc.* 2014;2014:1444-1447.

117. Kennedy CE, Aoki N, Mariscalco M, Turley JP. Using Time Series Analysis to Predict Cardiac Arrest in a PICU. *Pediatr Crit Care Med*. 2015;16(9):e332-339.
118. Kora P, Kalva SR. Improved Bat algorithm for the detection of myocardial infarction. *Springerplus*. 2015;4:666.
119. Naseri H, Homaeinezhad MR. Electrocardiogram signal quality assessment using an artificially reconstructed target lead. *Comput Methods Biomech Biomed Engin*. 2015;18(10):1126-1141.
120. Attin M, Feld G, Lemus H, et al. Electrocardiogram characteristics prior to in-hospital cardiac arrest. *J Clin Monit Comput*. 2015;29(3):385-392.
121. Zhang L, Guo T, Xi B, et al. Automatic recognition of cardiac arrhythmias based on the geometric patterns of Poincare plots. *Physiol Meas*. 2015;36(2):283-301.
122. He M, Gong Y, Li Y, et al. Combining multiple ECG features does not improve prediction of defibrillation outcome compared to single features in a large population of out-of-hospital cardiac arrests. *Crit Care*. 2015;19:425.
123. Lillo-Castellano JM, Mora-Jimenez I, Santiago-Mozos R, et al. Symmetrical compression distance for arrhythmia discrimination in cloud-based big-data services. *IEEE J Biomed Health Inform*. 2015;19(4):1253-1263.
124. Tanantong T, Nantajeewarawat E, Thiemjarus S. False alarm reduction in BSN-based cardiac monitoring using signal quality and activity type information. *Sensors (Basel)*. 2015;15(2):3952-3974.
125. Ravens U, Katircioglu-Ozturk D, Wettwer E, et al. Application of the RIMARC algorithm to a large data set of action potentials and clinical parameters for risk prediction of atrial fibrillation. *Med Biol Eng Comput*. 2015;53(3):263-273.
126. Jiang M, Zhang H, Zhu L, et al. Noninvasive reconstruction of cardiac transmembrane potentials using a kernelized extreme learning method. *Phys Med Biol*. 2015;60(8):3237-3253.
127. Kiranyaz S, Ince T, Gabbouj M. Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks. *IEEE Trans Biomed Eng*. 2016;63(3):664-675.
128. Rosier A, Mabo P, Temal L, et al. Personalized and automated remote monitoring of atrial fibrillation. *Europace*. 2016;18(3):347-352.
129. Rad AB, Engan K, Katsaggelos AK, et al. Automatic cardiac rhythm interpretation during resuscitation. *Resuscitation*. 2016;102:44-50.
130. Shandilya S, Kurz MC, Ward KR, Najarian K. Integration of Attributes from Non-Linear Characterization of Cardiovascular Time-Series for Prediction of Defibrillation Outcomes. *PLoS One*. 2016;11(1):e0141313.
131. Yu M, Zhang G, Wu T, et al. A new method without reference channels used for ventricular fibrillation detection during cardiopulmonary resuscitation. *Australas Phys Eng Sci Med*. 2016;39(2):391-401.
132. Mert A. ECG feature extraction based on the bandwidth properties of variational mode decomposition. *Physiol Meas*. 2016;37(4):530-543.
133. Elhaj FA, Salim N, Harris AR, Swee TT, Ahmed T. Arrhythmia recognition and classification using combined linear and nonlinear features of ECG signals. *Comput Methods Programs Biomed*. 2016;127:52-63.
134. Zahid S, Cochet H, Boyle PM, et al. Patient-derived models link re-entrant driver localization in atrial fibrillation to fibrosis spatial pattern. *Cardiovasc Res*. 2016;110(3):443-454.
135. Pant JK, Krishnan S. Efficient compressive sensing of ECG segments based on machine learning for QRS-based arrhythmia detection. *Conf Proc IEEE Eng Med Biol Soc*. 2016;2016:4731-4734.
136. Ansari S, Belle A, Ghanbari H, Salamango M, Najarian K. Suppression of false arrhythmia alarms in the ICU: a machine learning approach. *Physiol Meas*. 2016;37(8):1186-1203.
137. Immanuel SA, Sadrieh A, Baumert M, et al. T-wave morphology can distinguish healthy controls from LQTS patients. *Physiol Meas*. 2016;37(9):1456-1473.

138. Li X, Liu H, Du X, et al. Integrated Machine Learning Approaches for Predicting Ischemic Stroke and Thromboembolism in Atrial Fibrillation. *AMIA Annu Symp Proc.* 2016;2016:799-807.
139. Nemati S, Ghassemi MM, Ambai V, et al. Monitoring and detecting atrial fibrillation using wearable technology. *Conf Proc IEEE Eng Med Biol Soc.* 2016;2016:3394-3397.
140. Xiong P, Wang H, Liu M, Lin F, Hou Z, Liu X. A stacked contractive denoising auto-encoder for ECG signal denoising. *Physiol Meas.* 2016;37(12):2214-2230.
141. Yanting S, Yang Y, Parish S, Zhengming C, Clarke R, Clifton DA. Risk prediction for cardiovascular disease using ECG data in the China kadoorie biobank. *Conf Proc IEEE Eng Med Biol Soc.* 2016;2016:2419-2422.
142. Zhang C, Zhou Y, Gu S, et al. In silico prediction of hERG potassium channel blockage by chemical category approaches. *Toxicol Res (Camb).* 2016;5(2):570-582.
143. He M, Lu Y, Zhang L, Zhang H, Gong Y, Li Y. Combining Amplitude Spectrum Area with Previous Shock Information Using Neural Networks Improves Prediction Performance of Defibrillation Outcome for Subsequent Shocks in Out-Of-Hospital Cardiac Arrest Patients. *PLoS One.* 2016;11(2):e0149115.
144. Lancaster MC, Sobie EA. Improved Prediction of Drug-Induced Torsades de Pointes Through Simulations of Dynamics and Machine Learning Algorithms. *Clin Pharmacol Ther.* 2016;100(4):371-379.
145. Khan MS, Yousuf S. A cardiac electrical activity model based on a cellular automata system in comparison with neural network model. *Pak J Pharm Sci.* 2016;29(2):579-584.
146. Alqaraawi A, Alwosheel A, Alasaad A. Heart rate variability estimation in photoplethysmography signals using Bayesian learning approach. *Health Technol Lett.* 2016;3(2):136-142.
147. Kalidas V, Tamil LS. Cardiac arrhythmia classification using multi-modal signal analysis. *Physiol Meas.* 2016;37(8):1253-1272.
148. Lee H, Shin SY, Seo M, Nam GB, Joo S. Prediction of Ventricular Tachycardia One Hour before Occurrence Using Artificial Neural Networks. *Sci Rep.* 2016;6:32390.
149. Abdul-Kadir NA, Mat Safri N, Othman MA. Dynamic ECG features for atrial fibrillation recognition. *Comput Methods Programs Biomed.* 2016;136:143-150.
150. Erem B, Martinez Orellana R, Hyde DE, et al. Extensions to a manifold learning framework for time-series analysis on dynamic manifolds in bioelectric signals. *Phys Rev E.* 2016;93(4):042218.
151. Kurzendorfer T, Mewes PW, Maier A, Strobel N, Brost A. Cryo-Balloon Catheter Localization Based on a Support-Vector-Machine Approach. *IEEE Trans Med Imaging.* 2016;35(8):1892-1902.
152. Evangelista LS, Ghasemzadeh H, Lee JA, Fallahzadeh R, Sarrafzadeh M, Moser DK. Predicting adherence to use of remote health monitoring systems in a cohort of patients with chronic heart failure. *Technol Health Care.* 2017;25(3):425-433.
153. Giffard-Roisin S, Jackson T, Fovargue L, et al. Noninvasive Personalization of a Cardiac Electrophysiology Model From Body Surface Potential Mapping. *IEEE Trans Biomed Eng.* 2017;64(9):2206-2218.
154. Li H, Yuan D, Ma X, Cui D, Cao L. Genetic algorithm for the optimization of features and neural networks in ECG signals classification. *Sci Rep.* 2017;7:41011.
155. Mjihad A, Rosado-Munoz A, Bataller-Mompean M, Frances-Villora JV, Guerrero-Martinez JF. Ventricular Fibrillation and Tachycardia detection from surface ECG using time-frequency representation images as input dataset for machine learning. *Comput Methods Programs Biomed.* 2017;141:119-127.

156. Okutucu S, Katircioglu-Ozturk D, Oto E, et al. Data mining experiments on the Angiotensin II-Antagonist in Paroxysmal Atrial Fibrillation (ANTIPAF-AFNET 2) trial: 'exposing the invisible'. *Europace*. 2017;19(5):741-746.
157. Acharya UR, Oh SL, Hagiwara Y, et al. A deep convolutional neural network model to classify heartbeats. *Comput Biol Med*. 2017;89:389-396.
158. Arvanaghi R, Daneshvar S, Seyedarabi H, Goshvarpour A. Fusion of ECG and ABP signals based on wavelet transform for cardiac arrhythmias classification. *Comput Methods Programs Biomed*. 2017;151:71-78.
159. Tripathy RK, Dandapat S. Automated detection of heart ailments from 12-lead ECG using complex wavelet sub-band bi-spectrum features. *Healthc Technol Lett*. 2017;4(2):57-63.
160. Luo K, Li J, Wang Z, Cuschieri A. Patient-Specific Deep Architectural Model for ECG Classification. *J Healthc Eng*. 2017;2017:4108720.
161. Li P, Wang Y, He J, et al. High-Performance Personalized Heartbeat Classification Model for Long-Term ECG Signal. *IEEE Trans Biomed Eng*. 2017;64(1):78-86.
162. Pandit D, Zhang L, Liu C, Chattopadhyay S, Aslam N, Lim CP. A lightweight QRS detector for single lead ECG signals using a max-min difference algorithm. *Comput Methods Programs Biomed*. 2017;144:61-75.
163. Rad AB, Eftestol T, Engan K, et al. ECG-Based Classification of Resuscitation Cardiac Rhythms for Retrospective Data Analysis. *IEEE Trans Biomed Eng*. 2017;64(10):2411-2418.
164. Zhou FY, Jin LP, Dong J. Premature ventricular contraction detection combining deep neural networks and rules inference. *Artif Intell Med*. 2017;79:42-51.
165. Kiranyaz S, Ince T, Gabbouj M. Personalized Monitoring and Advance Warning System for Cardiac Arrhythmias. *Sci Rep*. 2017;7(1):9270.
166. Mustaqeem A, Anwar SM, Majid M, Khan AR. Wrapper method for feature selection to classify cardiac arrhythmia. *Conf Proc IEEE Eng Med Biol Soc*. 2017;2017:3656-3659.
167. Wang T, Lu M, Du Q, et al. An integrated anti-arrhythmic target network of a Chinese medicine compound, Wenxin Keli, revealed by combined machine learning and molecular pathway analysis. *Mol Biosyst*. 2017;13(5):1018-1030.
168. Sadrawi M, Lin CH, Lin YT, et al. Arrhythmia Evaluation in Wearable ECG Devices. *Sensors (Basel)*. 2017;17(11).
169. Lemkaddem A, Proenca M, Delgado-Gonzalo R, et al. An autonomous medical monitoring system: Validation on arrhythmia detection. *Conf Proc IEEE Eng Med Biol Soc*. 2017;2017:4553-4556.
170. Liu H, Li X, Xie G, et al. Precision Cohort Finding with Outcome-Driven Similarity Analytics: A Case Study of Patients with Atrial Fibrillation. *Stud Health Technol Inform*. 2017;245:491-495.
171. Sharifi M, Buzatu D, Harris S, Wilkes J. Development of models for predicting Torsade de Pointes cardiac arrhythmias using perceptron neural networks. *BMC Bioinformatics*. 2017;18(Suppl 14):497.
172. Sahoo S, Mohanty M, Behera S, Sabut SK. ECG beat classification using empirical mode decomposition and mixture of features. *J Med Eng Technol*. 2017;41(8):652-661.
173. Lenning M, Fortunato J, Le T, et al. Real-Time Monitoring and Analysis of Zebrafish Electrocardiogram with Anomaly Detection. *Sensors (Basel)*. 2017;18(1).
174. Clifford GD, Liu C, Moody B, et al. AF Classification from a Short Single Lead ECG Recording: the PhysioNet/Computing in Cardiology Challenge 2017. *Comput Cardiol (2010)*. 2017;44.
175. Yu M, Chen F, Zhang G, et al. [Research on malignant arrhythmia detection algorithm using neural network optimized by genetic algorithm]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*. 2017;34(3):421-430.

176. Ferrer-Albero A, Godoy EJ, Lozano M, et al. Non-invasive localization of atrial ectopic beats by using simulated body surface P-wave integral maps. *PLoS One*. 2017;12(7):e0181263.
177. Li X, Zhang Y, Li H, Zhao Y. Modeling of the hERG K<sup>+</sup> Channel Blockage Using Online Chemical Database and Modeling Environment (OCHEM). *Mol Inform*. 2017;36(12).
178. Marzec L, Raghavan S, Banaei-Kashani F, et al. Device-measured physical activity data for classification of patients with ventricular arrhythmia events: A pilot investigation. *PLoS One*. 2018;13(10):e0206153.
179. Iqbal U, Wah TY, Habib Ur Rehman M, Mujtaba G, Imran M, Shoaib M. Deep Deterministic Learning for Pattern Recognition of Different Cardiac Diseases through the Internet of Medical Things. *J Med Syst*. 2018;42(12):252.
180. Yao Z, Chen Y. Arrhythmia Classification from Single Lead ECG by Multi-Scale Convolutional Neural Networks. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:344-347.
181. Nguyen MT, Nguyen BV, Kim K. Deep Feature Learning for Sudden Cardiac Arrest Detection in Automated External Defibrillators. *Sci Rep*. 2018;8(1):17196.
182. Liu M, Kim Y. Classification of Heart Diseases Based On ECG Signals Using Long Short-Term Memory. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:2707-2710.
183. Yu R, Gao Y, Duan X, Zhu T, Wang Z, Jiao B. QRS Detection and Measurement Method of ECG Paper Based on Convolutional Neural Networks. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:4636-4639.
184. Malik A, Peng T, Trew ML. A machine learning approach to reconstruction of heart surface potentials from body surface potentials. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:4828-4831.
185. He Z, Zhang X, Cao Y, Liu Z, Zhang B, Wang X. LiteNet: Lightweight Neural Network for Detecting Arrhythmias at Resource-Constrained Mobile Devices. *Sensors (Basel)*. 2018;18(4).
186. Lee H, Whang M. Heart Rate Estimated from Body Movements at Six Degrees of Freedom by Convolutional Neural Networks. *Sensors (Basel)*. 2018;18(5).
187. Yildirim O. A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification. *Comput Biol Med*. 2018;96:189-202.
188. Sengupta PP, Kulkarni H, Narula J. Prediction of Abnormal Myocardial Relaxation From Signal Processed Surface ECG. *J Am Coll Cardiol*. 2018;71(15):1650-1660.
189. Mustaqeem A, Anwar SM, Majid M. Multiclass Classification of Cardiac Arrhythmia Using Improved Feature Selection and SVM Invariants. *Comput Math Methods Med*. 2018;2018:7310496.
190. Sadr N, Jayawardhana M, Pham TT, Tang R, Balaei AT, de Chazal P. A low-complexity algorithm for detection of atrial fibrillation using an ECG. *Physiol Meas*. 2018;39(6):064003.
191. Mathews SM, Kambhamettu C, Barner KE. A novel application of deep learning for single-lead ECG classification. *Comput Biol Med*. 2018;99:53-62.
192. Godoy EJ, Lozano M, Garcia-Fernandez I, et al. Atrial Fibrosis Hampers Non-invasive Localization of Atrial Ectopic Foci From Multi-Electrode Signals: A 3D Simulation Study. *Front Physiol*. 2018;9:404.
193. Shaik NA, Awan ZA, Verma PK, Elango R, Banaganapalli B. Protein phenotype diagnosis of autosomal dominant calmodulin mutations causing irregular heart rhythms. *J Cell Biochem*. 2018;119(10):8233-8248.
194. Dobkowska-Chudon W, Wrobel M, Karłowicz P, et al. Detecting cerebrovascular changes in the brain caused by hypertension in atrial fibrillation group using acoustocerebrography. *PLoS One*. 2018;13(7):e0199999.



