

Predicting Cardiovascular Disease Using Machine Learning: A Systematic Review

Josue Max Suasnabar Perez¹; Kevin Junior Sulca Gomez¹; Rose Mary Lozada Flores¹; Katherine Bernardo Herrera¹
¹Universidad Tecnológica del Perú, Perú, U20200226@utp.edu.pe, U20212200@utp.edu.pe, C16321@utp.edu.pe, C28608@utp.edu.pe

Abstract– Artificial intelligence (AI) is currently used in various techniques for disease prediction, including cardiovascular disease (CVD). Therefore, this review aims to demonstrate the most effective AI technology methods for the prediction of CVD risk factors and diagnoses during early detection in humans, compared to traditional methods such as electrocardiogram (ECG) signal reporters, intensive care unit (ICU) review, and echocardiographic examinations. PICO and PRISMA methodologies were used for the search and selection of relevant documents. Along these lines, 397 documents were identified, including articles and systematic reviews in the following databases: Scopus and IEEE. According to the inclusion and exclusion criteria, 26 open access articles were selected. Where, the use of AI technological methods allowed analyzing and capturing predictive values on people with or without CVD problems in an effective way, achieving a range greater than 0.99. It should be noted that the representations were made with spreadsheet tools and the data manipulation, analysis and graph generation with Python libraries. Finally, it is concluded that the most effective AI technological methods for CVD prediction are based on machine learning (ML) techniques for predicting measurement values such as sensitivity, specificity, precision, F1 score and area under the curve (AUROC), with measurement indices between 0.97 and 0.99, compared to deep learning (DL) techniques, whose indices are between 0.88 and 0.90.

Keywords– Artificial Intelligence, Cardiovascular Diseases, Machine Learning, Predictive Values, Risk Factors.

Predicción de enfermedades cardiovasculares mediante el uso de Aprendizaje Automático: Una Revisión Sistemática

Josue Max Suasnabar Perez¹; Kevin Junior Sulca Gomez¹; Rose Mary Lozada Flores¹; Katherine Bernardo Herrera¹

¹Universidad Tecnológica del Perú, Perú, U20200226@utp.edu.pe, U20212200@utp.edu.pe, C16321@utp.edu.pe.

C28608@utp.edu.pe

Resumen– La inteligencia artificial (IA) es empleada en diversas técnicas en la actualidad para la predicción de enfermedades, entre ellas las enfermedades cardiovasculares (ECV). Por ello, esta revisión tiene como objetivo evidenciar los métodos tecnológicos de IA más efectivos para la predicción de los factores de riesgo y diagnósticos de las ECV durante la detección temprana en personas, en comparación con las mediciones de métodos tradicionales, tales como los repositorios de señales de electrocardiogramas (ECG), revisión de unidad de cuidados intensivos (UCI) y exámenes ecocardiográficos. Se utilizaron las metodologías PICO y PRISMA para la búsqueda y selección de documentos relevantes. En esa línea, se identificaron 397 documentos, entre artículos y revisiones sistemáticas en las bases de datos: Scopus y IEEE. De acuerdo con los criterios de inclusión y exclusión fueron seleccionados 26 artículos de acceso abierto. Donde, el uso de los métodos tecnológicos de IA permitió analizar y capturar valores predictivos sobre personas con o sin problemas de ECV de forma efectiva, logrando un alcance mayor de 0.99. Cabe señalar que las representaciones se realizaron con herramientas de hojas de cálculo y la manipulación, análisis de datos y generación de gráficos con las librerías en Python. Finalmente, se concluye que los métodos tecnológicos de IA más efectivos para la predicción de ECV se basan en las técnicas de aprendizaje automático (ML) para predecir valores de medición como la sensibilidad, especificidad, precisión, puntaje F1 y el área bajo la curva (AUROC), con índices de medición comprendidos entre 0.97 y 0.99, en comparación con las técnicas de aprendizaje profundo (DL), cuyos índices se sitúan entre 0.88 y 0.90.

Palabras clave– Inteligencia Artificial, Enfermedades Cardiovasculares, Aprendizaje Automático, Valores Predictivos, Factores de Riesgo.

I. INTRODUCCIÓN

En la actualidad, el aprendizaje automático (ML) se ha convertido en un modelo tecnológico ampliamente utilizado en diversos ámbitos para la predicción de mediciones a partir de datos estructurados. Este enfoque emplea técnicas avanzadas de aprendizaje que implican el uso de complejas aproximaciones matemáticas y el análisis de grandes series de datos [1]. Debido a su eficacia en las predicciones, múltiples investigaciones han evidenciado el uso de estos modelos tanto en métodos de aprendizaje profundo (DL) [1-8] como en métodos de aprendizaje automático (ML) [4, 9-19], aplicados a la predicción de diversas enfermedades, incluyendo las enfermedades cardiovasculares (ECV). Las técnicas de ML permiten manejar diferentes estructuras de datos para extraer

características útiles que faciliten el entrenamiento del modelo. Estos métodos requieren volúmenes significativos de datos etiquetados para alcanzar una mayor precisión en aplicaciones de atención clínica [1] y [2]. La recopilación de datos suele incluir mediciones en personas [2, 9, 17, 18] y el uso de conjuntos de datos o registros clínicos [1, 3-8, 10, 12-16, 19]. Sin embargo, un desafío clave en la atención clínica es la integración de estas técnicas tradicionales con la información existente, así como el procesamiento y entrenamiento de los métodos tecnológicos para ofrecer atenciones cómodas a los pacientes mediante dispositivos tecnológicos modernos [11, 17, 20]. Las intervenciones clínicas apoyadas por tecnología moderna pueden mejorar significativamente las predicciones de las ECV, especialmente mediante el uso de dispositivos móviles [3, 7, 17] y dispositivos portátiles [18, 20], que facilitan una detección temprana accesible tanto para pacientes con diagnóstico como para aquellos con factores de riesgo. Esta revisión sistemática busca evidenciar los modelos de ML a través de diversos tipos de muestras recolectadas en los documentos revisados. El objetivo es identificar los modelos de ML más eficientes para la predicción de ECV, evaluando factores de riesgo y diagnósticos mediante métodos de DL y ML. Para ello, se emplean la metodología de preguntas PICO para una búsqueda adecuada de documentos y el enfoque PRISMA para la selección correcta de los artículos.

II. METODOLOGÍA

En la revisión se utilizó la metodología PICO para identificar el objeto de estudio y PRISMA para la selección de artículos que se utilizaron para el análisis.

A. Formulación de PICO

La herramienta PICO permite la elaboración de preguntas basadas en la problemática planteada y su formulación responde a preguntas adecuadas y respaldadas que enfocan el tema de interés en la búsqueda de registros. Esta herramienta utilizó los elementos: Problema (P), Intervención (I), Comparación (C) y Resultados (Outcomes, en inglés). Gracias a esta, se logró una selección precisa de artículos y facilitó la secuencia de las fases de esta metodología.

1) *Identificación de componentes PICO:* En esta fase se presentó el tema de investigación: Predicción de enfermedades cardiovasculares por medio de machine learning para aplicaciones móviles. En la Tabla I se detallan los componentes:

TABLA I
COMPONENTES PICO

P	Problema	Pacientes con enfermedades cardiovasculares o con factores de riesgo
I	Intervención	Modelos de aprendizaje automático
C	Comparación	Factores de riesgo y diagnósticos establecidos mediante criterios clínicos convencionales
O	Resultados	Precisión o efectividad de las predicciones clínicas generadas por los modelos

2) *Formulación de preguntas:* Inmediatamente se procedió a la siguiente fase, la formulación de la pregunta general PICO. ¿Qué modelos de aprendizaje automático aplicados a tecnologías clínicas se utilizaron para predecir factores de riesgo y diagnósticos establecidos mediante criterios clínicos convencionales en pacientes con enfermedades cardiovasculares o con factores de riesgo? A partir de la interrogante, se formularon preguntas específicas basadas en cada componente PICO que se muestra en la Tabla II, con el fin de identificar y seleccionar detalladamente los artículos que se eligieron más adelante.

TABLA II
PREGUNTAS PICO

P	Problema	Pacientes con enfermedades cardiovasculares o con factores de riesgo	¿Cuáles son los problemas en la detección temprana de las enfermedades cardiovasculares?
I	Intervención	Modelos de aprendizaje automático	¿Qué métodos se han implementado para mejorar la detección temprana de las enfermedades cardiovasculares?
C	Comparación	Factores de riesgo y diagnósticos establecidos mediante criterios clínicos convencionales	¿Cuáles son los resultados que se obtienen al entrenar los modelos predictivos?
O	Resultados	Precisión o efectividad de las predicciones clínicas generadas por los modelos	¿Qué nivel de precisión o efectividad presentan las predicciones generadas por estos modelos sobre enfermedades cardiovasculares?

3) *Identificación de palabras clave:* Seguidamente se identificaron las palabras clave para cada componente

generadas por la metodología PICO, a continuación, en la Tabla III.

TABLA III
PALABRAS CLAVE

P	Problema	Pacientes con enfermedades cardiovasculares o con factores de riesgo	Arterial hypertension, arterial hypertension problems, cardiovascular disease, high blood pressure, high blood pressure problems
I	Intervención	Modelos de aprendizaje automático	Machine learning, virtual consulting, KNN, k-nearest neighbors algorithm
C	Comparación	Factores de riesgo y diagnósticos establecidos mediante criterios clínicos convencionales	Photoplethysmography, measurement, medical history
O	Resultados	Precisión o efectividad de las predicciones clínicas generadas por los modelos	prediagnostic report, report, prediagnosis, decision making

4) *Implementación de operadores booleanos:* Posteriormente de identificar las palabras clave, se utilizaron los operadores booleanos con el conector OR para llevar a cabo la búsqueda. En la Tabla IV se muestran las palabras simples, sin comillas. Mientras que, las palabras compuestas se muestran entre comillas dobles (“”).

TABLA IV
OPERADOR OR Y PALABRAS COMPUESTAS

P	Problema	Pacientes con enfermedades cardiovasculares o con factores de riesgo	"Arterial hypertension" OR "arterial hypertension problems" OR "cardiovascular disease" OR "High blood pressure" OR "high blood pressure problems"
I	Intervención	Modelos de aprendizaje automático	"Machine learning" OR "virtual consulting" OR "KNN" OR "knearest neighbors algorithm"
C	Comparación	Factores de riesgo y diagnósticos establecidos mediante criterios clínicos convencionales	Photoplethysmography OR measurement OR "medical history"
O	Resultados	Precisión o efectividad de las predicciones clínicas generadas por los modelos	"prediagnostic report" OR report OR prediagnosis OR "decision making"

5) *Sintaxis de la fórmula PICO:* Para finalizar la formulación de las fases, se relacionan todos los componentes establecidos utilizando los operadores booleanos OR entre las palabras clave similares y el operador booleano AND entre los componentes en cada base de datos, ver Tabla V.