195. Parvaneh S, Rubin J, Rahman A, Conroy B, Babaeizadeh S. Analyzing single-lead short ECG recordings using dense convolutional neural networks and feature-based post-processing to detect atrial fibrillation. *Physiol Meas.* 2018;39(8):084003.
196. Bollepalli SC, Challa SS, Jana S. Robust Heartbeat Detection from Multimodal Data via CNN-based Generalizable Information Fusion. *IEEE Trans Biomed Eng.* 2018.
197. Warrick PA, Nabhan Homsy M. Ensembling convolutional and long short-term memory networks for electrocardiogram arrhythmia detection. *Physiol Meas.* 2018;39(11):114002.
198. Faust O, Shenfield A, Kareem M, San TR, Fujita H, Acharya UR. Automated detection of atrial fibrillation using long short-term memory network with RR interval signals. *Comput Biol Med.* 2018;102:327-335.
199. Xu X, Wei S, Ma C, Luo K, Zhang L, Liu C. Atrial Fibrillation Beat Identification Using the Combination of Modified Frequency Slice Wavelet Transform and Convolutional Neural Networks. *J Healthc Eng.* 2018;2018:2102918.
200. Amezcua-Sanchez JP, Valtierra-Rodriguez M, Adeli H, Perez-Ramirez CA. A Novel Wavelet Transform-Homogeneity Model for Sudden Cardiac Death Prediction Using ECG Signals. *J Med Syst.* 2018;42(10):176.
201. Rubin J, Parvaneh S, Rahman A, Conroy B, Babaeizadeh S. Densely connected convolutional networks for detection of atrial fibrillation from short single-lead ECG recordings. *J Electrocardiol.* 2018;51(6S):S18-S21.
202. Bond RR, Novotny T, Andrsova I, et al. Automation bias in medicine: The influence of automated diagnoses on interpreter accuracy and uncertainty when reading electrocardiograms. *J Electrocardiol.* 2018;51(6S):S6-S11.
203. Attia ZI, Sugrue A, Asirvatham SJ, et al. Noninvasive assessment of dofetilide plasma concentration using a deep learning (neural network) analysis of the surface electrocardiogram: A proof of concept study. *PLoS One.* 2018;13(8):e0201059.
204. Kropf M, Hayn D, Morris D, et al. Cardiac anomaly detection based on time and frequency domain features using tree-based classifiers. *Physiol Meas.* 2018;39(11):114001.
205. Zhao Z, Zhang Y, Deng Y, Zhang X. ECG authentication system design incorporating a convolutional neural network and generalized S-Transformation. *Comput Biol Med.* 2018;102:168-179.
206. Blecker S, Sontag D, Horwitz LI, et al. Early Identification of Patients With Acute Decompensated Heart Failure. *J Card Fail.* 2018;24(6):357-362.
207. Cikes M, Sanchez-Martinez S, Claggett B, et al. Machine learning-based phenogrouping in heart failure to identify responders to cardiac resynchronization therapy. *Eur J Heart Fail.* 2018.
208. Hearn J, Ross HJ, Mueller B, et al. Neural Networks for Prognostication of Patients With Heart Failure. *Circ Heart Fail.* 2018;11(8):e005193.
209. Sodmann P, Vollmer M, Nath N, Kaderali L. A convolutional neural network for ECG annotation as the basis for classification of cardiac rhythms. *Physiol Meas.* 2018;39(10):104005.
210. Xu SS, Mak MW, Cheung CC. Towards End-to-End ECG Classification with Raw Signal Extraction and Deep Neural Networks. *IEEE J Biomed Health Inform.* 2018.
211. Yang W, Si Y, Wang D, Guo B. Automatic recognition of arrhythmia based on principal component analysis network and linear support vector machine. *Comput Biol Med.* 2018;101:22-32.
212. Fan X, Yao Q, Cai Y, Miao F, Sun F, Li Y. Multiscaled Fusion of Deep Convolutional Neural Networks for Screening Atrial Fibrillation From Single Lead Short ECG Recordings. *IEEE J Biomed Health Inform.* 2018;22(6):1744-1753.

213. Plesinger F, Nejedly P, Viscor I, Halamek J, Jurak P. Parallel use of a convolutional neural network and bagged tree ensemble for the classification of Holter ECG. *Physiol Meas.* 2018;39(9):094002.
214. Das A, Pradhapan P, Groenendaal W, et al. Unsupervised heart-rate estimation in wearables with Liquid states and a probabilistic readout. *Neural Netw.* 2018;99:134-147.
215. Xiang Y, Lin Z, Meng J. Automatic QRS complex detection using two-level convolutional neural network. *Biomed Eng Online.* 2018;17(1):13.
216. Tison GH, Sanchez JM, Ballinger B, et al. Passive Detection of Atrial Fibrillation Using a Commercially Available Smartwatch. *JAMA Cardiol.* 2018;3(5):409-416.
217. Gliner V, Yaniv Y. An SVM approach for identifying atrial fibrillation. *Physiol Meas.* 2018;39(9):094007.
218. Faganeli Pucer J, Kukar M. A topological approach to delineation and arrhythmic beats detection in unprocessed long-term ECG signals. *Comput Methods Programs Biomed.* 2018;164:159-168.
219. Clerx M, Heijman J, Collins P, Volders PGA. Predicting changes to INa from missense mutations in human SCN5A. *Sci Rep.* 2018;8(1):12797.
220. Lawson BA, Burrage K, Burrage P, Drovandi CC, Bueno-Orovio A. Slow Recovery of Excitability Increases Ventricular Fibrillation Risk as Identified by Emulation. *Front Physiol.* 2018;9:1114.
221. He R, Wang K, Zhao N, et al. Automatic Detection of Atrial Fibrillation Based on Continuous Wavelet Transform and 2D Convolutional Neural Networks. *Front Physiol.* 2018;9:1206.
222. Isasi I, Irusta U, Elola A, et al. A Machine Learning Shock Decision Algorithm for use during Piston-driven Chest Compressions. *IEEE Trans Biomed Eng.* 2018.
223. Xie P, Wang G, Zhang C, et al. Bidirectional Recurrent Neural Network And Convolutional Neural Network (BiRCNN) For ECG Beat Classification. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:2555-2558.
224. Miyatani S, Fujiwara K, Kano M. Denoising Autoencoder-based Modification of RRI data with Premature Ventricular Contraction for Precise Heart Rate Variability Analysis. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:5018-5021.
225. Khamis H, Chen J, Stephen Redmond J, Lovell NH. Detection of Atrial Fibrillation from RR Intervals and PQRST Morphology using a Neural Network Ensemble. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:5998-6001.
226. Zhou X, Zhu X, Nakamura K, Mahito N. Premature Ventricular Contraction Detection from Ambulatory ECG Using Recurrent Neural Networks. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:2551-2554.
227. Hernandez F, Mendez D, Amado L, Altuve M. Atrial Fibrillation Detection in Short Single Lead ECG Recordings Using Wavelet Transform and Artificial Neural Networks. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:5982-5985.
228. Anwar SM, Gul M, Majid M, Alnowami M. Arrhythmia Classification of ECG Signals Using Hybrid Features. *Comput Math Methods Med.* 2018;2018:1380348.
229. Rizwan M, Whitaker BM, Anderson DV. AF detection from ECG recordings using feature selection, sparse coding, and ensemble learning. *Physiol Meas.* 2018;39(12):124007.
230. Yang T, Yu L, Jin Q, et al. Localization of Origins of Premature Ventricular Contraction by Means of Convolutional Neural Network From 12-Lead ECG. *IEEE Trans Biomed Eng.* 2018;65(7):1662-1671.
231. Xia Y, Wulan N, Wang K, Zhang H. Detecting atrial fibrillation by deep convolutional neural networks. *Comput Biol Med.* 2018;93:84-92.
232. Kamaleswaran R, Mahajan R, Akbilgic O. A robust deep convolutional neural network for the classification of abnormal cardiac rhythm using single lead electrocardiograms of variable length. *Physiol Meas.* 2018;39(3):035006.

233. Deng M, Wang C, Tang M, Zheng T. Extracting cardiac dynamics within ECG signal for human identification and cardiovascular diseases classification. *Neural Netw.* 2018;100:70-83.
234. Zhong W, Liao L, Guo X, Wang G. A deep learning approach for fetal QRS complex detection. *Physiol Meas.* 2018;39(4):045004.
235. Hajimolahoseini H, Hashemi J, Gazor S, Redfearn D. Inflection point analysis: A machine learning approach for extraction of IEGM active intervals during atrial fibrillation. *Artif Intell Med.* 2018;85:7-15.
236. Xiong Z, Nash MP, Cheng E, Fedorov VV, Stiles MK, Zhao J. ECG signal classification for the detection of cardiac arrhythmias using a convolutional recurrent neural network. *Physiol Meas.* 2018;39(9):094006.
237. Mena LJ, Felix VG, Ochoa A, et al. Mobile Personal Health Monitoring for Automated Classification of Electrocardiogram Signals in Elderly. *Comput Math Methods Med.* 2018;2018:9128054.
238. Teijeiro T, Garcia CA, Castro D, Felix P. Abductive reasoning as a basis to reproduce expert criteria in ECG atrial fibrillation identification. *Physiol Meas.* 2018;39(8):084006.
239. Poh MZ, Poh YC, Chan PH, et al. Diagnostic assessment of a deep learning system for detecting atrial fibrillation in pulse waveforms. *Heart.* 2018;104(23):1921-1928.
240. Oh SL, Ng EYK, Tan RS, Acharya UR. Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats. *Comput Biol Med.* 2018;102:278-287.
241. Kumar A, Komaragiri R, Kumar M. Heart rate monitoring and therapeutic devices: A wavelet transform based approach for the modeling and classification of congestive heart failure. *ISA Trans.* 2018;79:239-250.
242. Savalia S, Emamian V. Cardiac Arrhythmia Classification by Multi-Layer Perceptron and Convolution Neural Networks. *Bioengineering (Basel).* 2018;5(2).
243. McGillivray MF, Cheng W, Peters NS, Christensen K. Machine learning methods for locating re-entrant drivers from electrograms in a model of atrial fibrillation. *R Soc Open Sci.* 2018;5(4):172434.
244. G S, T KP, V KK. Classification of ECG beats using deep belief network and active learning. *Med Biol Eng Comput.* 2018;56(10):1887-1898.
245. Ma Y, Alhrishy M, Narayan SA, Mountney P, Rhode KS. A novel real-time computational framework for detecting catheters and rigid guidewires in cardiac catheterization procedures. *Med Phys.* 2018;45(11):5066-5079.
246. Zhu H, Pan Y, Cheng KT, Huan R. A lightweight piecewise linear synthesis method for standard 12-lead ECG signals based on adaptive region segmentation. *PLoS One.* 2018;13(10):e0206170.
247. Wu Y, Yang F, Liu Y, Zha X, Yuan S. A Comparison of 1-D and 2-D Deep Convolutional Neural Networks in ECG Classification. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:324-327.
248. Nagaraj SB, Tjepkema-Cloostermans MC, Ruijter BJ, Hofmeijer J, van Putten M. The revised Cerebral Recovery Index improves predictions of neurological outcome after cardiac arrest. *Clin Neurophysiol.* 2018;129(12):2557-2566.
249. Akca B, Erdil N, Colak MC, Disli OM, Battaloglu B, Colak C. Is There Any Difference in Risk Factors between Male and Female Patients in New-Onset Atrial Fibrillation after Coronary Artery Bypass Grafting? *Thorac Cardiovasc Surg.* 2018;66(6):483-490.
250. Kwon JM, Lee Y, Lee Y, Lee S, Park J. An Algorithm Based on Deep Learning for Predicting In-Hospital Cardiac Arrest. *J Am Heart Assoc.* 2018;7(13).
251. Sbröllini A, De Jongh MC, Ter Haar CC, et al. Serial electrocardiography to detect newly emerging or aggravating cardiac pathology: a deep-learning approach. *Biomedical Engineering Online.* 2019;18(1):15.

252. Erdenebayar U, Kim H, Park J-U, Kang D, Lee K-J. Automatic Prediction of Atrial Fibrillation Based on Convolutional Neural Network Using a Short-term Normal Electrocardiogram Signal. *Journal of Korean Medical Science*. 2019;34(7):e64.
253. Biswas D, Everson L, Liu M, et al. CorNET: Deep Learning Framework for PPG-Based Heart Rate Estimation and Biometric Identification in Ambulant Environment. *IEEE transactions on biomedical circuits and systems*. 2019;13(2):282-291.
254. Hannun AY, Rajpurkar P, Haghpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*. 2019;25(1):65-69.
255. Bhattacharya M, Lu D-Y, Kudchadkar SM, et al. Identifying Ventricular Arrhythmias and Their Predictors by Applying Machine Learning Methods to Electronic Health Records in Patients With Hypertrophic Cardiomyopathy (HCM-VAR-Risk Model). *The American Journal of Cardiology*. 2019.
256. Attia ZI, Kapa S, Lopez-Jimenez F, et al. Screening for cardiac contractile dysfunction using an artificial intelligence-enabled electrocardiogram. *Nature Medicine*. 2019;25(1):70-74.
257. Gradolewski D, Magenes G, Johansson S, Kulesza WJ. A Wavelet Transform-Based Neural Network Denoising Algorithm for Mobile Phonocardiography. *Sensors (Basel, Switzerland)*. 2019;19(4).
258. Karaoguz MR, Yurtseven E, Aslan G, et al. The quality of ECG data acquisition, and diagnostic performance of a novel adhesive patch for ambulatory cardiac rhythm monitoring in arrhythmia detection. *Journal of Electrocardiology*. 2019;54:28-35.
259. Lozoya RC, Berte B, Cochet H, Jais P, Ayache N, Sermesant M. Model-Based Feature Augmentation for Cardiac Ablation Target Learning From Images. *IEEE Trans Biomed Eng*. 2019;66(1):30-40.
260. Oliveira BRd, Abreu CCEd, Duarte MAQ, Vieira Filho J. Geometrical features for premature ventricular contraction recognition with analytic hierarchy process based machine learning algorithms selection. *Computer Methods and Programs in Biomedicine*. 2019;169:59-69.
261. Coult J, Blackwood J, Sherman L, Rea TD, Kudenchuk PJ, Kwok H. Ventricular Fibrillation Waveform Analysis During Chest Compressions to Predict Survival From Cardiac Arrest. *Circulation Arrhythmia and Electrophysiology*. 2019;12(1):e006924.
262. Gordon M, Williams C. PVC Detection Using a Convolutional Autoencoder and Random Forest Classifier. *Pacific Symposium on Biocomputing Pacific Symposium on Biocomputing*. 2019;24:42-53.
263. Sutton JR, Mahajan R, Akbilgic O, Kamaleswaran R. PhysOnline: An Open Source Machine Learning Pipeline for Real-Time Analysis of Streaming Physiological Waveform. *IEEE journal of biomedical and health informatics*. 2019;23(1):59-65.
264. Cazeau S, Toulemont M, Ritter P, Reygner J. Statistical ranking of electromechanical dyssynchrony parameters for CRT. *Open Heart*. 2019;6(1):e000933.
265. Alqudah AM, Albadarneh A, Abu-Qasmieh I, Alquran H. Developing of robust and high accurate ECG beat classification by combining Gaussian mixtures and wavelets features. *Australasian Physical & Engineering Sciences in Medicine*. 2019;42(1):149-157.
266. Xie L, Hou C, Jiang X, Zhao J, Li Y, Xiao T. A compound heterozygosity of Tecrl gene confirmed in a catecholaminergic polymorphic ventricular tachycardia family. *European Journal of Medical Genetics*. 2019.
267. Liu N, Koh ZX, Goh J, et al. Prediction of adverse cardiac events in emergency department patients with chest pain using machine learning for variable selection. *BMC Med Inform Decis Mak*. 2014;14:75.

268. Kumar PR, Priya M. Classification of atherosclerotic and non-atherosclerotic individuals using multiclass state vector machine. *Technol Health Care*. 2014;22(4):583-595.
269. Alshurafa N, Eastwood JA, Pourhomayoun M, Liu JJ, Sarrafzadeh M. Remote health monitoring: predicting outcome success based on contextual features for cardiovascular disease. *Conf Proc IEEE Eng Med Biol Soc*. 2014;2014:1777-1781.
270. Zellweger MJ, Brinkert M, Bucher U, Tsirkin A, Ruff P, Pfisterer ME. A new memetic pattern based algorithm to diagnose/exclude coronary artery disease. *Int J Cardiol*. 2014;174(1):184-186.
271. Guidi G, Pettenati MC, Melillo P, Iadanza E. A machine learning system to improve heart failure patient assistance. *IEEE J Biomed Health Inform*. 2014;18(6):1750-1756.
272. Liu N, Lee MA, Ho AF, et al. Risk stratification for prediction of adverse coronary events in emergency department chest pain patients with a machine learning score compared with the TIMI score. *Int J Cardiol*. 2014;177(3):1095-1097.
273. Nakajima K, Matsuo S, Wakabayashi H, et al. Diagnostic Performance of Artificial Neural Network for Detecting Ischemia in Myocardial Perfusion Imaging. *Circ J*. 2015;79(7):1549-1556.
274. Chen H, Yang B, Liu D, et al. Using Blood Indexes to Predict Overweight Statuses: An Extreme Learning Machine-Based Approach. *PLoS One*. 2015;10(11):e0143003.
275. Panahiazar M, Taslimitehrani V, Pereira NL, Pathak J. Using EHRs for Heart Failure Therapy Recommendation Using Multidimensional Patient Similarity Analytics. *Stud Health Technol Inform*. 2015;210:369-373.
276. Panahiazar M, Taslimitehrani V, Pereira N, Pathak J. Using EHRs and Machine Learning for Heart Failure Survival Analysis. *Stud Health Technol Inform*. 2015;216:40-44.
277. Torii M, Fan JW, Yang WL, et al. Risk factor detection for heart disease by applying text analytics in electronic medical records. *J Biomed Inform*. 2015;58 Suppl:S164-170.
278. Arsanjani R, Dey D, Khachatryan T, et al. Prediction of revascularization after myocardial perfusion SPECT by machine learning in a large population. *J Nucl Cardiol*. 2015;22(5):877-884.
279. Drenos F, Grossi E, Buscema M, Humphries SE. Networks in Coronary Heart Disease Genetics As a Step towards Systems Epidemiology. *PLoS One*. 2015;10(5):e0125876.
280. Roberts K, Shooshan SE, Rodriguez L, Abhyankar S, Kilicoglu H, Demner-Fushman D. The role of fine-grained annotations in supervised recognition of risk factors for heart disease from EHRs. *J Biomed Inform*. 2015;58 Suppl:S111-119.
281. Poddar MG, Kumar V, Sharma YP. Automated diagnosis of coronary artery diseased patients by heart rate variability analysis using linear and non-linear methods. *J Med Eng Technol*. 2015;39(6):331-341.
282. Sladojevic M, Cankovic M, Cemerlic S, Mihajlovic B, Adic F, Jarakovic M. Data Mining Approach for in-Hospital Treatment Outcome in Patients with Acute Coronary Syndrome. *Med Pregl*. 2015;68(5-6):157-161.
283. Yang H, Garibaldi JM. A hybrid model for automatic identification of risk factors for heart disease. *J Biomed Inform*. 2015;58 Suppl:S171-182.
284. Jonnagaddala J, Liaw ST, Ray P, Kumar M, Dai HJ, Hsu CY. Identification and Progression of Heart Disease Risk Factors in Diabetic Patients from Longitudinal Electronic Health Records. *Biomed Res Int*. 2015;2015:636371.
285. Chen Q, Li H, Tang B, et al. An automatic system to identify heart disease risk factors in clinical texts over time. *J Biomed Inform*. 2015;58 Suppl:S158-163.