TABLA V
ECUACIÓN DE BÚSQUEDA DE BASES DE DATOS

Scopus ("Arterial hypertension" OR "arterial hypertension problems" OR "cardiovascular disease" OR "High blood pressure" OR "high blood pressure problems") AND ("Machine learning" OR "decision tree algorithms" OR "decision making" OR "virtual consulting") AND ("photoplethysmography OR measurement OR "medical history") AND ("prediagnostic report" OR report OR prediagnosis OR "decision making")
IEEE ("All Metadata": machine learning) AND ("All Metadata": "Arterial hypertension" OR "All Metadata": "arterial hypertension problems" OR "All Metadata": "cardiovascular disease" OR "All Metadata": "High blood pressure" OR "All Metadata": "high blood pressure problems") AND ("All Metadata": "Photoplethysmography OR "All Metadata": "measurement OR "All Metadata": "medical history")

Al concluir con la metodología, se identificaron 61 resultados en las bases de datos Scopus. Estos artículos pasaron a otra fase de selección para delimitar las fuentes de la revisión. Asimismo, se realizó otra búsqueda en la base de datos IEEE siguiendo la metodología PICO, de esta, se recuperaron un total de 336 resultados y se adicionaron a los resultados de Scopus. En total se recuperaron 397 registros.

6) *Definición de criterios de inclusión y de exclusión:*

6.1) Criterios de inclusión

- I1: Registros de artículos y revisiones sistemáticas
- I2: Registros que abordaron temas de "métodos de ML"
- I3: Registros que abordaron temas de "factores de riesgo y diagnóstico de ECV"
- I4: Registros que contengan las métricas de sensibilidad, especificidad, precisión y área bajo la curva

6.2) *Criterios de exclusión*

- E1: Registros consolidados sin duplicidad
- E2: Registros completos que no cuentan con acceso abierto
- E3: Registros anteriores a los últimos 5 años
- E4: Registros con menos de una cita o ninguna

B. *Prisma*

La herramienta PRISMA emplea una metodología que permitió la selección y el análisis de los registros, se estableció procesos de extracción para la inclusión final de artículos en la revisión sistemática. En la Fig. 1 se muestra la secuencia.

1) *Proceso de identificación:* En la elaboración del proceso de identificación, se establecieron cinco etapas: identificación, duplicidad, elegibilidad, selección y sesgo. En la primera etapa, se utilizaron los registros recopilados de las bases de datos de literatura científica Scopus e IEEE, que ofrecen una amplia variedad de publicaciones confiables y prestigiosas. Luego, los registros se almacenaron en la aplicación y gestor de documentos de investigación Mendeley. Sin embargo, no todos los artículos encontrados se cargaron

completamente, por lo que, se utilizaron hojas de cálculos en Excel para cargar los registros completos e identificar y analizar qué y cuáles registros se incluyeron en la revisión sistemática.

2) *Proceso de eliminación de duplicados:* Esta etapa consistió en eliminar los artículos repetidos en las bases de datos Scopus e IEEE, lo que resultó en la eliminación de tres artículos. Además, se suprimieron artículos que no presentaban el Identificador de Objetos Digitales (DOI), eliminándose dos artículos en Scopus y uno en IEEE, permitiendo una elección correcta de los registros para el proceso.

3) *Proceso de elegibilidad:* Esta etapa tuvo dos partes. Por un lado, la eliminación de los registros completos que no contaron con acceso abierto, se recuperó 64 registros. Por otro lado, se delimitó un período de los cinco últimos años y se recuperó 51 registros.

4) *Proceso selección:* Esta etapa permitió segmentar los enfoques de artículos y revisiones sistemáticas, resultando 47 registros. Por otra parte, se seleccionaron artículos con una o más citas, recuperándose 42 registros.

5) *Proceso de sesgo:* Finalmente, se excluyeron los artículos que no abordaban los temas de "métodos de ML", "factores de riesgo y diagnóstico de ECV" y "métricas de sensibilidad, especificidad, precisión y área bajo la curva", obteniéndose 26 artículos para la revisión sistemática, tal como se muestra en el siguiente diagrama.

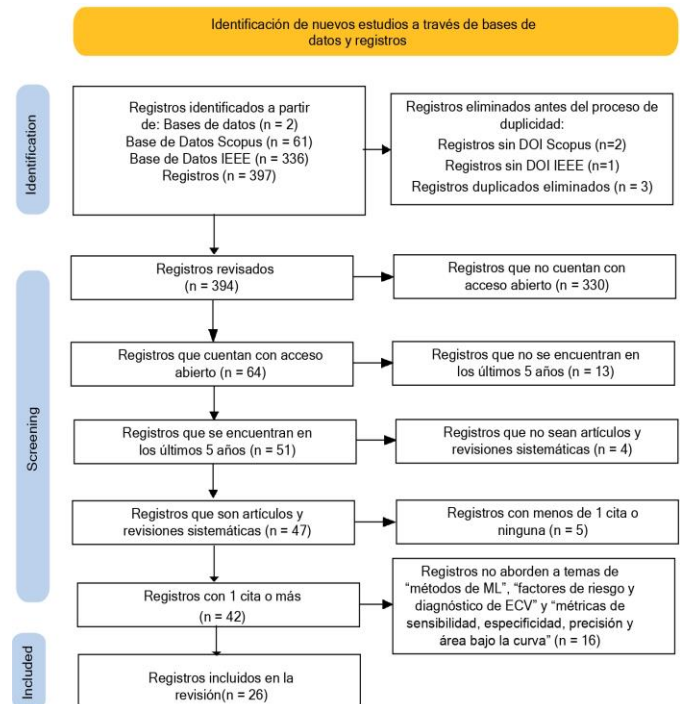


Fig. 1 Diagrama PRISMA

III. RESULTADOS

En la sección de resultados, se analizó y recopiló información de un total de 26 documentos seleccionados, entre artículos y revisiones sistemáticas. Se diseñaron tablas y gráficos para las representaciones visuales por categorías. Para la representación gráfica, se utilizaron las herramientas en hojas de cálculo en Excel y las librerías de Python, como pandas, matplotlib.pyplot y numpy, en las Fig. 2, 3 y 4. Además, los resultados responden a las preguntas PICO planteadas en el proceso de redacción de este trabajo y que se presentan a continuación:

A. RQ 1: ¿Cuáles son los problemas en la detección temprana de las enfermedades cardiovasculares?

En esta sección, se identificaron los problemas relacionados con la detección temprana de factores de riesgo y diagnósticos de enfermedades cardiovasculares.

Factores de riesgo y diagnósticos de ECV

Los factores de riesgo y diagnósticos de ECV se muestran en la Tabla VI. Los autores establecieron estos valores para verificar y desarrollar métodos tecnológicos de ML en sus estudios [1 – 26].

TABLA VI
MODELOS DEFINIDOS POR EL AUTOR SOBRE MÉTODOS DE ML

Factores de riesgo y diagnósticos de ECV	Tipos de factores de riesgo y diagnósticos
Diagnóstico	Onda de pulso radial
	ECG
	Función cardíaca
	Insuficiencia cardíaca
	Edad, Sexo, Peso
	Consumo de alcohol
	Hábitos de fumar
	Pulso arterial
Factores de riesgo	Insuficiencia cardíaca
	Presión arterial
	Enfermedad coronaria isquémica
	Infartos de miocardio, ECV
	Fibrilación auricular
	Enfermedad renal crónica
	Enfermedad oclusiva de la arteria periférica
	Enfermedad cardíaca coronaria
	Hipertensión
	Comorbilidades Charlson y Elixhauser
	Enfermedad pulmonar obstructiva crónica
	Diabetes, Depresión, Colesterol, Dieta, Estrés, Masa Corporal, Obesidad

B. RQ 2: ¿Qué métodos se han implementado para mejorar la detección temprana de las enfermedades cardiovasculares?

En esta sección, se identificaron los distintos métodos utilizados para la detección temprana de enfermedades cardiovasculares.

1) *Modelos definidos por los autores sobre métodos de ML:* Estos métodos definidos se muestran en la Tabla VII. Tanto para el aprendizaje profundo (DL) [1-8] como para ML [4, 9-19], los autores emplearon estos modelos para el análisis y recopilación de información.

TABLA VII
MODELOS DEFINIDOS POR EL AUTOR SOBRE MÉTODOS DE ML

Modelos de Aprendizaje Automático	Métodos de Aprendizaje	Nombre definido por el autor	Referencias
DL	CNN	FCQ-CNN	1
		CNN-DDM	2
		CNN ID	3
		CNN	4,6
		EDCNN	5
		HMM-CNN (ecgAI)	7
		Prueba Interna y prueba Externa	8
	RNN	RNN	4
ML	DT	DT	16
	LR	LR	13
	SVM	Clasificación Binaria	19
	RF	Clasificación Multietiqueta y jerárquica	18
		Clasificación Binaria	19
		Metamodelos ponderados - Modelo de votación híbrido	10
	XGBoost	Clasificación Multietiqueta y jerárquica	18
		XGBoost	4,17
	AutoPrognosis	AutoPrognosis	12
	CART	CART	14
	CatBoost	Modelos de ML	9
	HyperSMURF	HyperSMURF	16
	ET	Metamodelos ponderados - Modelo de votación híbrido	10
	NB	Modelo de dos etapas - NB - SVM	11
	Otros (Dura)	Metamodelos ponderados - Modelo de votación híbrido	10

Otros (Blanda)	Metamodelos ponderados - Modelo de votación híbrido	10
Otros (auto-ML)	Esquema autoML	14
Otros (ML)	Modelo de predicción optimizado	15

2) *Valores predictivos por factores de riesgo, diagnóstico y métodos utilizados por ML:* Estos se muestran en la Fig. 2. Este análisis permite a futuros investigadores profundizar en el tema de investigación basado en las palabras recopiladas. Las palabras más destacadas y relevantes en los artículos como valores predictivos son insuficiencia cardiaca [1, 2, 5, 9, 11, 18, 19, 22], diagnóstico [1, 2, 4-6, 8, 10, 18] y ML [1-5, 9, 10, 14-21, 23, 25, 26].

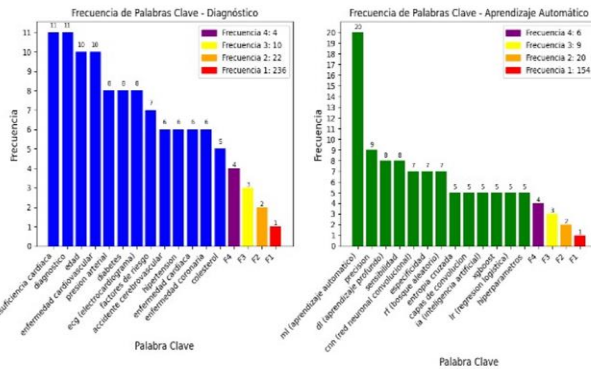


Fig. 2 Valores predictivos por factores de riesgo, diagnóstico y métodos utilizados por ML

- F1: Referencia a las palabras clave mencionadas al menos una vez
- F2: Referencia a las palabras clave mencionadas al menos dos veces
- F3: Referencia a las palabras clave mencionadas al menos tres veces
- F4: Referencia a las palabras clave mencionadas al menos cuatro veces

C. RQ 3: *¿Cuáles son los resultados que se obtienen al entrenar los modelos predictivos?*

En esta sección, se identificaron los resultados de las mediciones de factores de riesgos y diagnósticos de las ECV.