286. Isma'eel HA, Cremer PC, Khalaf S, et al. Artificial neural network modeling enhances risk stratification and can reduce downstream testing for patients with suspected acute coronary syndromes, negative cardiac biomarkers, and normal ECGs. *Int J Cardiovasc Imaging*. 2016;32(4):687-696.
287. Tylman W, Waszyrowski T, Napieralski A, et al. Real-time prediction of acute cardiovascular events using hardware-implemented Bayesian networks. *Comput Biol Med*. 2016;69:245-253.
288. Churpek MM, Yuen TC, Winslow C, Meltzer DO, Kattan MW, Edelson DP. Multicenter Comparison of Machine Learning Methods and Conventional Regression for Predicting Clinical Deterioration on the Wards. *Crit Care Med*. 2016;44(2):368-374.
289. Wiharto W, Kusnanto H, Herianto H. Intelligence System for Diagnosis Level of Coronary Heart Disease with K-Star Algorithm. *Healthc Inform Res*. 2016;22(1):30-38.
290. Hoyt R, Linnville S, Thaler S, Moore J. Digital Family History Data Mining with Neural Networks: A Pilot Study. *Perspect Health Inf Manag*. 2016;13:1c.
291. Araki T, Ikeda N, Shukla D, et al. A new method for IVUS-based coronary artery disease risk stratification: A link between coronary & carotid ultrasound plaque burdens. *Comput Methods Programs Biomed*. 2016;124:161-179.
292. Araki T, Ikeda N, Shukla D, et al. PCA-based polling strategy in machine learning framework for coronary artery disease risk assessment in intravascular ultrasound: A link between carotid and coronary grayscale plaque morphology. *Comput Methods Programs Biomed*. 2016;128:137-158.
293. Celutkiene J, Burneikaite G, Petkevicius L, Balkeviciene L, Laucevicius A. Combination of single quantitative parameters into multiparametric model for ischemia detection is not superior to visual assessment during dobutamine stress echocardiography. *Cardiovasc Ultrasound*. 2016;14:13.
294. Motwani M, Dey D, Berman DS, et al. Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: a 5-year multicentre prospective registry analysis. *Eur Heart J*. 2017;38(7):500-507.
295. Afarideh M, Aryan Z, Ghajar A, et al. Complex association of serum alanine aminotransferase with the risk of future cardiovascular disease in type 2 diabetes. *Atherosclerosis*. 2016;254:42-51.
296. Liu Y, Scirica BM, Stultz CM, Guttag JV. Beatquency domain and machine learning improve prediction of cardiovascular death after acute coronary syndrome. *Sci Rep*. 2016;6:34540.
297. Ross EG, Shah NH, Dalman RL, Nead KT, Cooke JP, Leeper NJ. The use of machine learning for the identification of peripheral artery disease and future mortality risk. *J Vasc Surg*. 2016;64(5):1515-1522 e1513.
298. Lopez-de-Andres A, Hernandez-Barrera V, Lopez R, et al. Predictors of in-hospital mortality following major lower extremity amputations in type 2 diabetic patients using artificial neural networks. *BMC Med Res Methodol*. 2016;16(1):160.
299. Hu D, Huang Z, Chan TM, Dong W, Lu X, Duan H. Utilizing Chinese Admission Records for MACE Prediction of Acute Coronary Syndrome. *Int J Environ Res Public Health*. 2016;13(9).
300. Narain R, Saxena S, Goyal AK. Cardiovascular risk prediction: a comparative study of Framingham and quantum neural network based approach. *Patient Prefer Adherence*. 2016;10:1259-1270.
301. Spertus JV, SL TN, Wolf R, Cioffi M, Lovett A, Rose S. Assessing Hospital Performance After Percutaneous Coronary Intervention Using Big Data. *Circ Cardiovasc Qual Outcomes*. 2016;9(6):659-669.
302. Mirian NS, Sedehi M, Kheiri S, Ahmadi A. A Hybrid ANN-GA Model to Prediction of Bivariate Binary Responses: Application to Joint Prediction of Occurrence of Heart Block and Death in Patients with Myocardial Infarction. *J Res Health Sci*. 2016;16(4):190-194.

303. Orphanou K, Stassopoulou A, Keravnou E. DBN-Extended: A Dynamic Bayesian Network Model Extended With Temporal Abstractions for Coronary Heart Disease Prognosis. *IEEE J Biomed Health Inform.* 2016;20(3):944-952.
304. Berikol GB, Yildiz O, Ozcan IT. Diagnosis of Acute Coronary Syndrome with a Support Vector Machine. *J Med Syst.* 2016;40(4):84.
305. Verma L, Srivastava S, Negi PC. A Hybrid Data Mining Model to Predict Coronary Artery Disease Cases Using Non-Invasive Clinical Data. *J Med Syst.* 2016;40(7):178.
306. McRae MP, Simmons G, Wong J, McDevitt JT. Programmable Bio-nanochip Platform: A Point-of-Care Biosensor System with the Capacity To Learn. *Acc Chem Res.* 2016;49(7):1359-1368.
307. Karimi A, Rahmati SM, Sera T, Kudo S, Navidbakhsh M. A Combination of Constitutive Damage Model and Artificial Neural Networks to Characterize the Mechanical Properties of the Healthy and Atherosclerotic Human Coronary Arteries. *Artif Organs.* 2017;41(9):E103-E117.
308. Schafer S, de Marvao A, Adami E, et al. Titin-truncating variants affect heart function in disease cohorts and the general population. *Nat Genet.* 2017;49(1):46-53.
309. Arabasadi Z, Alizadehsani R, Roshanzamir M, Moosaei H, Yarifard AA. Computer aided decision making for heart disease detection using hybrid neural network-Genetic algorithm. *Comput Methods Programs Biomed.* 2017;141:19-26.
310. Podda GM, Grossi E, Palmerini T, et al. Prediction of high on-treatment platelet reactivity in clopidogrel-treated patients with acute coronary syndromes. *Int J Cardiol.* 2017;240:60-65.
311. Alghamdi M, Al-Mallah M, Keteyian S, Brawner C, Ehrman J, Sakr S. Predicting diabetes mellitus using SMOTE and ensemble machine learning approach: The Henry Ford Exercise Testing (FIT) project. *PLoS One.* 2017;12(7):e0179805.
312. Gao H, Aderhold A, Mangion K, Luo X, Husmeier D, Berry C. Changes and classification in myocardial contractile function in the left ventricle following acute myocardial infarction. *J R Soc Interface.* 2017;14(132).
313. Shah RV, Yeri AS, Murthy VL, et al. Association of Multiorgan Computed Tomographic Phenomap With Adverse Cardiovascular Health Outcomes: The Framingham Heart Study. *JAMA Cardiol.* 2017;2(11):1236-1246.
314. Oguz C, Sen SK, Davis AR, Fu YP, O'Donnell CJ, Gibbons GH. Genotype-driven identification of a molecular network predictive of advanced coronary calcium in ClinSeq(R) and Framingham Heart Study cohorts. *BMC Syst Biol.* 2017;11(1):99.
315. Hu D, Huang Z, Chan TM, Dong W, Lu X, Duan H. Acute Coronary Syndrome Risk Prediction Based on GRACE Risk Score. *Stud Health Technol Inform.* 2017;245:398-402.
316. Karimi A, Rahmati SM, Sera T, Kudo S, Navidbakhsh M. A combination of experimental and numerical methods to investigate the role of strain rate on the mechanical properties and collagen fiber orientations of the healthy and atherosclerotic human coronary arteries. *Bioengineered.* 2017;8(2):154-170.
317. Lee KE, Ryu AJ, Shin ES, Shim EB. Physiome approach for the analysis of vascular flow reserve in the heart and brain. *Pflugers Arch.* 2017;469(5-6):613-628.
318. Forssen H, Patel R, Fitzpatrick N, et al. Evaluation of Machine Learning Methods to Predict Coronary Artery Disease Using Metabolomic Data. *Stud Health Technol Inform.* 2017;235:111-115.
319. Takeuchi M, Inuzuka R, Hayashi T, et al. Novel Risk Assessment Tool for Immunoglobulin Resistance in Kawasaki Disease: Application Using a Random Forest Classifier. *Pediatr Infect Dis J.* 2017;36(9):821-826.
320. Buchan K, Filannino M, Uzuner O. Automatic prediction of coronary artery disease from clinical narratives. *J Biomed Inform.* 2017;72:23-32.

321. Ambale-Venkatesh B, Yang X, Wu CO, et al. Cardiovascular Event Prediction by Machine Learning: The Multi-Ethnic Study of Atherosclerosis. *Circ Res.* 2017;121(9):1092-1101.
322. Basu S, Sussman JB, Berkowitz SA, Hayward RA, Yudkin JS. Development and validation of Risk Equations for Complications Of type 2 Diabetes (RECODE) using individual participant data from randomised trials. *Lancet Diabetes Endocrinol.* 2017;5(10):788-798.
323. Kim JK, Kang S. Neural Network-Based Coronary Heart Disease Risk Prediction Using Feature Correlation Analysis. *J Healthc Eng.* 2017;2017.
324. El-Askary H, LaHaye N, Linstead E, Sprigg WA, Yacoub M. Remote sensing observation of annual dust cycles and possible causality of Kawasaki disease outbreaks in Japan. *Glob Cardiol Sci Pract.* 2017;2017(3):e201722.
325. Isma'eel HA, Sakr GE, Serhan M, et al. Artificial neural network-based model enhances risk stratification and reduces non-invasive cardiac stress imaging compared to Diamond-Forrester and Morise risk assessment models: A prospective study. *J Nucl Cardiol.* 2018;25(5):1601-1609.
326. Tao R, Zhang S, Huang X, et al. Magnetocardiography based Ischemic Heart Disease Detection and Localization using Machine Learning Methods. *IEEE Trans Biomed Eng.* 2018.
327. Hae H, Kang SJ, Kim WJ, et al. Machine learning assessment of myocardial ischemia using angiography: Development and retrospective validation. *PLoS Med.* 2018;15(11):e1002693.
328. Shi P, Xin J, Liu S, Deng Y, Zheng N. Vulnerable Plaque Recognition Based on Attention Model with Deep Convolutional Neural Network. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:834-837.
329. Pieszko K, Hiczekiewicz J, Budzianowski P, et al. Machine-learned models using hematological inflammation markers in the prediction of short-term acute coronary syndrome outcomes. *J Transl Med.* 2018;16(1):334.
330. van Rosendaal AR, Maliakal G, Kollu KK, et al. Maximization of the usage of coronary CTA derived plaque information using a machine learning based algorithm to improve risk stratification; insights from the CONFIRM registry. *J Cardiovasc Comput Tomogr.* 2018;12(3):204-209.
331. Duggal B, Subramanian J, Duggal M, et al. Survival outcomes post percutaneous coronary intervention: Why the hype about stent type? Lessons from a healthcare system in India. *PLoS One.* 2018;13(5):e0196830.
332. Alizadehsani R, Hosseini MJ, Khosravi A, et al. Non-invasive detection of coronary artery disease in high-risk patients based on the stenosis prediction of separate coronary arteries. *Comput Methods Programs Biomed.* 2018;162:119-127.
333. Xiao R, Xu Y, Pelter MM, Mortara DW, Hu X. A Deep Learning Approach to Examine Ischemic ST Changes in Ambulatory ECG Recordings. *AMIA Jt Summits Transl Sci Proc.* 2018;2017:256-262.
334. Rose S, Normand SL. Double robust estimation for multiple unordered treatments and clustered observations: Evaluating drug-eluting coronary artery stents. *Biometrics.* 2018.
335. Dozio E, Vianello E, Grossi E, Menicanti L, Schmitz G, Corsi Romanelli MM. Plasma fatty acid profile as biomarker of coronary artery disease: a pilot study using fourth generation artificial neural networks. *J Biol Regul Homeost Agents.* 2018;32(4):1007-1013.
336. Steele AJ, Denaxas SC, Shah AD, Hemingway H, Luscombe NM. Machine learning models in electronic health records can outperform conventional survival models for predicting patient mortality in coronary artery disease. *PLoS One.* 2018;13(8):e0202344.
337. Zellweger MJ, Tsirkin A, Vasilchenko V, et al. A new non-invasive diagnostic tool in coronary artery disease: artificial intelligence as an essential element of predictive, preventive, and personalized medicine. *EPMA J.* 2018;9(3):235-247.



338. Body R. Acute coronary syndromes diagnosis, version 2.0: Tomorrow's approach to diagnosing acute coronary syndromes? *Turk J Emerg Med.* 2018;18(3):94-99.
339. Yahyaie M, Tarokh MJ, Mahmoodiyar MA. Use of Internet of Things to Provide a New Model for Remote Heart Attack Prediction. *Telemed J E Health.* 2018.
340. Brown NJL, Coyne JC. Does Twitter language reliably predict heart disease? A commentary on Eichstaedt et al. (2015a). *PeerJ.* 2018;6:e5656.
341. McAllister P, Zheng H, Bond R, Moorhead A. Combining deep residual neural network features with supervised machine learning algorithms to classify diverse food image datasets. *Comput Biol Med.* 2018;95:217-233.
342. Zhang L, Wahle A, Chen Z, Lopez JJ, Kovarnik T, Sonka M. Predicting Locations of High-Risk Plaques in Coronary Arteries in Patients Receiving Statin Therapy. *IEEE Trans Med Imaging.* 2018;37(1):151-161.
343. Huang Z, Dong W, Duan H, Liu J. A Regularized Deep Learning Approach for Clinical Risk Prediction of Acute Coronary Syndrome Using Electronic Health Records. *IEEE Trans Biomed Eng.* 2018;65(5):956-968.
344. Tan JH, Hagiwara Y, Pang W, et al. Application of stacked convolutional and long short-term memory network for accurate identification of CAD ECG signals. *Comput Biol Med.* 2018;94:19-26.
345. Margulis K, Zhou Z, Fang Q, Sievers RE, Lee RJ, Zare RN. Combining Desorption Electrospray Ionization Mass Spectrometry Imaging and Machine Learning for Molecular Recognition of Myocardial Infarction. *Anal Chem.* 2018;90(20):12198-12206.
346. Yang F, Yang X, Teo SK, et al. Multi-dimensional proprio-proximus machine learning for assessment of myocardial infarction. *Comput Med Imaging Graph.* 2018;70:63-72.
347. Kigka VI, Georga EI, Sakellarios AI, et al. A Machine Learning Approach for the Prediction of the Progression of Cardiovascular Disease based on Clinical and Non-Invasive Imaging Data. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:6108-6111.
348. Singh P, Singh S, Pandi-Jain GS. Effective heart disease prediction system using data mining techniques. *Int J Nanomedicine.* 2018;13:121-124.
349. Haro Alonso D, Wernick MN, Yang Y, Germano G, Berman DS, Slomka P. Prediction of cardiac death after adenosine myocardial perfusion SPECT based on machine learning. *J Nucl Cardiol.* 2018.
350. Naushad SM, Hussain T, Indumathi B, Samreen K, Alrokayan SA, Kutala VK. Machine learning algorithm-based risk prediction model of coronary artery disease. *Mol Biol Rep.* 2018.
351. Stuckey TD, Gammon RS, Goswami R, et al. Cardiac Phase Space Tomography: A novel method of assessing coronary artery disease utilizing machine learning. *PLoS One.* 2018;13(8):e0198603.
352. Good WW, Erem B, Zenger B, Coll-Font J, Brooks DH, MacLeod RS. Temporal Performance of Laplacian Eigenmaps and 3D Conduction Velocity in Detecting Ischemic Stress. *J Electrocardiol.* 2018;51(6S):S116-S120.
353. Chen ZF, Young L, Yu CH, Shiao SPK. A Meta-Prediction of Methylenetetrahydrofolate-Reductase Polymorphisms and Air Pollution Increased the Risk of Ischemic Heart Diseases Worldwide. *Int J Environ Res Public Health.* 2018;15(7).
354. Öman O, Mäkelä T, Salli E, Savolainen S, Kangasniemi M. 3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke. *European Radiology Experimental.* 2019;3(1):8.
355. Heo J, Yoon JG, Park H, Kim YD, Nam HS, Heo JH. Machine Learning-Based Model for Prediction of Outcomes in Acute Stroke. *Stroke.* 2019:STROKEAHA118024293.

356. Vallée A, Cinaud A, Blachier V, Lelong H, Safar ME, Blacher J. Coronary heart disease diagnosis by artificial neural networks including aortic pulse wave velocity index and clinical parameters. *Journal of Hypertension*. 2019.
357. Balanescu DV, Monlezun DJ, Donisan T, et al. A Cancer Paradox: Machine-Learning Backed Propensity-Score Analysis of Coronary Angiography Findings in Cardio-Oncology. *The Journal of Invasive Cardiology*. 2019;31(1):21-26.
358. Zabihollahy F, White JA, Ukwatta E. Convolutional neural network-based approach for segmentation of left ventricle myocardial scar from 3D late gadolinium enhancement MR images. *Medical Physics*. 2019;46(4):1740-1751.
359. Yu M, Lu Z, Shen C, et al. The best predictor of ischemic coronary stenosis: subtended myocardial volume, machine learning-based FFRCT, or high-risk plaque features? *European Radiology*. 2019.
360. Vernon ST, Hansen T, Kott KA, Yang JY, O'Sullivan JF, Figtree GA. Utilizing state-of-the-art "omics" technology and bioinformatics to identify new biological mechanisms and biomarkers for coronary artery disease. *Microcirculation (New York, NY: 1994)*. 2019;26(2):e12488.
361. Chen T, Brewster P, Tuttle KR, et al. Prediction of cardiovascular outcomes with machine learning techniques: application to the Cardiovascular Outcomes in Renal Atherosclerotic Lesions (CORAL) study. *International Journal of Nephrology and Renovascular Disease*. 2019;12:49-58.
362. Li Y, Liu SH, Niu L, Liu B. Unhealthy Behaviors, Prevention Measures, and Neighborhood Cardiovascular Health: A Machine Learning Approach. *Journal of public health management and practice: JPHMP*. 2019;25(1):E25-E28.
363. Freiman M, Manjeshwar R, Goshen L. Unsupervised abnormality detection through mixed structure regularization (MSR) in deep sparse autoencoders. *Medical Physics*. 2019.
364. Goto S, Kimura M, Katsumata Y, et al. Artificial intelligence to predict needs for urgent revascularization from 12-leads electrocardiography in emergency patients. *PloS One*. 2019;14(1):e0210103.
365. Pieszko K, Hiczkiewicz J, Budzianowski P, et al. Predicting Long-Term Mortality after Acute Coronary Syndrome Using Machine Learning Techniques and Hematological Markers. *Disease Markers*. 2019;2019:9056402.
366. Cho H, Lee J-G, Kang S-J, et al. Angiography-Based Machine Learning for Predicting Fractional Flow Reserve in Intermediate Coronary Artery Lesions. *Journal of the American Heart Association*. 2019;8(4):e011685.
367. Duan H, Sun Z, Dong W, Huang Z. Utilizing dynamic treatment information for MACE prediction of acute coronary syndrome. *BMC medical informatics and decision making*. 2019;19(1):5.
368. Bazylev VV, Bartosh FL, Babukov RM, Mikuliak AI. [Predictors of early occlusion of coronary bypass grafts in patients with extremely low ejection fraction]. *Angiologïia I Sosudistaia Khirurgïia = Angiology and Vascular Surgery*. 2019;25(1):153-157.
369. Adela R, Reddy PNC, Ghosh TS, et al. Serum protein signature of coronary artery disease in type 2 diabetes mellitus. *Journal of Translational Medicine*. 2019;17(1):17.
370. Sharma A, Sun J-L, Lokhnygina Y, et al. Patient Phenotypes, Cardiovascular Risk, and Ezetimibe Treatment in Patients After Acute Coronary Syndromes (from IMPROVE-IT). *The American Journal of Cardiology*. 2019;123(8):1193-1201.
371. Myers J, de Souza CR, Borghi-Silva A, et al. A neural network approach to predicting outcomes in heart failure using cardiopulmonary exercise testing. *Int J Cardiol*. 2014;171(2):265-269.
372. Schmitz B, De Maria R, Gatsios D, et al. Identification of genetic markers for treatment success in heart failure patients: insight from cardiac resynchronization therapy. *Circ Cardiovasc Genet*. 2014;7(6):760-770.

373. Ooi HL, Ng SC, Lim E, Salamonsen RF, Avolio AP, Lovell NH. Robust aortic valve non-opening detection for different cardiac conditions. *Artif Organs*. 2014;38(3):E57-67.
374. Bondar G, Cadeiras M, Wisniewski N, et al. Comparison of whole blood and peripheral blood mononuclear cell gene expression for evaluation of the perioperative inflammatory response in patients with advanced heart failure. *PLoS One*. 2014;9(12):e115097.
375. Summers RL, Pipke M, Wegerich S, Conkright G, Isom KC. Functionality of empirical model-based predictive analytics for the early detection of hemodynamic instability. *Biomed Sci Instrum*. 2014;50:219-224.
376. Guidi G, Pollonini L, Dacso CC, Iadanza E. A multi-layer monitoring system for clinical management of Congestive Heart Failure. *BMC Med Inform Decis Mak*. 2015;15 Suppl 3:S5.
377. Nilsson J, Ohlsson M, Hoglund P, Ekmeahag B, Koul B, Andersson B. The International Heart Transplant Survival Algorithm (IHTSA): a new model to improve organ sharing and survival. *PLoS One*. 2015;10(3):e0118644.
378. Meystre SM, Kim Y, Heavirland J, Williams J, Bray BE, Garvin J. Heart Failure Medications Detection and Prescription Status Classification in Clinical Narrative Documents. *Stud Health Technol Inform*. 2015;216:609-613.
379. Xu Z, Yang M, Wang X, Wang Z. Multi-objective optimization of pulsatile ventricular assist device hemocompatibility based on neural networks and a genetic algorithm. *Int J Artif Organs*. 2015;38(6):325-336.
380. Alonso-Betanzos A, Bolon-Canedo V, Heyndrickx GR, Kerkhof PL. Exploring Guidelines for Classification of Major Heart Failure Subtypes by Using Machine Learning. *Clin Med Insights Cardiol*. 2015;9(Suppl 1):57-71.
381. Mohammadzadeh N, Safdari R. Chronic Heart Failure Follow-up Management Based on Agent Technology. *Healthc Inform Res*. 2015;21(4):307-314.
382. Liu NT, Kramer GC, Khan MN, Kinsky MP, Salinas J. Blood pressure and heart rate from the arterial blood pressure waveform can reliably estimate cardiac output in a conscious sheep model of multiple hemorrhages and resuscitation using computer machine learning approaches. *J Trauma Acute Care Surg*. 2015;79(4 Suppl 2):S85-92.
383. Steventon A, Ariti C, Fisher E, Bardsley M. Effect of telehealth on hospital utilisation and mortality in routine clinical practice: a matched control cohort study in an early adopter site. *BMJ Open*. 2016;6(2):e009221.
384. Sideris C, Pourhomayoun M, Kalantarian H, Sarrafzadeh M. A flexible data-driven comorbidity feature extraction framework. *Comput Biol Med*. 2016;73:165-172.
385. Hou F, Wu X, Yi Z, Zhang J, Hai B. [Investigation into Feasibility of Congestive Heart Failure Diagnosis Based on Analysis of Very Short-term Heart Rate Variability]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*. 2016;33(3):559-563.
386. Blecker S, Katz SD, Horwitz LI, et al. Comparison of Approaches for Heart Failure Case Identification From Electronic Health Record Data. *JAMA Cardiol*. 2016;1(9):1014-1020.
387. Karanasiou GS, Tripoliti EE, Papadopoulos TG, et al. Predicting adherence of patients with HF through machine learning techniques. *Healthc Technol Lett*. 2016;3(3):165-170.
388. Altan G, Kutlu Y, Allahverdi N. A new approach to early diagnosis of congestive heart failure disease by using Hilbert-Huang transform. *Comput Methods Programs Biomed*. 2016;137:23-34.
389. Farhan W, Wang Z, Huang Y, Wang S, Wang F, Jiang X. A Predictive Model for Medical Events Based on Contextual Embedding of Temporal Sequences. *JMIR Med Inform*. 2016;4(4):e39.
390. Argerich S, Herrera S, Benito S, Giraldo BF. Evaluation of periodic breathing in respiratory flow signal of elderly patients using SVM and linear discriminant analysis. *Conf Proc IEEE Eng Med Biol Soc*. 2016;2016:4276-4279.

391. Ng K, Steinhubl SR, deFilippi C, Dey S, Stewart WF. Early Detection of Heart Failure Using Electronic Health Records: Practical Implications for Time Before Diagnosis, Data Diversity, Data Quantity, and Data Density. *Circ Cardiovasc Qual Outcomes*. 2016;9(6):649-658.
392. Tsai PF, Chen PC, Chen YY, et al. Length of Hospital Stay Prediction at the Admission Stage for Cardiology Patients Using Artificial Neural Network. *J Healthc Eng*. 2016;2016.
393. Masetic Z, Subasi A. Congestive heart failure detection using random forest classifier. *Comput Methods Programs Biomed*. 2016;130:54-64.
394. Buzaev IV, Plechev VV, Nikolaeva IE, Galimova RM. Artificial intelligence: Neural network model as the multidisciplinary team member in clinical decision support to avoid medical mistakes. *Chronic Dis Transl Med*. 2016;2(3):166-172.
395. Choi E, Schuetz A, Stewart WF, Sun J. Using recurrent neural network models for early detection of heart failure onset. *J Am Med Inform Assoc*. 2017;24(2):361-370.
396. Katz DH, Deo RC, Aguilar FG, et al. Phenomapping for the Identification of Hypertensive Patients with the Myocardial Substrate for Heart Failure with Preserved Ejection Fraction. *J Cardiovasc Transl Res*. 2017;10(3):275-284.
397. Zhang R, Ma S, Shanahan L, Munroe J, Horn S, Speedie S. Automatic Methods to Extract New York Heart Association Classification from Clinical Notes. *Proceedings (IEEE Int Conf Bioinformatics Biomed)*. 2017;2017:1296-1299.
398. Sanchez-Martinez S, Duchateau N, Erdei T, Fraser AG, Bijmens BH, Piella G. Characterization of myocardial motion patterns by unsupervised multiple kernel learning. *Med Image Anal*. 2017;35:70-82.
399. Meystre SM, Kim Y, Gobbel GT, et al. Congestive heart failure information extraction framework for automated treatment performance measures assessment. *J Am Med Inform Assoc*. 2017;24(e1):e40-e46.
400. Gleeson S, Liao YW, Dugo C, et al. ECG-derived spatial QRS-T angle is associated with ICD implantation, mortality and heart failure admissions in patients with LV systolic dysfunction. *PLoS One*. 2017;12(3):e0171069.
401. Xing J, Lu W, Liu R, et al. Machine-Learning-Assisted Approach for Discovering Novel Inhibitors Targeting Bromodomain-Containing Protein 4. *J Chem Inf Model*. 2017;57(7):1677-1690.
402. Iborra-Egea O, Galvez-Monton C, Roura S, et al. Mechanisms of action of sacubitril/valsartan on cardiac remodeling: a systems biology approach. *NPJ Syst Biol Appl*. 2017;3:12.
403. Oskouie SK, Prenner SB, Shah SJ, Sauer AJ. Differences in Repolarization Heterogeneity Among Heart Failure With Preserved Ejection Fraction Phenotypic Subgroups. *Am J Cardiol*. 2017;120(4):601-606.
404. Steinhoff G, Nesteruk J, Wolfien M, et al. Cardiac Function Improvement and Bone Marrow Response -: Outcome Analysis of the Randomized PERFECT Phase III Clinical Trial of Intramyocardial CD133(+) Application After Myocardial Infarction. *EBioMedicine*. 2017;22:208-224.
405. Al-Mallah MH, Elshawi R, Ahmed AM, et al. Using Machine Learning to Define the Association between Cardiorespiratory Fitness and All-Cause Mortality (from the Henry Ford Exercise Testing Project). *Am J Cardiol*. 2017;120(11):2078-2084.
406. Myers PD, Scirica BM, Stultz CM. Machine Learning Improves Risk Stratification After Acute Coronary Syndrome. *Sci Rep*. 2017;7(1):12692.
407. van der Ster BJP, Bennis FC, Delhaas T, Westerhof BE, Stok WJ, van Lieshout JJ. Support Vector Machine Based Monitoring of Cardio-Cerebrovascular Reserve during Simulated Hemorrhage. *Front Physiol*. 2017;8:1057.
408. Wenhui C, Guanzheng L, Su S, Qing J, Hung N. A CHF detection method based on deep learning with RR intervals. *Conf Proc IEEE Eng Med Biol Soc*. 2017;2017:3369-3372.

409. Mustaqeem A, Anwar SM, Khan AR, Majid M. A statistical analysis based recommender model for heart disease patients. *Int J Med Inform.* 2017;108:134-145.
410. Sakr S, Elshawi R, Ahmed AM, et al. Comparison of machine learning techniques to predict all-cause mortality using fitness data: the Henry ford exercise testing (FIT) project. *BMC Med Inform Decis Mak.* 2017;17(1):174.
411. Au-Yeung WM, Reinhall PG, Bardy GH, Brunton SL. Development and validation of warning system of ventricular tachyarrhythmia in patients with heart failure with heart rate variability data. *PLoS One.* 2018;13(11):e0207215.
412. Wang L, Zhou W, Liu N, Xing Y, Zhou X. CHF Detection with LSTM Neural Network. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:514-517.
413. Zhang Y, Yang Q, Pang W, et al. Congestive Heart Failure Detection Via Short-Time Electrocardiographic Monitoring For Fast Reference Advice In Urgent Medical Conditions. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:2256-2259.
414. McKinley D, Moye-Dickerson P, Davis S, Akil A. Impact of a Pharmacist-Led Intervention on 30-Day Readmission and Assessment of Factors Predictive of Readmission in African American Men With Heart Failure. *Am J Mens Health.* 2018:1557988318814295.
415. Xiao C, Ma T, Dieng AB, Blei DM, Wang F. Readmission prediction via deep contextual embedding of clinical concepts. *PLoS One.* 2018;13(4):e0195024.
416. Wang Z, Yao L, Li D, Ruan T, Liu M, Gao J. Mortality prediction system for heart failure with orthogonal relief and dynamic radius means. *Int J Med Inform.* 2018;115:10-17.
417. Mahajan SM, Mahajan AS, King R, Negahban S. Predicting Risk of 30-Day Readmissions Using Two Emerging Machine Learning Methods. *Stud Health Technol Inform.* 2018;250:250-255.
418. Rasmy L, Wu Y, Wang N, et al. A study of generalizability of recurrent neural network-based predictive models for heart failure onset risk using a large and heterogeneous EHR data set. *J Biomed Inform.* 2018;84:11-16.
419. Lindholm D, Fukaya E, Leeper NJ, Ingelsson E. Bioimpedance and New-Onset Heart Failure: A Longitudinal Study of >500 000 Individuals From the General Population. *J Am Heart Assoc.* 2018;7(13).
420. Kwon BC, Choi MJ, Kim JT, et al. RetainVis: Visual Analytics with Interpretable and Interactive Recurrent Neural Networks on Electronic Medical Records. *IEEE Trans Vis Comput Graph.* 2018.
421. Kakarmath S, Golas S, Felsted J, Kvedar J, Jethwani K, Agboola S. Validating a Machine Learning Algorithm to Predict 30-Day Re-Admissions in Patients With Heart Failure: Protocol for a Prospective Cohort Study. *JMIR Res Protoc.* 2018;7(9):e176.
422. Tabassian M, Sunderji I, Erdei T, et al. Diagnosis of Heart Failure With Preserved Ejection Fraction: Machine Learning of Spatiotemporal Variations in Left Ventricular Deformation. *J Am Soc Echocardiogr.* 2018;31(12):1272-1284 e1279.
423. Bose E, Radhakrishnan K. Using Unsupervised Machine Learning to Identify Subgroups Among Home Health Patients With Heart Failure Using Telehealth. *Comput Inform Nurs.* 2018;36(5):242-248.
424. Michalik J, Cacko A, Polinski J, et al. An interactive assistant for patients with cardiac implantable electronic devices: A study protocol of the LUCY trial. *Medicine (Baltimore).* 2018;97(39):e12556.
425. Seah JCY, Tang JSN, Kitchen A, Gaillard F, Dixon AF. Chest Radiographs in Congestive Heart Failure: Visualizing Neural Network Learning. *Radiology.* 2018:180887.
426. Kashihara K. Nonlinear System Identification Based on Convolutional Neural Networks for Multiple Drug Interactions. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:1-4.

427. Etemadi M, Inan OT. Wearable ballistocardiogram and seismocardiogram systems for health and performance. *J Appl Physiol (1985)*. 2018;124(2):452-461.
428. Dogan MV, Grumbach IM, Michaelson JJ, Philibert RA. Integrated genetic and epigenetic prediction of coronary heart disease in the Framingham Heart Study. *PLoS One*. 2018;13(1):e0190549.
429. Yoon J, Zame WR, Banerjee A, Cadeiras M, Alaa AM, van der Schaar M. Personalized survival predictions via Trees of Predictors: An application to cardiac transplantation. *PLoS One*. 2018;13(3):e0194985.
430. Samad MD, Wehner GJ, Arbabshirani MR, et al. Predicting deterioration of ventricular function in patients with repaired tetralogy of Fallot using machine learning. *Eur Heart J Cardiovasc Imaging*. 2018;19(7):730-738.
431. Guaricci AI, Masci PG, Lorenzoni V, Schwitter J, Pontone G. CarDiac MagnEtic Resonance for Primary Prevention Implantable CardioVerter DebrillAtor ThErapy international registry: Design and rationale of the DERIVATE study. *Int J Cardiol*. 2018;261:223-227.
432. Nirschl JJ, Janowczyk A, Peyster EG, et al. A deep-learning classifier identifies patients with clinical heart failure using whole-slide images of H&E tissue. *PLoS One*. 2018;13(4):e0192726.
433. Ahmad T, Lund LH, Rao P, et al. Machine Learning Methods Improve Prognostication, Identify Clinically Distinct Phenotypes, and Detect Heterogeneity in Response to Therapy in a Large Cohort of Heart Failure Patients. *J Am Heart Assoc*. 2018;7(8).
434. Golas SB, Shibahara T, Agboola S, et al. A machine learning model to predict the risk of 30-day readmissions in patients with heart failure: a retrospective analysis of electronic medical records data. *BMC Med Inform Decis Mak*. 2018;18(1):44.
435. Zhang R, Ma S, Shanahan L, Munroe J, Horn S, Speedie S. Discovering and identifying New York heart association classification from electronic health records. *BMC Med Inform Decis Mak*. 2018;18(Suppl 2):48.
436. Puyol E, Ruijsink B, Gerber B, et al. Regional Multi-view Learning for Cardiac Motion Analysis: Application to Identification of Dilated Cardiomyopathy Patients. *IEEE Trans Biomed Eng*. 2018.
437. Sahli Costabal F, Choy JS, Sack KL, Guccione JM, Kassab GS, Kuhl E. Multiscale characterization of heart failure. *Acta Biomaterialia*. 2019;86:66-76.
438. Awan SE, Bennamoun M, Sohel F, Sanfilippo FM, Dwivedi G. Machine learning-based prediction of heart failure readmission or death: implications of choosing the right model and the right metrics. *ESC heart failure*. 2019;6(2):428-435.
439. Przewlocka-Kosmala M, Marwick TH, Dabrowski A, Kosmala W. Contribution of Cardiovascular Reserve to Prognostic Categories of Heart Failure With Preserved Ejection Fraction: A Classification Based on Machine Learning. *Journal of the American Society of Echocardiography: Official Publication of the American Society of Echocardiography*. 2019.
440. Maragatham G, Devi S. LSTM Model for Prediction of Heart Failure in Big Data. *Journal of Medical Systems*. 2019;43(5):111.
441. Al'Aref SJ, Singh G, van Rosendael AR, et al. Determinants of In-Hospital Mortality After Percutaneous Coronary Intervention: A Machine Learning Approach. *Journal of the American Heart Association*. 2019;8(5):e011160.
442. Chung NC, Mirza B, Choi H, et al. Unsupervised classification of multi-omics data during cardiac remodeling using deep learning. *Methods (San Diego, Calif)*. 2019.
443. Wang L, Zhou X. Detection of Congestive Heart Failure Based on LSTM-Based Deep Network via Short-Term RR Intervals. *Sensors (Basel, Switzerland)*. 2019;19(7).