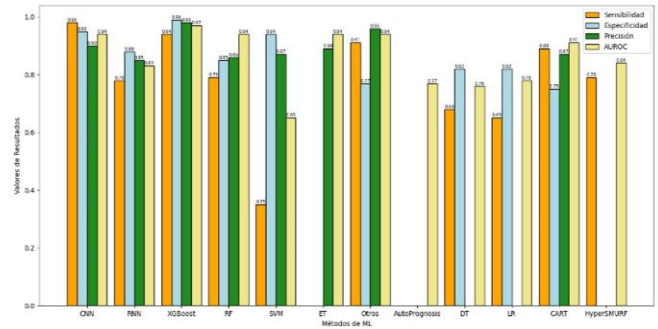


Fig 3. Mediciones de resultados por métodos de ML en conjunto de datos, registros hospitalarios y repositorios clínicos

Mediciones de resultados por métodos de ML en personas y dispositivos portátiles: Estas mediciones de resultados se grafican en la Fig. 4. Los autores utilizan muestras de mediciones en personas y dispositivos portátiles para entrenar métodos de ML y DL. Se destacan varios valores de medición: sensibilidad con un valor máximo de 1 utilizando el método RF, especificidad con un valor máximo de 0.91 para los métodos de RF, XGBoost y LR [18], precisión más efectiva para el método de CatBoost [9], y AUROC con valor máximo de 0.98 para los métodos de RF y XGBoost [18].

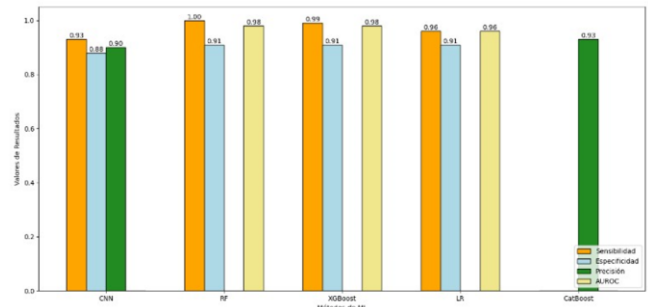


Fig 4 Mediciones de resultados por métodos de ML en personas y dispositivos portátiles

D. RQ 4: *¿Qué nivel de precisión o efectividad presentan las predicciones generadas por estos modelos sobre enfermedades cardiovasculares?*

En esta sección, se identificaron los valores de las predicciones finales de las mediciones de los métodos empleados de ML para la predicción de ECV.

1) *Valores de mediciones de resultados finales por métodos de ML:* Estos valores se muestran en la Tabla VIII. Las mediciones finales, basados en los dos tipos de muestras evaluados, resaltan los valores de los métodos más altos de los métodos utilizados por los autores.

TABLA VIII

VALORES DE MEDICIONES DE RESULTADOS FINALES POR MÉTODOS DE ML

Métodos de ML	Conjunto de datos, registros hospitalarios y repositorios clínicos				Personas y dispositivos portátiles			
	SE	SP	Precisión	AUR OC	SE	SP	Precisión	AUR OC
CNN	0.98	0.95	0.90	0.94	0.93	0.88	0.90	0.00
XGBoost	0.94	0.99	0.98	0.97	0.99	0.91	0.00	0.98
RF	0.79	0.85	0.86	0.94	1.00	0.91	0.00	0.98
LR	0.65	0.82	0.00	0.78	0.96	0.91	0.00	0.96
CatBoost	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.00

2) *Tendencia anual de publicaciones:* La tendencia anual de publicaciones se muestra en la Fig. 5. El 2021 fue el año que tuvo mayor cantidad de publicaciones y los últimos años disminuyendo. A continuación, el detalle: En el año 2019, se publicaron 5 artículos [2, 3, 7, 12, 24]; en el año 2020, 6 artículos [4, 5, 13, 18, 21, 22]; en el año del 2021, 8 artículos [6, 9, 14-16, 19, 23, 25]; en el año 2022, 4 artículos [1, 8, 17, 26] y en el año 2023, 3 artículos [10, 11, 20].

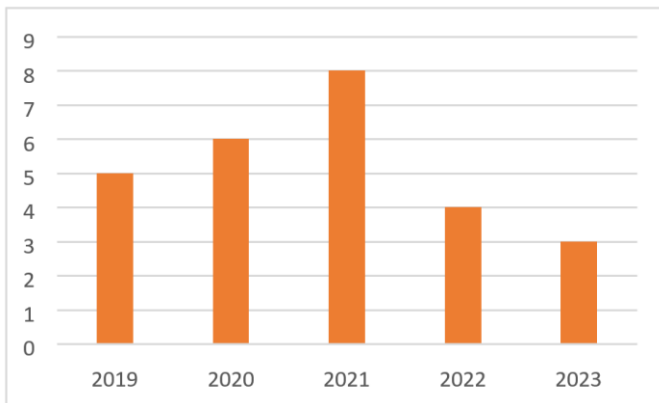


Fig. 5 Tendencia anual de publicaciones

3) *Estudios más citados por base de datos:* en la Tabla IX se muestran los estudios más citados por base de datos. Además, se identificó dos artículos con mayor número de citas. En Scopus, los artículos más citados tuvieron de 115 y 293 citas [7, 12], y en IEEE, 64 y 116 citas [3, 5].

TABLA IX
ESTUDIOS MÁS CITADOS POR BASE DE DATOS

Base de Datos	Referencia	Cantidad de Citaciones
Scopus	[7]	115
	[12]	293
IEEE	[3]	64
	[5]	116

4) *Origen de artículos seleccionados:* El origen de los artículos seleccionados se muestra en la Fig. 6. Los artículos se estructuraron por continentes y países según el origen de su publicación. Se revisaron 4 continentes: América [7, 15, 26], Asia [2-6, 8-11, 14, 17, 20, 23], Europa [1, 12, 13, 16, 18, 19, 21, 22, 24] y Oceanía [25]. Se especifican los países en sus respectivas ubicaciones.



Fig 6. Origen de artículos seleccionados

5) *Tipos de investigación por tema y estructura:*

Los tipos de investigación por tema se muestran en la Tabla X. Los autores se refieren a los temas relacionados con 2 tipos de investigación en general.

TABLA X

TIPOS DE INVESTIGACIÓN POR TEMA Y ESTRUCTURA DE INVESTIGACIÓN

Tipo de Investigación	Tema y estructura de Investigación	Cantidad de Artículos
Cuantitativo	“Uso de métodos tecnológicos de ML en los factores de riesgos y diagnósticos de enfermedades cardiovasculares”	21
Cualitativo	“Comparación entre métodos tradicionales y tecnológicos de medición de enfermedades cardiovasculares”	1
	“Estructuración y análisis de temas de investigación en RSL”	4

6). *Publicaciones en Revistas Científicas:* Estas publicaciones se muestran en la Tabla XI. Se presentan las referencias utilizadas en esta revisión y las revistas científicas en las que se publicaron.

TABLA XI
PUBLICACIONES EN REVISTAS CIENTÍFICAS

REVISTA	REFERENCIA
IEEE Access	1,2,3,4,5,6,9,10,11,19,23
Circulation: Cardiovascular Quality and Outcomes	7
IEEE Sensors Journal	18
Australian & New Zealand Journal of Psychiatry	21

International Journal of Environmental Research and Public Health	13,14
PLOS ONE	12,16
Diagnostics	25
Scientific Reports	15
Vascular Health and Risk Management	26
Frontiers in Cardiovascular Medicine	8,17
F1000Research	24
Journal of Vascular Surgery	22
Microsystems & Nanoengineering	20

IV. DISCUSIÓN

En este estudio, la metodología PICO permitió realizar una selección rigurosa de estudios basados en factores de riesgo, tipos de diagnóstico y el rendimiento de los modelos de ML y DL aplicados a la detección temprana de ECV. Los resultados obtenidos evidencian avances significativos en la aplicación de estas técnicas, así como desafíos persistentes y oportunidades de mejora. En el primer requerimiento, los problemas identificados para la detección temprana se clasificaron en dos categorías: factores de riesgo y diagnósticos, como se detalla en la Tabla VI. En esta tabla, los autores destacan variables claves utilizadas para entrenar los modelos de ML [1-26]. En el segundo requerimiento, los autores designaron nombres a los métodos de ML empleados en sus investigaciones, observando una mayor recurrencia a métodos de ML, como se muestra en la Tabla VII. Asimismo, se identificaron las palabras clave de los artículos en base a los valores predictivos mostrados, abarcando un total de 769 palabras (considerando cada palabra mencionada al menos una vez) representadas en la Fig. 2, incluyendo la frecuencia y conteo total de palabras, lo cual evidencia una tendencia metodológica orientada a la optimización de métricas. Respecto al tercer requerimiento, se identificaron dos tipos de muestras para las mediciones de los resultados. Por una parte, las muestras de conjuntos de datos, registros hospitalarios y repositorios clínicos, se utilizaron las métricas de SE, SP, precisión y AUROC. El modelo CNN alcanzó una sensibilidad (SE) de 0.98, superior a la de XGBoost, que obtuvo un valor de 0.94, logrando el mejor desempeño en este aspecto [14], como se visualiza en la Fig. 3. No obstante, los valores de especificidad (SP), precisión y área bajo la curva (AUROC) más altos fueron obtenidos con el método de XGBoost, con valores de 0.99, 0.98 y 0.97 respectivamente [4], en comparación con otros métodos de CNN y ML. Por otra parte, las muestras obtenidas de personas y dispositivos portátiles, el método RF alcanzó una SE de 1 con el método de RF, en comparación al método de XGBoost con un valor de 0.99 [18]. La especificidad tuvo un valor máximo de 0.91 utilizando los métodos de RF, XGBoost y LR, como se visualiza en la Fig. 4. Estos resultados sugieren que los modelos pueden comportarse de manera distinta según el tipo