444. Li R, He H, Fang S, et al. Time Series Characteristics of Serum Branched-Chain Amino Acids for Early Diagnosis of Chronic Heart Failure. *Journal of Proteome Research*. 2019.
445. Sheet D, Karamalis A, Eslami A, et al. Joint learning of ultrasonic backscattering statistical physics and signal confidence primal for characterizing atherosclerotic plaques using intravascular ultrasound. *Med Image Anal*. 2014;18(1):103-117.
446. Sun L, Li Y, Zhang YT, et al. A computer-aided diagnostic algorithm improves the accuracy of transesophageal echocardiography for left atrial thrombi: a single-center prospective study. *J Ultrasound Med*. 2014;33(1):83-91.
447. Takx RA, de Jong PA, Leiner T, et al. Automated coronary artery calcification scoring in non-gated chest CT: agreement and reliability. *PLoS One*. 2014;9(3):e91239.
448. Afshin M, Ben Ayed I, Punithakumar K, et al. Regional assessment of cardiac left ventricular myocardial function via MRI statistical features. *IEEE Trans Med Imaging*. 2014;33(2):481-494.
449. Sheet D, Karamalis A, Eslami A, et al. Hunting for necrosis in the shadows of intravascular ultrasound. *Comput Med Imaging Graph*. 2014;38(2):104-112.
450. Yamak D, Panse P, Pavlicek W, Boltz T, Akay M. Non-calcified coronary atherosclerotic plaque characterization by dual energy computed tomography. *IEEE J Biomed Health Inform*. 2014;18(3):939-945.
451. Wei J, Zhou C, Chan HP, et al. Computerized detection of noncalcified plaques in coronary CT angiography: evaluation of topological soft gradient prescreening method and luminal analysis. *Med Phys*. 2014;41(8):081901.
452. Sankaran S, Grady LJ, Taylor CA. Real-time sensitivity analysis of blood flow simulations to lumen segmentation uncertainty. *Med Image Comput Comput Assist Interv*. 2014;17(Pt 2):1-8.
453. Ralovich K, John M, Camus E, Navab N, Heimann T. 6DoF catheter detection, application to intracardiac echocardiography. *Med Image Comput Comput Assist Interv*. 2014;17(Pt 2):635-642.
454. Otaki Y, Hell M, Slomka PJ, et al. Relationship of epicardial fat volume from noncontrast CT with impaired myocardial flow reserve by positron emission tomography. *J Cardiovasc Comput Tomogr*. 2015;9(4):303-309.
455. Xiong G, Kola D, Heo R, Elmore K, Cho I, Min JK. Myocardial perfusion analysis in cardiac computed tomography angiographic images at rest. *Med Image Anal*. 2015;24(1):77-89.
456. Gopalakrishnan V, Menon PG, Madan S. cMRI-BED: A novel informatics framework for cardiac MRI biomarker extraction and discovery applied to pediatric cardiomyopathy classification. *Biomed Eng Online*. 2015;14 Suppl 2:S7.
457. Zhao W, Jenkins MW, Linderman GC, et al. 3-D Stent Detection in Intravascular OCT Using a Bayesian Network and Graph Search. *IEEE Trans Med Imaging*. 2015;34(7):1549-1561.
458. Sankaran S, Grady L, Taylor CA. Fast Computation of Hemodynamic Sensitivity to Lumen Segmentation Uncertainty. *IEEE Trans Med Imaging*. 2015;34(12):2562-2571.
459. Kang D, Dey D, Slomka PJ, et al. Structured learning algorithm for detection of nonobstructive and obstructive coronary plaque lesions from computed tomography angiography. *J Med Imaging (Bellingham)*. 2015;2(1):014003.
460. Born S, Sundermann SH, Russ C, et al. Stent Maps--Comparative Visualization for the Prediction of Adverse Events of Transcatheter Aortic Valve Implantations. *IEEE Trans Vis Comput Graph*. 2014;20(12):2704-2713.
461. Dey D, Diaz Zamudio M, Schuhbaeck A, et al. Relationship Between Quantitative Adverse Plaque Features From Coronary Computed Tomography Angiography and Downstream Impaired Myocardial Flow Reserve by <sup>13</sup>N-

- Ammonia Positron Emission Tomography: A Pilot Study. *Circ Cardiovasc Imaging*. 2015;8(10):e003255.
462. Macedo MM, Guimaraes WV, Galon MZ, Takimura CK, Lemos PA, Gutierrez MA. A bifurcation identifier for IV-OCT using orthogonal least squares and supervised machine learning. *Comput Med Imaging Graph*. 2015;46 Pt 2:237-248.
463. Emad O, Yassine IA, Fahmy AS. Automatic localization of the left ventricle in cardiac MRI images using deep learning. *Conf Proc IEEE Eng Med Biol Soc*. 2015;2015:683-686.
464. Shalev R, Bezerra HG, Ray S, Prabhu D, Wilson DL. Classification of calcium in intravascular OCT images for the purpose of intervention planning. *Proc SPIE Int Soc Opt Eng*. 2016;9786.
465. Nam HS, Kim CS, Lee JJ, Song JW, Kim JW, Yoo H. Automated detection of vessel lumen and stent struts in intravascular optical coherence tomography to evaluate stent apposition and neointimal coverage. *Med Phys*. 2016;43(4):1662.
466. Moghaddasi H, Nourian S. Automatic assessment of mitral regurgitation severity based on extensive textural features on 2D echocardiography videos. *Comput Biol Med*. 2016;73:47-55.
467. Yu SC, Liu W, Wong RH, Underwood M, Wang D. The Potential of Computational Fluid Dynamics Simulation on Serial Monitoring of Hemodynamic Change in Type B Aortic Dissection. *Cardiovasc Intervent Radiol*. 2016;39(8):1090-1098.
468. Leni PE, Laurent R, Salomon M, Gschwind R, Makovicka L, Henriët J. Development of a 4D numerical chest phantom with customizable breathing. *Phys Med*. 2016;32(6):795-800.
469. Ma Y, Wang L, Ma Y, Dong M, Du S, Sun X. An SPCNN-GVF-based approach for the automatic segmentation of left ventricle in cardiac cine MR images. *Int J Comput Assist Radiol Surg*. 2016;11(11):1951-1964.
470. Zhou H, Sun P, Ha S, Lundine D, Xiong G. Watertight modeling and segmentation of bifurcated Coronary arteries for blood flow simulation using CT imaging. *Comput Med Imaging Graph*. 2016;53:43-53.
471. Ciompi F, Balocco S, Rigla J, Carrillo X, Mauri J, Radeva P. Computer-aided detection of intracoronary stent in intravascular ultrasound sequences. *Med Phys*. 2016;43(10):5616.
472. Norlen A, Alven J, Molnar D, et al. Automatic pericardium segmentation and quantification of epicardial fat from computed tomography angiography. *J Med Imaging (Bellingham)*. 2016;3(3):034003.
473. Nasr-Esfahani E, Samavi S, Karimi N, et al. Vessel extraction in X-ray angiograms using deep learning. *Conf Proc IEEE Eng Med Biol Soc*. 2016;2016:643-646.
474. Alzubiedi S, Saleh MI. Pharmacogenetic-guided Warfarin Dosing Algorithm in African-Americans. *J Cardiovasc Pharmacol*. 2016;67(1):86-92.
475. Rodrigues EO, Morais FF, Morais NA, Conci LS, Neto LV, Conci A. A novel approach for the automated segmentation and volume quantification of cardiac fats on computed tomography. *Comput Methods Programs Biomed*. 2016;123:109-128.
476. Itu L, Rapaka S, Passerini T, et al. A machine-learning approach for computation of fractional flow reserve from coronary computed tomography. *J Appl Physiol (1985)*. 2016;121(1):42-52.
477. Wong KC, Tee M, Chen M, Bluemke DA, Summers RM, Yao J. Regional infarction identification from cardiac CT images: a computer-aided biomechanical approach. *Int J Comput Assist Radiol Surg*. 2016;11(9):1573-1583.
478. Lungu A, Swift AJ, Capener D, Kiely D, Hose R, Wild JM. Diagnosis of pulmonary hypertension from magnetic resonance imaging-based computational models and decision tree analysis. *Pulm Circ*. 2016;6(2):181-190.



479. Hagenah J, Werrmann E, Scharfschwerdt M, Ernst F, Metzner C. Prediction of individual prosthesis size for valve-sparing aortic root reconstruction based on geometric features. *Conf Proc IEEE Eng Med Biol Soc.* 2016;2016:3273-3276.
480. Liang L, Kong F, Martin C, et al. Machine learning-based 3-D geometry reconstruction and modeling of aortic valve deformation using 3-D computed tomography images. *Int J Numer Method Biomed Eng.* 2017;33(5).
481. Su S, Hu Z, Lin Q, Hau WK, Gao Z, Zhang H. An artificial neural network method for lumen and media-adventitia border detection in IVUS. *Comput Med Imaging Graph.* 2017;57:29-39.
482. Dawes TJW, de Marvao A, Shi W, et al. Machine Learning of Three-dimensional Right Ventricular Motion Enables Outcome Prediction in Pulmonary Hypertension: A Cardiac MR Imaging Study. *Radiology.* 2017;283(2):381-390.
483. Avendi MR, Kheradvar A, Jafarkhani H. Automatic segmentation of the right ventricle from cardiac MRI using a learning-based approach. *Magn Reson Med.* 2017;78(6):2439-2448.
484. Yu L, Guo Y, Wang Y, Yu J, Chen P. Segmentation of Fetal Left Ventricle in Echocardiographic Sequences Based on Dynamic Convolutional Neural Networks. *IEEE Trans Biomed Eng.* 2017;64(8):1886-1895.
485. Liao F, Chen X, Hu X, Song S. Estimation of the Volume of the Left Ventricle From MRI Images Using Deep Neural Networks. *IEEE Trans Cybern.* 2017.
486. Betancur J, Rubeaux M, Fuchs TA, et al. Automatic Valve Plane Localization in Myocardial Perfusion SPECT/CT by Machine Learning: Anatomic and Clinical Validation. *J Nucl Med.* 2017;58(6):961-967.
487. Freiman M, Nickisch H, Prevrhal S, et al. Improving CCTA-based lesions' hemodynamic significance assessment by accounting for partial volume modeling in automatic coronary lumen segmentation. *Med Phys.* 2017;44(3):1040-1049.
488. Wang J, Ding H, Bidgoli FA, et al. Detecting Cardiovascular Disease from Mammograms With Deep Learning. *IEEE Trans Med Imaging.* 2017;36(5):1172-1181.
489. Rodrigues EO, Pinheiro VHA, Liatsis P, Conci A. Machine learning in the prediction of cardiac epicardial and mediastinal fat volumes. *Comput Biol Med.* 2017;89:520-529.
490. Lian J, Shi B, Li M, Nan Z, Ma Y. An automatic segmentation method of a parameter-adaptive PCNN for medical images. *Int J Comput Assist Radiol Surg.* 2017;12(9):1511-1519.
491. Tan LK, Liew YM, Lim E, McLaughlin RA. Convolutional neural network regression for short-axis left ventricle segmentation in cardiac cine MR sequences. *Med Image Anal.* 2017;39:78-86.
492. Larroza A, Materka A, Lopez-Lereu MP, Monmeneu JV, Bodi V, Moratal D. Differentiation between acute and chronic myocardial infarction by means of texture analysis of late gadolinium enhancement and cine cardiac magnetic resonance imaging. *Eur J Radiol.* 2017;92:78-83.
493. Mazo C, Alegre E, Trujillo M. Classification of cardiovascular tissues using LBP based descriptors and a cascade SVM. *Comput Methods Programs Biomed.* 2017;147:1-10.
494. Molaei S, Shiri M, Horan K, Kahrobaei D, Nallamothu B, Najarian K. Deep Convolutional Neural Networks for left ventricle segmentation. *Conf Proc IEEE Eng Med Biol Soc.* 2017;2017:668-671.
495. Yang F, He Y, Hussain M, Xie H, Lei P. Convolutional Neural Network for the Detection of End-Diastole and End-Systole Frames in Free-Breathing Cardiac Magnetic Resonance Imaging. *Comput Math Methods Med.* 2017;2017:1640835.
496. Duguay TM, Tesche C, Vliedhart R, et al. Coronary Computed Tomographic Angiography-Derived Fractional Flow Reserve Based on Machine Learning for

- Risk Stratification of Non-Culprit Coronary Narrowings in Patients with Acute Coronary Syndrome. *Am J Cardiol.* 2017;120(8):1260-1266.
497. Nakajima K, Kudo T, Nakata T, et al. Diagnostic accuracy of an artificial neural network compared with statistical quantitation of myocardial perfusion images: a Japanese multicenter study. *Eur J Nucl Med Mol Imaging.* 2017;44(13):2280-2289.
498. Mengdi X, Jun C, Annan L, et al. Fibroatheroma identification in Intravascular Optical Coherence Tomography images using deep features. *Conf Proc IEEE Eng Med Biol Soc.* 2017;2017:1501-1504.
499. Yu L, Guo Y, Wang Y, Yu J, Chen P. Determination of Fetal Left Ventricular Volume Based on Two-Dimensional Echocardiography. *J Healthc Eng.* 2017;2017:4797315.
500. Jeganathan J, Knio Z, Amador Y, et al. Artificial intelligence in mitral valve analysis. *Ann Card Anaesth.* 2017;20(2):129-134.
501. Tabassian M, Alessandrini M, Herbots L, et al. Machine learning of the spatio-temporal characteristics of echocardiographic deformation curves for infarct classification. *Int J Cardiovasc Imaging.* 2017;33(8):1159-1167.
502. Yang D, Wu P, Tan C, Pohl KM, Axel L, Metaxas D. 3D Motion Modeling and Reconstruction of Left Ventricle Wall in Cardiac MRI. *Funct Imaging Model Heart.* 2017;10263:481-492.
503. Wolterink JM, Leiner T, Viergever MA, Isgum I. Generative Adversarial Networks for Noise Reduction in Low-Dose CT. *IEEE Trans Med Imaging.* 2017;36(12):2536-2545.
504. Xue W, Islam A, Bhaduri M, Li S. Direct Multitype Cardiac Indices Estimation via Joint Representation and Regression Learning. *IEEE Trans Med Imaging.* 2017;36(10):2057-2067.
505. Zahnd G, Hoogendoorn A, Combaret N, et al. Contour segmentation of the intima, media, and adventitia layers in intracoronary OCT images: application to fully automatic detection of healthy wall regions. *Int J Comput Assist Radiol Surg.* 2017;12(11):1923-1936.
506. Solecki M, Kruk M, Demkow M, et al. What is the optimal anatomic location for coronary artery pressure measurement at CT-derived FFR? *J Cardiovasc Comput Tomogr.* 2017;11(5):397-403.
507. Banchhor SK, Londhe ND, Araki T, et al. Wall-based measurement features provides an improved IVUS coronary artery risk assessment when fused with plaque texture-based features during machine learning paradigm. *Comput Biol Med.* 2017;91:198-212.
508. Tesche C, Vliegenthart R, Duguay TM, et al. Coronary Computed Tomographic Angiography-Derived Fractional Flow Reserve for Therapeutic Decision Making. *Am J Cardiol.* 2017;120(12):2121-2127.
509. Yong YL, Tan LK, McLaughlin RA, Chee KH, Liew YM. Linear-regression convolutional neural network for fully automated coronary lumen segmentation in intravascular optical coherence tomography. *J Biomed Opt.* 2017;22(12):1-9.
510. Dangi S, Linte CA, Yaniv Z. Cine Cardiac MRI Slice Misalignment Correction Towards Full 3D Left Ventricle Segmentation. *Proc SPIE Int Soc Opt Eng.* 2018;10576.
511. Wardziak L, Kruk M, Pleban W, et al. Coronary CTA enhanced with CTA based FFR analysis provides higher diagnostic value than invasive coronary angiography in patients with intermediate coronary stenosis. *J Cardiovasc Comput Tomogr.* 2018.
512. Abdolmanafi A, Duong L, Dahdah N, Adib IR, Cheriet F. Characterization of coronary artery pathological formations from OCT imaging using deep learning. *Biomed Opt Express.* 2018;9(10):4936-4960.
513. Mazo C, Bernal J, Trujillo M, Alegre E. Transfer learning for classification of cardiovascular tissues in histological images. *Comput Methods Programs Biomed.* 2018;165:69-76.

514. Yao J, Tridandapani S, Auffermann WF, Wick CA, Bhatti PT. An Adaptive Seismocardiography (SCG)-ECG Multimodal Framework for Cardiac Gating Using Artificial Neural Networks. *IEEE J Transl Eng Health Med.* 2018;6:1900611.
515. van Hamersvelt RW, Zreik M, Voskuil M, Viergever MA, Isgum I, Leiner T. Deep learning analysis of left ventricular myocardium in CT angiographic intermediate-degree coronary stenosis improves the diagnostic accuracy for identification of functionally significant stenosis. *Eur Radiol.* 2018.
516. Yang G, Chen J, Gao Z, et al. Multiview Sequential Learning and Dilated Residual Learning for a Fully Automatic Delineation of the Left Atrium and Pulmonary Veins from Late Gadolinium-Enhanced Cardiac MRI Images. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:1123-1127.
517. Moccia S, Banali R, Martini C, et al. Development and testing of a deep learning-based strategy for scar segmentation on CMR-LGE images. *MAGMA.* 2018.
518. Masuda T, Nakaura T, Funama Y, et al. Machine-learning integration of CT histogram analysis to evaluate the composition of atherosclerotic plaques: Validation with IB-IVUS. *J Cardiovasc Comput Tomogr.* 2018.
519. Kolluru C, Prabhu D, Gharaibeh Y, Bezerra H, Guagliumi G, Wilson D. Deep neural networks for A-line-based plaque classification in coronary intravascular optical coherence tomography images. *J Med Imaging (Bellingham).* 2018;5(4):044504.
520. Zreik M, van Hamersvelt RW, Wolterink JM, Leiner T, Viergever MA, Isgum I. A Recurrent CNN for Automatic Detection and Classification of Coronary Artery Plaque and Stenosis in Coronary CT Angiography. *IEEE Trans Med Imaging.* 2018.
521. Toth D, Miao S, Kurzendorfer T, et al. 3D/2D model-to-image registration by imitation learning for cardiac procedures. *Int J Comput Assist Radiol Surg.* 2018.
522. Yu M, Lu Z, Li W, Wei M, Yan J, Zhang J. CT morphological index provides incremental value to machine learning based CT-FFR for predicting hemodynamically significant coronary stenosis. *Int J Cardiol.* 2018;265:256-261.
523. Slomka PJ, Betancur J, Liang JX, et al. Rationale and design of the REgistry of Fast Myocardial Perfusion Imaging with NExt generation SPECT (REFINE SPECT). *J Nucl Cardiol.* 2018.
524. Jin C, Feng J, Wang L, et al. Left Atrial Appendage Segmentation Using Fully Convolutional Neural Networks and Modified Three-Dimensional Conditional Random Fields. *IEEE J Biomed Health Inform.* 2018;22(6):1906-1916.
525. Hu X, Yang M, Han L, Du Y. Diagnostic performance of machine-learning-based computed fractional flow reserve (FFR) derived from coronary computed tomography angiography for the assessment of myocardial ischemia verified by invasive FFR. *Int J Cardiovasc Imaging.* 2018;34(12):1987-1996.
526. Wong KCL, Syeda-Mahmood T, Moradi M. Building medical image classifiers with very limited data using segmentation networks. *Med Image Anal.* 2018;49:105-116.
527. Cano-Espinosa C, Gonzalez G, Washko GR, Cazorla M, Estepar RSJ. Automated Agatston Score Computation in non-ECG Gated CT Scans Using Deep Learning. *Proc SPIE Int Soc Opt Eng.* 2018;10574.
528. Parikh SA, Gomez R, Thirugnanasambandam M, et al. Decision Tree Based Classification of Abdominal Aortic Aneurysms Using Geometry Quantification Measures. *Ann Biomed Eng.* 2018;46(12):2135-2147.
529. Gessert N, Lutz M, Heyder M, et al. Automatic Plaque Detection in IVOCT Pullbacks Using Convolutional Neural Networks. *IEEE Trans Med Imaging.* 2018.

530. Dormer JD, Ma L, Halicek M, Reilly CM, Schreibmann E, Fei B. Heart Chamber Segmentation from CT Using Convolutional Neural Networks. *Proc SPIE Int Soc Opt Eng.* 2018;10578.
531. Garcia EV, Klein JL, Moncayo V, et al. Diagnostic performance of an artificial intelligence-driven cardiac-structured reporting system for myocardial perfusion SPECT imaging. *J Nucl Cardiol.* 2018.
532. Bai W, Sinclair M, Tarroni G, et al. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *J Cardiovasc Magn Reson.* 2018;20(1):65.
533. Xu C, Xu L, Gao Z, et al. Direct delineation of myocardial infarction without contrast agents using a joint motion feature learning architecture. *Med Image Anal.* 2018;50:82-94.
534. Sulas E, Ortu E, Raffo L, Urru M, Tumbarello R, Pani D. Automatic Recognition of Complete Atrioventricular Activity in Fetal Pulsed-Wave Doppler Signals. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:917-920.
535. Duchateau N, Sermesant M, Delingette H, Ayache N. Model-Based Generation of Large Databases of Cardiac Images: Synthesis of Pathological Cine MR Sequences From Real Healthy Cases. *IEEE Trans Med Imaging.* 2018;37(3):755-766.
536. Ouzir N, Basarab A, Liebgott H, Harbaoui B, Tourneret JY. Motion Estimation in Echocardiography Using Sparse Representation and Dictionary Learning. *IEEE Trans Image Process.* 2018;27(1):64-77.
537. Luo G, Dong S, Wang K, Zuo W, Cao S, Zhang H. Multi-Views Fusion CNN for Left Ventricular Volumes Estimation on Cardiac MR Images. *IEEE Trans Biomed Eng.* 2018;65(9):1924-1934.
538. Mannil M, von Spiczak J, Manka R, Alkadhi H. Texture Analysis and Machine Learning for Detecting Myocardial Infarction in Noncontrast Low-Dose Computed Tomography: Unveiling the Invisible. *Invest Radiol.* 2018;53(6):338-343.
539. Dey D, Gaur S, Ovrehus KA, et al. Integrated prediction of lesion-specific ischaemia from quantitative coronary CT angiography using machine learning: a multicentre study. *Eur Radiol.* 2018;28(6):2655-2664.
540. Larroza A, Lopez-Lereu MP, Monmeneu JV, et al. Texture analysis of cardiac cine magnetic resonance imaging to detect nonviable segments in patients with chronic myocardial infarction. *Med Phys.* 2018;45(4):1471-1480.
541. Kolluru C, Prabhu D, Gharaibeh Y, Wu H, Wilson DL. Voxel-based plaque classification in coronary intravascular optical coherence tomography images using decision trees. *Proc SPIE Int Soc Opt Eng.* 2018;10575.
542. Nakajima K, Okuda K, Watanabe S, et al. Artificial neural network retrained to detect myocardial ischemia using a Japanese multicenter database. *Ann Nucl Med.* 2018;32(5):303-310.
543. Dormer JD, Halicek M, Ma L, Reilly CM, Schreibmann E, Fei B. Convolutional Neural Networks for the Detection of Diseased Hearts Using CT Images and Left Atrium Patches. *Proc SPIE Int Soc Opt Eng.* 2018;10575.
544. Shao XN, Sun YJ, Xiao KT, et al. Texture analysis of magnetic resonance T1 mapping with dilated cardiomyopathy: A machine learning approach. *Medicine (Baltimore).* 2018;97(37):e12246.
545. Dong S, Luo G, Wang K, Cao S, Li Q, Zhang H. A Combined Fully Convolutional Networks and Deformable Model for Automatic Left Ventricle Segmentation Based on 3D Echocardiography. *Biomed Res Int.* 2018;2018:5682365.
546. Betancur JA, Hu LH, Commandeur F, et al. Deep Learning Analysis of Upright-Supine High-Efficiency SPECT Myocardial Perfusion Imaging for Prediction of Obstructive Coronary Artery Disease: A Multicenter Study. *J Nucl Med.* 2018.
547. Shibutani T, Nakajima K, Wakabayashi H, et al. Accuracy of an artificial neural network for detecting a regional abnormality in myocardial perfusion SPECT. *Ann Nucl Med.* 2018.

548. Narang A, Mor-Avi V, Prado A, et al. Machine learning based automated dynamic quantification of left heart chamber volumes. *Eur Heart J Cardiovasc Imaging*. 2018.
549. Hasegawa A, Lee Y, Takeuchi Y, Ichikawa K. [Automated Classification of Calcification and Stent on Computed Tomography Coronary Angiography Using Deep Learning]. *Nihon Hoshasen Gijutsu Gakkai Zasshi*. 2018;74(10):1138-1143.
550. Abdolmanafi A, Duong L, Cheriet F, Dahdah NS. Intra-slice motion correction of intravascular OCT images using deep features. *IEEE J Biomed Health Inform*. 2018.
551. Du X, Tang R, Yin S, Zhang Y, Li S. Direct Segmentation-based Full Quantification for Left Ventricle via Deep Multi-task Regression Learning Network. *IEEE J Biomed Health Inform*. 2018.
552. Mastrodicasa D, Albrecht MH, Schoepf UJ, et al. Artificial intelligence machine learning-based coronary CT fractional flow reserve (CT-FFRML): Impact of iterative and filtered back projection reconstruction techniques. *J Cardiovasc Comput Tomogr*. 2018.
553. Jun TJ, Kang SJ, Lee JG, et al. Automated detection of vulnerable plaque in intravascular ultrasound images. *Med Biol Eng Comput*. 2018.
554. Huang W, Huang L, Lin Z, et al. Coronary Artery Segmentation by Deep Learning Neural Networks on Computed Tomographic Coronary Angiographic Images. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:608-611.
555. Que Q, Tang Z, Wang R, et al. CardioXNet: Automated Detection for Cardiomegaly Based on Deep Learning. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:612-615.
556. Nasr-Esfahani M, Mohrekesh M, Akbari M, et al. Left Ventricle Segmentation in Cardiac MR Images Using Fully Convolutional Network. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:1275-1278.
557. Omar HA, Patra A, Domingos JS, Leeson P, Noblel AJ. Automated Myocardial Wall Motion Classification using Handcrafted Features vs a Deep CNN-based mapping. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:3140-3143.
558. Leng S, Yang X, Zhao X, et al. Computational Platform Based on Deep Learning for Segmenting Ventricular Endocardium in Long-axis Cardiac MR Imaging. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:4500-4503.
559. Ostvik A, Smistad E, Aase SA, Haugen BO, Lovstakken L. Real-Time Standard View Classification in Transthoracic Echocardiography Using Convolutional Neural Networks. *Ultrasound Med Biol*. 2019;45(2):374-384.
560. Xue Z, Long R, Jaeger S, Folio L, George Thoma R, Antani AS. Extraction of Aortic Knuckle Contour in Chest Radiographs Using Deep Learning. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:5890-5893.
561. Kang E, Koo HJ, Yang DH, Seo JB, Ye JC. Cycle-consistent adversarial denoising network for multiphase coronary CT angiography. *Med Phys*. 2018.
562. Lossau T, Nickisch H, Wissel T, et al. Motion artifact recognition and quantification in coronary CT angiography using convolutional neural networks. *Med Image Anal*. 2018;52:68-79.
563. Zhang L, Gooya A, Pereanez M, et al. Automatic Assessment of Full Left Ventricular Coverage in Cardiac Cine Magnetic Resonance Imaging with Fisher Discriminative 3D CNN. *IEEE Trans Biomed Eng*. 2018.
564. Huang C, Murugiah K, Mahajan S, et al. Enhancing the prediction of acute kidney injury risk after percutaneous coronary intervention using machine learning techniques: A retrospective cohort study. *PLoS Med*. 2018;15(11):e1002703.
565. Nous FMA, Coenen A, Boersma E, et al. Comparison of the Diagnostic Performance of Coronary Computed Tomography Angiography-Derived Fractional Flow Reserve in Patients With Versus Without Diabetes Mellitus (from the MACHINE Consortium). *Am J Cardiol*. 2018.

566. Kwon JM, Kim KH, Jeon KH, Park J. Deep learning for predicting in-hospital mortality among heart disease patients based on echocardiography. *Echocardiography*. 2018.
567. von Knebel Doeberitz PL, De Cecco CN, Schoepf UJ, et al. Coronary CT angiography-derived plaque quantification with artificial intelligence CT fractional flow reserve for the identification of lesion-specific ischemia. *Eur Radiol*. 2018.
568. Han D, Lee JH, Rizvi A, et al. Incremental role of resting myocardial computed tomography perfusion for predicting physiologically significant coronary artery disease: A machine learning approach. *J Nucl Cardiol*. 2018;25(1):223-233.
569. Isgum I, de Vos BD, Wolterink JM, et al. Automatic determination of cardiovascular risk by CT attenuation correction maps in Rb-82 PET/CT. *J Nucl Cardiol*. 2018;25(6):2133-2142.
570. Baessler B, Mannil M, Oebel S, Maintz D, Alkadhi H, Manka R. Subacute and Chronic Left Ventricular Myocardial Scar: Accuracy of Texture Analysis on Nonenhanced Cine MR Images. *Radiology*. 2018;286(1):103-112.
571. Oktay O, Ferrante E, Kamnitsas K, et al. Anatomically Constrained Neural Networks (ACNNs): Application to Cardiac Image Enhancement and Segmentation. *IEEE Trans Med Imaging*. 2018;37(2):384-395.
572. Xue W, Brahm G, Pandey S, Leung S, Li S. Full left ventricle quantification via deep multitask relationships learning. *Med Image Anal*. 2018;43:54-65.
573. Wojnarski CM, Roselli EE, Idrees JJ, et al. Machine-learning phenotypic classification of bicuspid aortopathy. *J Thorac Cardiovasc Surg*. 2018;155(2):461-469 e464.
574. Tan LK, McLaughlin RA, Lim E, Abdul Aziz YF, Liew YM. Fully automated segmentation of the left ventricle in cine cardiac MRI using neural network regression. *J Magn Reson Imaging*. 2018;48(1):140-152.
575. Winther HB, Hundt C, Schmidt B, et al. nu-net: Deep Learning for Generalized Biventricular Mass and Function Parameters Using Multicenter Cardiac MRI Data. *JACC Cardiovasc Imaging*. 2018;11(7):1036-1038.
576. Lessmann N, van Ginneken B, Zreik M, et al. Automatic Calcium Scoring in Low-Dose Chest CT Using Deep Neural Networks With Dilated Convolutions. *IEEE Trans Med Imaging*. 2018;37(2):615-625.
577. Betancur J, Commandeur F, Motlagh M, et al. Deep Learning for Prediction of Obstructive Disease From Fast Myocardial Perfusion SPECT: A Multicenter Study. *JACC Cardiovasc Imaging*. 2018;11(11):1654-1663.
578. Nakanishi R, Sankaran S, Grady L, et al. Automated estimation of image quality for coronary computed tomographic angiography using machine learning. *Eur Radiol*. 2018;28(9):4018-4026.
579. Gernaat SAM, van Velzen SGM, Koh V, et al. Automatic quantification of calcifications in the coronary arteries and thoracic aorta on radiotherapy planning CT scans of Western and Asian breast cancer patients. *Radiother Oncol*. 2018;127(3):487-492.
580. Baessler B, Mannil M, Maintz D, Alkadhi H, Manka R. Texture analysis and machine learning of non-contrast T1-weighted MR images in patients with hypertrophic cardiomyopathy-Preliminary results. *Eur J Radiol*. 2018;102:61-67.
581. Fu X, Liu T, Xiong Z, Smaill BH, Stiles MK, Zhao J. Segmentation of histological images and fibrosis identification with a convolutional neural network. *Comput Biol Med*. 2018;98:147-158.
582. Coenen A, Kim YH, Kruk M, et al. Diagnostic Accuracy of a Machine-Learning Approach to Coronary Computed Tomographic Angiography-Based Fractional Flow Reserve: Result From the MACHINE Consortium. *Circ Cardiovasc Imaging*. 2018;11(6):e007217.