de muestra y fuente de datos, lo cual es relevante para futuras aplicaciones en entornos clínicos reales. En el cuarto requerimiento, se compararon los tipos de muestra en base a las mediciones finales obtenidas del análisis de los artículos. El método CNN tiene un valor final de SE de 0.98, SP de 0.95 [5] y AUROC de 0.94 [7] para la muestra de conjunto de datos, y precisión de 0.90 [2, 4] para ambas muestras. El método XGBoost muestra con un valor final de SE de 0.99 y AUROC de 0.98 [18] para la muestra de personas y dispositivos portátiles, y con una SP de 0.99 y precisión de 0.98 [4] para la muestra de conjunto de datos. El método RF tiene valores de SE, SP y AUROC de 1, 0.91 y 0.98 respectivamente [18] para la muestra de personas y dispositivos portátiles, y una precisión de 0.86 [10] para la muestra de conjunto de datos. El método de LR con valores de SE, SP y AUROC de 0.96, 0.91 y 0.96 respectivamente [18] para la muestra de personas y dispositivos portátiles. El método CatBoost, con un valor de precisión de 0.93, mostró buenos resultados en el diagnóstico, particularmente en la muestra de personas y dispositivos portátiles, aunque no superó a otros métodos como XGBoost en cuanto a AUROC y sensibilidad [9]. Estos hallazgos indican que la elección del modelo debe considerar tanto el tipo de datos como el entorno de aplicación, siendo los modelos híbridos y adaptativos una posible vía para mejorar la precisión y la generalización. Los artículos revisados por la tendencia anual se publicaron con mayor frecuencia en 2021, con de 8 publicaciones, en contraste con los años 2019, 2020, 2022 y 2023 con 5, 6, 4 y 3 publicaciones correlativamente, como se muestra en la Fig. 5. De igual modo, los artículos con más citaciones se clasificaron por las bases de datos seleccionadas, considerando los 2 artículos con mayor cantidad de citaciones. En la base de datos en Scopus, se recuperaron 115 y 293 citaciones. En la base de datos de IEEE, se recuperaron 64 y 116 citaciones, utilizando los valores actualizados en las bases de datos correspondientes, como se visualiza en la Tabla IX. Los artículos evaluados por su origen mostraron mayor frecuencia en el continente de Asia con una cantidad de 13 países ubicados, de forma distinta a los continentes de Oceanía, América y Europa con una cantidad 1, 3 y 9 países ubicados de publicación respectivamente, como se muestra en la Fig. 6. Los autores se refieren a temas de investigación relacionados con dos tipos de investigaciones en general. Para el tipo de investigación cuantitativo, se identificaron 21 artículos que exploran el tema “Uso de métodos tecnológicos de ML en los factores de riesgos y diagnósticos de enfermedades cardiovasculares” [1-21, 23]. Para las investigaciones cualitativas, se contabilizó un artículo que aborda el tema de la “Comparación entre métodos tradicionales y tecnológicos de medición de enfermedades cardiovasculares” [26], y 4 artículos que tratan sobre la “Estructuración y análisis de temas de investigación en RSL” [22, 24, 25]. Del mismo modo, las publicaciones en revistas científicas otorgan un valor a los documentos de investigación por su almacenamiento en las bases de datos confiable, se clasificaron por cada documento, como se muestra en la Tabla

XI, lo que permitió verificar su inclusión en fuentes confiables y su contribución al avance del conocimiento. Durante el análisis de los estudios revisados, se identificaron diversos desafíos y limitaciones comunes en el desarrollo e implementación de modelos de aprendizaje automático aplicados a enfermedades cardiovasculares. Uno de los principales retos es el desequilibrio de clases en los conjuntos de datos clínicos, especialmente cuando los eventos de interés, como los infartos, son poco frecuentes. Además, los datos longitudinales, que consisten en registros repetidos del mismo paciente a lo largo del tiempo, pueden agregar complejidad tanto estadística como computacional a los modelos. En muchos casos, los estudios optan por simplificaciones metodológicas que pueden afectar la precisión y utilidad clínica de los resultados. Asimismo, existe una tendencia predominante a evaluar los modelos exclusivamente con métricas de rendimiento diagnóstico como la AUROC o la SE, sin considerar otras variables relevantes para la práctica clínica, como el impacto en los flujos de trabajo, la reducción de reingresos o los desenlaces clínicos a largo plazo. En términos de implementación, se observó una brecha significativa entre los modelos desarrollados en contextos retrospectivos y su validación prospectiva en entornos reales. La escasa evaluación clínica limita su adopción en la práctica médica, a pesar de sus capacidades técnicas. Además, la diversidad en las metodologías empleadas dificulta la comparación directa entre estudios, impidiendo la identificación de modelos consistentes o mejores prácticas estandarizadas. Pese a estos desafíos, los estudios analizados ofrecen aportes relevantes para la evolución del diagnóstico asistido por inteligencia artificial. Se destaca la capacidad de los modelos para identificar factores de riesgo clínico relevantes, como la hipertensión, la diabetes o el índice de masa corporal, especialmente cuando se integran variables temporales que mejoran la precisión de las predicciones. También se reconoce el valor del enfoque preventivo que ofrecen estos sistemas, permitiendo la detección temprana de condiciones comórbidas y la optimización de recursos en entornos hospitalarios. Se recomienda que los estudios futuros no solo evalúen el rendimiento diagnóstico, sino también otros factores clínicos cruciales, como la aceptabilidad de los modelos por parte de los profesionales de salud, el impacto en los flujos de trabajo y la mejora de los resultados a largo plazo para los pacientes. En conjunto, estos hallazgos destacan el valor de integrar enfoques de ML y DL en la práctica clínica, siempre que se aborden cuidadosamente los retos metodológicos, éticos y de implementación.

V. CONCLUSIÓN

La presente revisión sistemática evidenció el avance significativo del uso de algoritmos de ML y DL en la detección temprana de ECV, destacando su capacidad para optimizar el diagnóstico a partir de diferentes tipos de muestras y métricas de rendimiento. Entre los modelos

analizados, CNN y XGBoost demostraron desempeños destacados en sensibilidad, especificidad y AUROC, aunque con variaciones según la fuente de datos utilizada. Se identificaron desafíos recurrentes, como el desequilibrio de clases, la falta de validación prospectiva y la limitada incorporación de variables clínicas integrales. A pesar de ello, los estudios revisados aportan evidencia sólida sobre el potencial de estos enfoques para mejorar la identificación temprana de factores de riesgo y optimizar procesos clínicos. En conclusión, la efectividad de los modelos está influenciada por la calidad de los datos, el tipo de muestra y el contexto de aplicación, y se recomienda priorizar investigaciones que integren métricas clínicas, evaluación en entornos reales y adaptabilidad de los modelos, promoviendo así su aplicabilidad práctica y su adopción en la atención médica.