583. Cruz-Aceves I, Cervantes-Sanchez F, Avila-Garcia MS. A Novel Multiscale Gaussian-Matched Filter Using Neural Networks for the Segmentation of X-Ray Coronary Angiograms. *J Healthc Eng.* 2018;2018:5812059.
584. Vigneault DM, Xie W, Ho CY, Bluemke DA, Noble JA. Omega-Net (Omega-Net): Fully automatic, multi-view cardiac MR detection, orientation, and segmentation with deep neural networks. *Med Image Anal.* 2018;48:95-106.
585. Commandeur F, Goeller M, Betancur J, et al. Deep Learning for Quantification of Epicardial and Thoracic Adipose Tissue From Non-Contrast CT. *IEEE Trans Med Imaging.* 2018;37(8):1835-1846.
586. Pavoni M, Chang Y, Park SH, Smedby O. Convolutional neural network-based image enhancement for x-ray percutaneous coronary intervention. *J Med Imaging (Bellingham).* 2018;5(2):024006.
587. Al WA, Jung HY, Yun ID, Jang Y, Park HB, Chang HJ. Automatic aortic valve landmark localization in coronary CT angiography using colonial walk. *PLoS One.* 2018;13(7):e0200317.
588. Qin C, Schlemper J, Caballero J, Price AN, Hajnal JV, Rueckert D. Convolutional Recurrent Neural Networks for Dynamic MR Image Reconstruction. *IEEE Trans Med Imaging.* 2019;38(1):280-290.
589. Chen J, Zhang H, Zhang W, Du X, Zhang Y, Li S. Correlated Regression Feature Learning for Automated Right Ventricle Segmentation. *IEEE J Transl Eng Health Med.* 2018;6:1800610.
590. Wang D, Zhang R, Zhu J, et al. Neural network fusion: a novel CT-MR Aortic Aneurysm image segmentation method. *Proc SPIE Int Soc Opt Eng.* 2018;10574.
591. Green M, Marom EM, Konen E, Kiryati N, Mayer A. 3-D Neural denoising for low-dose Coronary CT Angiography (CCTA). *Comput Med Imaging Graph.* 2018;70:185-191.
592. Strange G, Celermajer DS, Marwick T, et al. The National Echocardiography Database Australia (NEDA): Rationale and methodology. *Am Heart J.* 2018;204:186-189.
593. Yang H, Sun J, Li H, Wang L, Xu Z. Neural multi-atlas label fusion: Application to cardiac MR images. *Med Image Anal.* 2018;49:60-75.
594. Zotti C, Luo Z, Lalande A, Jodoin PM. Convolutional Neural Network with Shape Prior Applied to Cardiac MRI Segmentation. *IEEE J Biomed Health Inform.* 2018.
595. Fahmy AS, Rausch J, Neisius U, et al. Automated Cardiac MR Scar Quantification in Hypertrophic Cardiomyopathy Using Deep Convolutional Neural Networks. *JACC Cardiovasc Imaging.* 2018;11(12):1917-1918.
596. Bratt A, Kim J, Pollie M, et al. Machine learning derived segmentation of phase velocity encoded cardiovascular magnetic resonance for fully automated aortic flow quantification. *Journal of Cardiovascular Magnetic Resonance: Official Journal of the Society for Cardiovascular Magnetic Resonance.* 2019;21(1):1.
597. Togo R, Hirata K, Manabe O, et al. Cardiac sarcoidosis classification with deep convolutional neural network-based features using polar maps. *Comput Biol Med.* 2019;104:81-86.
598. Curiale AH, Colavecchia FD, Mato G. Automatic quantification of the LV function and mass: A deep learning approach for cardiovascular MRI. *Computer Methods and Programs in Biomedicine.* 2019;169:37-50.
599. Biswas S, Aggarwal HK, Jacob M. Dynamic MRI using model-based deep learning and STORM priors: MoDL-STORM. *Magnetic Resonance in Medicine.* 2019.
600. Xiong Z, Fedorov VV, Fu X, Cheng E, Macleod R, Zhao J. Fully Automatic Left Atrium Segmentation From Late Gadolinium Enhanced Magnetic Resonance Imaging Using a Dual Fully Convolutional Neural Network. *IEEE transactions on medical imaging.* 2019;38(2):515-524.

601. Bahrami N, Retson T, Blansit K, Wang K, Hsiao A. Automated selection of myocardial inversion time with a convolutional neural network: Spatial temporal ensemble myocardium inversion network (STEMI-NET). *Magnetic Resonance in Medicine*. 2019;81(5):3283-3291.
602. Goldfarb JW, Craft J, Cao JJ. Water-fat separation and parameter mapping in cardiac MRI via deep learning with a convolutional neural network. *Journal of magnetic resonance imaging: JMRI*. 2019.
603. Du X, Yin S, Tang R, Zhang Y, Li S. Cardiac-DeepIED: Automatic Pixel-Level Deep Segmentation for Cardiac Bi-Ventricle Using Improved End-to-End Encoder-Decoder Network. *IEEE journal of translational engineering in health and medicine*. 2019;7:1900110.
604. Khened M, Kollerathu VA, Krishnamurthi G. Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Med Image Anal*. 2019;51:21-45.
605. Wang T, Lei Y, Tang H, et al. A learning-based automatic segmentation and quantification method on left ventricle in gated myocardial perfusion SPECT imaging: A feasibility study. *Journal of Nuclear Cardiology: Official Publication of the American Society of Nuclear Cardiology*. 2019.
606. Duan J, Bello G, Schlemper J, et al. Automatic 3D bi-ventricular segmentation of cardiac images by a shape-refined multi-task deep learning approach. *IEEE transactions on medical imaging*. 2019.
607. Priya C, Sudha S. Adaptive Fruitfly Based Modified Region Growing Algorithm for Cardiac Fat Segmentation Using Optimal Neural Network. *Journal of Medical Systems*. 2019;43(5):104.
608. Hauptmann A, Arridge S, Lucka F, Muthurangu V, Steeden JA. Real-time cardiovascular MR with spatio-temporal artifact suppression using deep learning-proof of concept in congenital heart disease. *Magn Reson Med*. 2019;81(2):1143-1156.
609. Liu J, Xie H, Zhang S, Gu L. Multi-sequence myocardium segmentation with cross-constrained shape and neural network-based initialization. *Comput Med Imaging Graph*. 2019;71:49-57.
610. Zhao J, Ai D, Yang Y, et al. Deep feature regression (DFR) for 3D vessel segmentation. *Physics in Medicine and Biology*. 2019.
611. de Vos BD, Berendsen FF, Viergever MA, Sokooti H, Staring M, Išgum I. A deep learning framework for unsupervised affine and deformable image registration. *Medical Image Analysis*. 2019;52:128-143.
612. Wang Z-Q, Zhou Y-J, Zhao Y-X, et al. Diagnostic accuracy of a deep learning approach to calculate FFR from coronary CT angiography. *Journal of geriatric cardiology: JGC*. 2019;16(1):42-48.
613. Murphy EK, Amoh J, Arshad SH, Halter RJ, Odame KM. Noise-robust bioimpedance approach for cardiac output measurement. *Physiological Measurement*. 2019.
614. van Hamersvelt RW, Išgum I, de Jong PA, et al. Application of speCtraL computed tomogrAphy to impRove specIficity of cardiac compuTed tomographY (CLARITY study): rationale and design. *BMJ open*. 2019;9(3):e025793.
615. Wu D, Wang X, Bai J, et al. Automated anatomical labeling of coronary arteries via bidirectional tree LSTMs. *Int J Comput Assist Radiol Surg*. 2018.
616. Wolterink JM, van Hamersvelt RW, Viergever MA, Leiner T, Išgum I. Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier. *Med Image Anal*. 2019;51:46-60.
617. Tao Q, Yan W, Wang Y, et al. Deep Learning-based Method for Fully Automatic Quantification of Left Ventricle Function from Cine MR Images: A Multivendor, Multicenter Study. *Radiology*. 2019;290(1):81-88.



618. Leclerc S, Smistad E, Pedrosa J, et al. Deep Learning for Segmentation using an Open Large-Scale Dataset in 2D Echocardiography. *IEEE transactions on medical imaging*. 2019.
619. Volpato V, Mor-Avi V, Narang A, et al. Automated, machine learning-based, 3D echocardiographic quantification of left ventricular mass. *Echocardiography (Mount Kisco, NY)*. 2019;36(2):312-319.
620. Goldstein BA, Chang TI, Mitani AA, Assimes TL, Winkelmayr WC. Near-term prediction of sudden cardiac death in older hemodialysis patients using electronic health records. *Clin J Am Soc Nephrol*. 2014;9(1):82-91.
621. Ahmad SM, Busser BW, Huang D, et al. Machine learning classification of cell-specific cardiac enhancers uncovers developmental subnetworks regulating progenitor cell division and cell fate specification. *Development*. 2014;141(4):878-888.
622. I AW, Pratt NL, Kalisch LM, Roughead EE. Comparing time to adverse drug reaction signals in a spontaneous reporting database and a claims database: a case study of rofecoxib-induced myocardial infarction and rosiglitazone-induced heart failure signals in Australia. *Drug Saf*. 2014;37(1):53-64.
623. Liu KE, Lo CL, Hu YH. Improvement of adequate use of warfarin for the elderly using decision tree-based approaches. *Methods Inf Med*. 2014;53(1):47-53.
624. Perry TE, Zha H, Zhou K, Frias P, Zeng D, Braunstein M. Supervised embedding of textual predictors with applications in clinical diagnostics for pediatric cardiology. *J Am Med Inform Assoc*. 2014;21(e1):e136-142.
625. Gharehbaghi A, Dutoit T, Ask P, Sornmo L. Detection of systolic ejection click using time growing neural network. *Med Eng Phys*. 2014;36(4):477-483.
626. Elgendi M, Bobhate P, Jain S, et al. Spectral analysis of the heart sounds in children with and without pulmonary artery hypertension. *Int J Cardiol*. 2014;173(1):92-99.
627. Elgendi M, Bobhate P, Jain S, et al. Time-domain analysis of heart sound intensity in children with and without pulmonary artery hypertension: a pilot study using a digital stethoscope. *Pulm Circ*. 2014;4(4):685-695.
628. Nouei MT, Kamyad AV, Sarzaem M, Ghazalbash S. Developing a genetic fuzzy system for risk assessment of mortality after cardiac surgery. *J Med Syst*. 2014;38(10):102.
629. Gorospe G, Zhu R, Millrod MA, Zambidis ET, Tung L, Vidal R. Automated grouping of action potentials of human embryonic stem cell-derived cardiomyocytes. *IEEE Trans Biomed Eng*. 2014;61(9):2389-2395.
630. LaFaro RJ, Pothula S, Kubal KP, et al. Neural Network Prediction of ICU Length of Stay Following Cardiac Surgery Based on Pre-Incision Variables. *PLoS One*. 2015;10(12):e0145395.
631. Economou GP, Sourla E, Stamatopoulou KM, et al. Exploiting expert systems in cardiology: a comparative study. *Adv Exp Med Biol*. 2015;820:79-89.
632. Han X, Ge M, Dong J, Wang Z, He J, Yang R. The relationship between left ventricle myocardial performance index of healthy women and geographical factors. *Int J Biometeorol*. 2015;59(11):1557-1565.
633. Huang RS, Nedelcu E, Bai Y, et al. Post-operative bleeding risk stratification in cardiac pulmonary bypass patients using artificial neural network. *Ann Clin Lab Sci*. 2015;45(2):181-186.
634. Nahato KB, Harichandran KN, Arputharaj K. Knowledge mining from clinical datasets using rough sets and backpropagation neural network. *Comput Math Methods Med*. 2015;2015:460189.
635. Wise ES, Hocking KM, Brophy CM. Prediction of in-hospital mortality after ruptured abdominal aortic aneurysm repair using an artificial neural network. *J Vasc Surg*. 2015;62(1):8-15.
636. Lee EK, Kurokawa YK, Tu R, George SC, Khine M. Machine learning plus optical flow: a simple and sensitive method to detect cardioactive drugs. *Sci Rep*. 2015;5:11817.

637. Mendes RG, de Souza CR, Machado MN, et al. Predicting reintubation, prolonged mechanical ventilation and death in post-coronary artery bypass graft surgery: a comparison between artificial neural networks and logistic regression models. *Arch Med Sci.* 2015;11(4):756-763.
638. Liu T, Lin Z, Ong ME, et al. Manifold ranking based scoring system with its application to cardiac arrest prediction: A retrospective study in emergency department patients. *Comput Biol Med.* 2015;67:74-82.
639. Luo G, Stone BL, Sakaguchi F, Sheng X, Murtaugh MA. Using Computational Approaches to Improve Risk-Stratified Patient Management: Rationale and Methods. *JMIR Res Protoc.* 2015;4(4):e128.
640. Mohamadipanah H, Hoberock LL, Andalibi M. Predictive Model Reference Adaptive Controller to Compensate Heart Motion in Minimally Invasive CABG Surgery. *Cardiovasc Eng Technol.* 2015;6(3):329-339.
641. Gharehbaghi A, Dutoit T, Sepehri AA, Kocharian A, Linden M. A Novel Method for Screening Children with Isolated Bicuspid Aortic Valve. *Cardiovasc Eng Technol.* 2015;6(4):546-556.
642. Elgendi M, Bobhate P, Jain S, et al. The unique heart sound signature of children with pulmonary artery hypertension. *Pulm Circ.* 2015;5(4):631-639.
643. Bashir S, Qamar U, Khan FH. IntelliHealth: A medical decision support application using a novel weighted multi-layer classifier ensemble framework. *J Biomed Inform.* 2016;59:185-200.
644. Monsalve-Torra A, Ruiz-Fernandez D, Marin-Alonso O, Soriano-Paya A, Camacho-Mackenzie J, Carreno-Jaimes M. Using machine learning methods for predicting inhospital mortality in patients undergoing open repair of abdominal aortic aneurysm. *J Biomed Inform.* 2016;62:195-201.
645. Wang S, Sun H, Liu H, Li D, Li Y, Hou T. ADMET Evaluation in Drug Discovery. 16. Predicting hERG Blockers by Combining Multiple Pharmacophores and Machine Learning Approaches. *Mol Pharm.* 2016;13(8):2855-2866.
646. Antink CH, Leonhardt S, Walter M. Reducing false alarms in the ICU by quantifying self-similarity of multimodal biosignals. *Physiol Meas.* 2016;37(8):1233-1252.
647. Hu SB, Wong DJ, Correa A, Li N, Deng JC. Prediction of Clinical Deterioration in Hospitalized Adult Patients with Hematologic Malignancies Using a Neural Network Model. *PLoS One.* 2016;11(8):e0161401.
648. Seyres D, Ghavi-Helm Y, Junion G, et al. Identification and in silico modeling of enhancers reveals new features of the cardiac differentiation network. *Development.* 2016;143(23):4533-4542.
649. Hansen PW, Clemmensen L, Sehested TS, et al. Identifying Drug-Drug Interactions by Data Mining: A Pilot Study of Warfarin-Associated Drug Interactions. *Circ Cardiovasc Qual Outcomes.* 2016;9(6):621-628.
650. Thompson PC, Dalman RL, Harris EJ, Chandra V, Lee JT, Mell MW. Predictive models for mortality after ruptured aortic aneurysm repair do not predict futility and are not useful for clinical decision making. *J Vasc Surg.* 2016;64(6):1617-1622.
651. Waardenberg AJ, Homan B, Mohamed S, Harvey RP, Bouveret R. Prediction and validation of protein-protein interactors from genome-wide DNA-binding data using a knowledge-based machine-learning approach. *Open Biol.* 2016;6(9).
652. Ruiz-Fernandez D, Monsalve Torra A, Soriano-Paya A, Marin-Alonso O, Triana Palencia E. Aid decision algorithms to estimate the risk in congenital heart surgery. *Comput Methods Programs Biomed.* 2016;126:118-127.
653. Wolf MJ, Lee EK, Nicolson SC, et al. Rationale and methodology of a collaborative learning project in congenital cardiac care. *Am Heart J.* 2016;174:129-137.

654. Tang H, Zhang J, Chen H, Mondal A, Park Y. A non-invasive approach to investigation of ventricular blood pressure using cardiac sound features. *Physiol Meas.* 2017;38(2):289-309.
655. Kario K. Perfect 24-h management of hypertension: clinical relevance and perspectives. *J Hum Hypertens.* 2017;31(4):231-243.
656. Flores DL, Gomez C, Cervantes D, Abaroa A, Castro C, Castaneda-Martinez RA. Predicting the physiological response of Tivela stultorum hearts with digoxin from cardiac parameters using artificial neural networks. *Biosystems.* 2017;151:1-7.
657. Schiltz NK, Warner DF, Sun J, et al. Identifying Specific Combinations of Multimorbidity that Contribute to Health Care Resource Utilization: An Analytic Approach. *Med Care.* 2017;55(3):276-284.
658. Allyn J, Allou N, Augustin P, et al. A Comparison of a Machine Learning Model with EuroSCORE II in Predicting Mortality after Elective Cardiac Surgery: A Decision Curve Analysis. *PLoS One.* 2017;12(1):e0169772.
659. Wise ES, Stonko DP, Glaser ZA, et al. Prediction of Prolonged Ventilation after Coronary Artery Bypass Grafting: Data from an Artificial Neural Network. *Heart Surg Forum.* 2017;20(1):E007-E014.
660. Pereira F, Bueno A, Rodriguez A, et al. Automated detection of coarctation of aorta in neonates from two-dimensional echocardiograms. *J Med Imaging (Bellingham).* 2017;4(1):014502.
661. Li H, Luo M, Zheng J, et al. An artificial neural network prediction model of congenital heart disease based on risk factors: A hospital-based case-control study. *Medicine (Baltimore).* 2017;96(6):e6090.
662. Zhang J, Tang H. [Research on prediction method of left ventricular blood pressure based on external heart sounds]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi.* 2017;34(3):335-341.
663. Gharehbaghi A, Linden M, Babic A. A Decision Support System for Cardiac Disease Diagnosis Based on Machine Learning Methods. *Stud Health Technol Inform.* 2017;235:43-47.
664. Radchenko EV, Rulev YA, Safanyaev AY, Palyulin VA, Zefirov NS. Computer-aided estimation of the hERG-mediated cardiotoxicity risk of potential drug components. *Dokl Biochem Biophys.* 2017;473(1):128-131.
665. Eslamizadeh G, Barati R. Heart murmur detection based on wavelet transformation and a synergy between artificial neural network and modified neighbor annealing methods. *Artif Intell Med.* 2017;78:23-40.
666. Sudha M. Evolutionary and Neural Computing Based Decision Support System for Disease Diagnosis from Clinical Data Sets in Medical Practice. *J Med Syst.* 2017;41(11):178.
667. Awad A, Bader-El-Den M, McNicholas J, Briggs J. Early hospital mortality prediction of intensive care unit patients using an ensemble learning approach. *Int J Med Inform.* 2017;108:185-195.
668. Xiao H, Qasem A, Butlin M, Avolio A. Estimation of aortic systolic blood pressure from radial systolic and diastolic blood pressures alone using artificial neural networks. *J Hypertens.* 2017;35(8):1577-1585.
669. Indja B, Fanning JP, Maller JJ, et al. Neural network imaging to characterize brain injury in cardiac procedures: the emerging utility of connectomics. *Br J Anaesth.* 2017;118(5):680-688.
670. Liang L, Liu M, Martin C, Elefteriades JA, Sun W. A machine learning approach to investigate the relationship between shape features and numerically predicted risk of ascending aortic aneurysm. *Biomech Model Mechanobiol.* 2017;16(5):1519-1533.
671. Abdollahpur M, Ghaffari A, Ghiasi S, Mollakazemi MJ. Detection of pathological heart sounds. *Physiol Meas.* 2017;38(8):1616-1630.

672. [Quantitative structure-activity relationship model for prediction of cardiotoxicity of chemical components in traditional Chinese medicines]. *Beijing Da Xue Xue Bao Yi Xue Ban.* 2017;49(3):551-556.
673. Gharehbaghi A, Sepehri AA, Linden M, Babic A. Intelligent Phonocardiography for Screening Ventricular Septal Defect Using Time Growing Neural Network. *Stud Health Technol Inform.* 2017;238:108-111.
674. Silva S, Ait Aissa D, Cocquet P, et al. Combined Thoracic Ultrasound Assessment during a Successful Weaning Trial Predicts Postextubation Distress. *Anesthesiology.* 2017;127(4):666-674.
675. Ebert LC, Heimer J, Schweitzer W, et al. Automatic detection of hemorrhagic pericardial effusion on PMCT using deep learning - a feasibility study. *Forensic Sci Med Pathol.* 2017;13(4):426-431.
676. Kay E, Agarwal A. DropConnected neural networks trained on time-frequency and inter-beat features for classifying heart sounds. *Physiol Meas.* 2017;38(8):1645-1657.
677. Kublanov VS, Dolganov AY, Belo D, Gamboa H. Comparison of Machine Learning Methods for the Arterial Hypertension Diagnostics. *Appl Bionics Biomech.* 2017;2017:5985479.
678. Attallah O, Karthikesalingam A, Holt PJ, et al. Using multiple classifiers for predicting the risk of endovascular aortic aneurysm repair re-intervention through hybrid feature selection. *Proc Inst Mech Eng H.* 2017;231(11):1048-1063.
679. Liang L, Liu M, Sun W. A deep learning approach to estimate chemically-treated collagenous tissue nonlinear anisotropic stress-strain responses from microscopy images. *Acta Biomater.* 2017;63:227-235.
680. Conant G, Ahadian S, Zhao Y, Radisic M. Kinase inhibitor screening using artificial neural networks and engineered cardiac biowires. *Sci Rep.* 2017;7(1):11807.
681. Vivekanandan T, Sriman Narayana Iyengar NC. Optimal feature selection using a modified differential evolution algorithm and its effectiveness for prediction of heart disease. *Comput Biol Med.* 2017;90:125-136.
682. Lee EK, Tran DD, Keung W, et al. Machine Learning of Human Pluripotent Stem Cell-Derived Engineered Cardiac Tissue Contractility for Automated Drug Classification. *Stem Cell Reports.* 2017;9(5):1560-1572.
683. Bergau DM, Liu C, Magin RL, Lu H. Machine-Learning Prediction of Drug-Induced Cardiac Arrhythmia: Analysis of Gene Expression and Clustering. *Crit Rev Biomed Eng.* 2018;46(3):245-275.
684. Amiriparian S, Schmitt M, Cummins N, Qian K, Dong F, Schuller B. Deep Unsupervised Representation Learning for Abnormal Heart Sound Classification. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:4776-4779.
685. Dey J, Gaurav A, Tiwari VN. InstaBP: Cuff-less Blood Pressure Monitoring on Smartphone using Single PPG Sensor. *Conf Proc IEEE Eng Med Biol Soc.* 2018;2018:5002-5005.
686. Thompson WR, Reinisch AJ, Unterberger MJ, Schriebl AJ. Artificial Intelligence-Assisted Auscultation of Heart Murmurs: Validation by Virtual Clinical Trial. *Pediatr Cardiol.* 2018.
687. Cai C, Fang J, Guo P, et al. In Silico Pharmacoepidemiologic Evaluation of Drug-Induced Cardiovascular Complications Using Combined Classifiers. *J Chem Inf Model.* 2018;58(5):943-956.
688. Pan X, Hu X, Zhang YH, et al. Identifying Patients with Atrioventricular Septal Defect in Down Syndrome Populations by Using Self-Normalizing Neural Networks and Feature Selection. *Genes (Basel).* 2018;9(4).
689. Elgendi M, Bobhate P, Jain S, et al. The Voice of the Heart: Vowel-Like Sound in Pulmonary Artery Hypertension. *Diseases.* 2018;6(2).
690. Burghardt TP, Ajtai K. Neural/Bayes network predictor for inheritable cardiac disease pathogenicity and phenotype. *J Mol Cell Cardiol.* 2018;119:19-27.

691. Zhu K, Fu Z, Tao P, Zhu S. [Left ventricle segmentation in echocardiography based on adaptive mean shift]. *Sheng Wu Yi Xue Gong Cheng Xue Za Zhi*. 2018;35(2):273-279.
692. Maharlou H, Niakan Kalhori SR, Shahbazi S, Ravangard R. Predicting Length of Stay in Intensive Care Units after Cardiac Surgery: Comparison of Artificial Neural Networks and Adaptive Neuro-fuzzy System. *Healthc Inform Res*. 2018;24(2):109-117.
693. Liem DA, Murali S, Sigdel D, et al. Phrase mining of textual data to analyze extracellular matrix protein patterns across cardiovascular disease. *Am J Physiol Heart Circ Physiol*. 2018;315(4):H910-H924.
694. Juhola M, Joutsijoki H, Penttinen K, Aalto-Setälä K. Detection of genetic cardiac diseases by Ca(2+) transient profiles using machine learning methods. *Sci Rep*. 2018;8(1):9355.
695. McCarthy CP, Ibrahim NE, van Kimmenade RRJ, et al. A clinical and proteomics approach to predict the presence of obstructive peripheral arterial disease: From the Catheter Sampled Blood Archive in Cardiovascular Diseases (CASABLANCA) Study. *Clin Cardiol*. 2018;41(7):903-909.
696. Hatib F, Jian Z, Buddi S, et al. Machine-learning Algorithm to Predict Hypotension Based on High-fidelity Arterial Pressure Waveform Analysis. *Anesthesiology*. 2018;129(4):663-674.
697. Parreco J, Hidalgo A, Parks JJ, Kozol R, Rattan R. Using artificial intelligence to predict prolonged mechanical ventilation and tracheostomy placement. *J Surg Res*. 2018;228:179-187.
698. Giri S, Bose JC, Chandrasekar A, Tiwary BK, Gajalakshmi P, Chatterjee S. Increased Plasma Nitrite and von Willebrand Factor Indicates Early Diagnosis of Vascular Diseases in Chemotherapy Treated Cancer Patients. *Cardiovasc Toxicol*. 2018.
699. Kan CD, Wang JN, Lin CH, et al. Handmade trileaflet valve design and validation for patch-valved conduit reconstruction using generalized regression machine learning model. *Technol Health Care*. 2018;26(4):605-620.
700. Rosler AM, Fraportti J, Nectoux P, et al. Development and Application of a System Based on Artificial Intelligence for Transcatheter Aortic Prosthesis Selection. *Braz J Cardiovasc Surg*. 2018;33(4):391-397.
701. Luo Y, Mao C, Yang Y, et al. Integrating hypertension phenotype and genotype with hybrid non-negative matrix factorization. *Bioinformatics*. 2018.
702. Luis Ahumadal M, Peck J, Guerra J, et al. Prediction of One-Year Transplant-Free Survival after Norwood Procedure Based on the Pre-Operative Data. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:3995-3998.
703. Lu H, Zhang H, Lin Z, Huat NS. A Novel Deep Learning based Neural Network for Heartbeat Detection in Ballistocardiograph. *Conf Proc IEEE Eng Med Biol Soc*. 2018;2018:2563-2566.
704. Chen HL, Yu SJ, Xu Y, et al. Artificial Neural Network: A Method for Prediction of Surgery-Related Pressure Injury in Cardiovascular Surgical Patients. *J Wound Ostomy Continence Nurs*. 2018;45(1):26-30.
705. Aviles-Rivero AI, Alsaleh SM, Casals A. Sliding to predict: vision-based beating heart motion estimation by modeling temporal interactions. *Int J Comput Assist Radiol Surg*. 2018;13(3):353-361.
706. Kang SH, Joe B, Yoon Y, Cho GY, Shin I, Suh JW. Cardiac Auscultation Using Smartphones: Pilot Study. *JMIR Mhealth Uhealth*. 2018;6(2):e49.
707. Lluçia-Valldeperas A, Bragos R, Soler-Botija C, et al. Unravelling the effects of mechanical physiological conditioning on cardiac adipose tissue-derived progenitor cells in vitro and in silico. *Sci Rep*. 2018;8(1):499.
708. Matam BR, Duncan H, Lowe D. Machine learning based framework to predict cardiac arrests in a paediatric intensive care unit : Prediction of cardiac arrests. *J Clin Monit Comput*. 2018.

709. Jordanski M, Radovic M, Milosevic Z, Filipovic N, Obradovic Z. Machine Learning Approach for Predicting Wall Shear Distribution for Abdominal Aortic Aneurysm and Carotid Bifurcation Models. *IEEE J Biomed Health Inform.* 2018;22(2):537-544.
710. Diaz J, Acosta J, Gonzalez R, Cota J, Sifuentes E, Nebot A. Modeling the control of the central nervous system over the cardiovascular system using support vector machines. *Comput Biol Med.* 2018;93:75-83.
711. Feng J, Huang Z, Zhou C, Ye X. Study of continuous blood pressure estimation based on pulse transit time, heart rate and photoplethysmography-derived hemodynamic covariates. *Australas Phys Eng Sci Med.* 2018;41(2):403-413.
712. Wacker S, Noskov SY. Performance of Machine Learning Algorithms for Qualitative and Quantitative Prediction Drug Blockade of hERG1 channel. *Comput Toxicol.* 2018;6:55-63.
713. Siramshetty VB, Chen Q, Devarakonda P, Preissner R. The Catch-22 of Predicting hERG Blockade Using Publicly Accessible Bioactivity Data. *J Chem Inf Model.* 2018;58(6):1224-1233.
714. Lu P, Guo S, Zhang H, et al. Research on Improved Depth Belief Network-Based Prediction of Cardiovascular Diseases. *J Healthc Eng.* 2018;2018:8954878.
715. Vervier K, Michaelson JJ. TiSAn: estimating tissue-specific effects of coding and non-coding variants. *Bioinformatics.* 2018;34(18):3061-3068.
716. Bozkurt B, Germanakis I, Stylianou Y. A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection. *Comput Biol Med.* 2018;100:132-143.
717. Messner E, Zohrer M, Pernkopf F. Heart Sound Segmentation-An Event Detection Approach Using Deep Recurrent Neural Networks. *IEEE Trans Biomed Eng.* 2018;65(9):1964-1974.
718. Mejia OAV, Antunes MJ, Goncharov M, et al. Predictive performance of six mortality risk scores and the development of a novel model in a prospective cohort of patients undergoing valve surgery secondary to rheumatic fever. *PLoS One.* 2018;13(7):e0199277.
719. Marateb HR, Mohebian MR, Javanmard SH, et al. Prediction of dyslipidemia using gene mutations, family history of diseases and anthropometric indicators in children and adolescents: The CASPIAN-III study. *Comput Struct Biotechnol J.* 2018;16:121-130.
720. Jalali A, Simpao AF, Galvez JA, Licht DJ, Nataraj C. Prediction of Periventricular Leukomalacia in Neonates after Cardiac Surgery Using Machine Learning Algorithms. *J Med Syst.* 2018;42(10):177.
721. Cai C, Guo P, Zhou Y, et al. Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity. *Journal of Chemical Information and Modeling.* 2019;59(3):1073-1084.
722. Diller G-P, Babu-Narayan S, Li W, et al. Utility of machine learning algorithms in assessing patients with a systemic right ventricle. *European Heart Journal Cardiovascular Imaging.* 2019.
723. Harford S, Darabi H, Del Rios M, et al. A machine learning based model for Out of Hospital cardiac arrest outcome classification and sensitivity analysis. *Resuscitation.* 2019;138:134-140.
724. Blomberg SN, Folke F, Ersbøll AK, et al. Machine learning as a supportive tool to recognize cardiac arrest in emergency calls. *Resuscitation.* 2019.
725. Casclang-Verzosa G, Shrestha S, Khalil MJ, et al. Network Tomography for Understanding Phenotypic Presentations in Aortic Stenosis. *JACC Cardiovascular imaging.* 2019;12(2):236-248.
726. Bello GA, Dawes TJW, Duan J, et al. Deep learning cardiac motion analysis for human survival prediction. *Nature Machine Intelligence.* 2019;1:95-104.
727. Huo D, Kou B, Zhou Z, Lv M. A machine learning model to classify aortic dissection patients in the early diagnosis phase. *Scientific Reports.* 2019;9(1):2701.

728. Diller G-P, Kempny A, Babu-Narayan SV, et al. Machine learning algorithms estimating prognosis and guiding therapy in adult congenital heart disease: data from a single tertiary centre including 10 019 patients. *European Heart Journal*. 2019;40(13):1069-1077.
729. Zhao J, Feng Q, Wu P, et al. Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction. *Scientific Reports*. 2019;9(1):717.
730. Daghistani TA, Elshawi R, Sakr S, Ahmed AM, Al-Thwayee A, Al-Mallah MH. Predictors of in-hospital length of stay among cardiac patients: A machine learning approach. *International Journal of Cardiology*. 2019.
731. Sparapani R, Dabbouseh NM, Gutterman D, et al. Detection of Left Ventricular Hypertrophy Using Bayesian Additive Regression Trees: The MESA. *Journal of the American Heart Association*. 2019;8(5):e009959.
732. Pulido M, Melin P, Prado-Arechiga G. Blood Pressure Classification Using the Method of the Modular Neural Networks. *International Journal of Hypertension*. 2019;2019:7320365.
733. Ibrahim NE, McCarthy CP, Shrestha S, et al. A clinical, proteomics, and artificial intelligence-driven model to predict acute kidney injury in patients undergoing coronary angiography. *Clinical Cardiology*. 2019;42(2):292-298.
734. van Assen M, De Cecco CN, Eid M, et al. Prognostic value of CT myocardial perfusion imaging and CT-derived fractional flow reserve for major adverse cardiac events in patients with coronary artery disease. *Journal of Cardiovascular Computed Tomography*. 2019.
735. Ruiz VM, Saenz L, Lopez-Magallon A, et al. Early prediction of critical events for infants with single-ventricle physiology in critical care using routinely collected data. *The Journal of Thoracic and Cardiovascular Surgery*. 2019.
736. Ris T, Teixeira-Carvalho A, Coelho RMP, et al. Inflammatory biomarkers in infective endocarditis: machine learning to predict mortality. *Clinical and Experimental Immunology*. 2019.
737. Hsich EM, Thuita L, McNamara DM, et al. Variables of importance in the Scientific Registry of Transplant Recipients database predictive of heart transplant waitlist mortality. *American Journal of Transplantation: Official Journal of the American Society of Transplantation and the American Society of Transplant Surgeons*. 2019.
738. Emre İE, Erol N, Ayhan Yİ, Özkan Y, Erol Ç. The analysis of the effects of acute rheumatic fever in childhood on cardiac disease with data mining. *International Journal of Medical Informatics*. 2019;123:68-75.
739. Sun X, Feng Z, Yuan X, Zhang W, Rong P. [Radiomics strategy based on cardiac magnetic resonance imaging cine sequence for assessing the severity of mitral valve regurgitation]. *Zhong Nan Da Xue Xue Bao Yi Xue Ban = Journal of Central South University Medical Sciences*. 2019;44(3):290-296.

## **Anexo 6. Aprobación CEIm proyecto ANGIO**

### **Resumen proyecto ANGIO**

Este proyecto pretende desarrollar sistemas de inteligencia artificial para la detección, clasificación y predicción de evolución de patologías cardiacas diagnosticables a través de angiografía coronaria. En concreto, se aprovechará la información disponible en forma de imagen angiográfica, parámetros angiográficos e historia clínica de pacientes con diversas patologías cardiacas del Hospital Universitario de Salamanca (HUS). Para ello se utilizarán datos retrospectivos almacenados en la base de datos Intellispace Cardiovascular (ISCV) del Servicio de Cardiología del HUS y datos también retrospectivos procedentes del estudio ya publicado **GRACIA-3** (NCT00306228), en el cual participó el HUS. **El punto de partida del proyecto consistiría en el estudio de la evolución de la reestenosis en pacientes sometidos a intervención coronaria percutánea.**

Anexo 6. Aprobación ANGIO. PDF 1 cara.



**EL COMITE DE ETICA DE LA INVESTIGACION CON MEDICAMENTOS DEL  
AREA DE SALUD DE SALAMANCA,**

**INFORMA**

Que el Proyecto de Investigación presentado por

D. Pedro Luis Sánchez Fernández

Titulado:

**ANGIO. Sistema de detección y predicción de patologías cardiacas a partir de angiografía  
coronaria basado en inteligencia artificial**

Que presenta como Investigadora responsable, SE AJUSTA A LAS NORMAS ÉTICAS Y DE  
BUENA PRÁCTICA CLÍNICA, establecidas para tales estudios.

Código CEIm: PI 2019 02 178

Y para que conste, lo firma en Salamanca con fecha 12 de febrero de 2019

MARIA BELEN  
VIDRIALES  
VICENTE - DNI  
07851455Z

Digitally signed by MARIA  
BELEN VIDRIALES VICENTE  
- DNI 07851455Z  
Date: 2019.02.12 18:51:18  
+01'00'

LA SECRETARIA TÉCNICA

Fdo.: Dra. D.<sup>a</sup> María Belén Vidriales Vicente