REFERENCIAS

- [1] U. Ullah, A. G. O. Jurado, I. D. Gonzalez, and B. Garcia-Zapirain, "A Fully Connected Quantum Convolutional Neural Network for Classifying Ischemic Cardiopathy," *IEEE Access*, vol. 10, pp. 134592–134605, May 2022, doi: 10.1109/ACCESS.2022.3232307.
- [2] L. Zhao, C. Liu, S. Wei, C. Liu, and J. Li, "Enhancing Detection Accuracy for Clinical Heart Failure Utilizing Pulse Transit Time Variability and Machine Learning," *IEEE Access*, vol. 7, pp. 17716–17724, May 2019, doi: 10.1109/ACCESS.2019.2895230.
- [3] S. Baek, J. Jang, and S. Yoon, "End-to-end blood pressure prediction via fully convolutional networks," *IEEE Access*, vol. 7, pp. 185458–185468, 2019, doi: 10.1109/ACCESS.2019.2960844.
- [4] L. Ibrahim, M. Mesinovic, K.-W. Yang, and M. A. Eid, "Explainable Prediction of Acute Myocardial Infarction Using Machine Learning and Shapley Values," *IEEE Access*, vol. 8, pp. 210410–210417, May 2020, doi: 10.1109/ACCESS.2020.3040166.
- [5] Y. Pan, M. Fu, B. Cheng, X. Tao, and J. Guo, "Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction on the Internet of Medical Things Platform," *IEEE Access*, vol. 8, pp. 189503–189512, May 2020, doi: 10.1109/ACCESS.2020.3026214.
- [6] S. Kim, J.-O. Hahn, and B. D. Youn, "Deep Learning-Based Diagnosis of Peripheral Artery Disease via Continuous Property-Adversarial Regularization: Preliminary in Silico Study," *IEEE Access*, vol. 9, pp. 127433–127443, May 2021, doi: 10.1109/ACCESS.2021.3112678.
- [7] G. H. Tison, J. Zhang, F. N. Dellling, and R. C. Deo, "Automated and Interpretable Patient ECG Profiles for Disease Detection, Tracking, and Discovery," *Circ Cardiovasc Qual Outcomes*, vol. 12, no. 9, 2019, doi: 10.1161/CIRCOUTCOMES.118.005289.
- [8] X. Lin et al., "Echocardiography-based AI detection of regional wall motion abnormalities and quantification of cardiac function in myocardial infarction," *Front Cardiovasc Med*, vol. 9, 2022, doi: 10.3389/fcvm.2022.903660.
- [9] H. R. H. Al-Absi, M. A. Refaee, A. U. Rehman, M. T. Islam, S. B. Belhauari, and T. Alam, "Risk Factors and Comorbidities Associated to Cardiovascular Disease in Qatar: A Machine Learning Based Case-Control Study," *IEEE Access*, vol. 9, pp. 29929–29941, May 2021, doi: 10.1109/ACCESS.2021.3059469.
- [10] S. C. Patra, B. U. Maheswari, and P. B. Pati, "Forecasting Coronary Heart Disease Risk With a 2-Step Hybrid Ensemble Learning Method and Forward Feature Selection Algorithm," *IEEE Access*, vol. 11, pp. 136758–136769, Jul. 2023, doi: 10.1109/ACCESS.2023.3338369.
- [11] C. H. Lee and S. H. Kim, "ECG Measurement System for Vehicle Implementation and Heart Disease Classification Using Machine Learning," *IEEE Access*, vol. 11, pp. 17968–17982, 2023, doi: 10.1109/ACCESS.2023.3245565.
- [12] A. M. Alaa, T. Bolton, E. D. Angelantonio, J. H. F. Rudd, and M. van der Schaar, "Cardiovascular disease risk prediction using automated

- machine learning: A prospective study of 423,604 UK Biobank participants,” *PLoS One*, vol. 14, no. 5, 2019, doi: 10.1371/journal.pone.0213653.
- [13] D. Barbieri et al., “Predicting cardiovascular risk in athletes: Resampling improves classification performance,” *Int J Environ Res Public Health*, vol. 17, no. 21, pp. 1–9, 2020, doi: 10.3390/ijerph17217923. N. B. Sønderoft et al., “The intestinal microbiome is a co-determinant of the postprandial plasma glucose response,” *PLoS One*, vol. 15, no. 9, September, 2020, doi: 10.1371/journal.pone.0238648.
- [14] C.-C. Shih, S.-H. Chen, G.-D. Chen, C.-C. Chang, and Y.-L. Shih, “Development of a longitudinal diagnosis and prognosis in patients with chronic kidney disease: Intelligent clinical decision-making scheme,” *Int J Environ Res Public Health*, vol. 18, no. 23, 2021, doi: 10.3390/ijerph182312807.
- [15] P. Pournazari et al., “Cardiac involvement in hospitalized patients with COVID-19 and its incremental value in outcomes prediction,” *Sci Rep*, vol. 11, no. 1, 2021, doi: 10.1038/s41598-021-98773-4.
- [16] B. Aguilaniu et al., “A machine learning approach to predict extreme inactivity in COPD patients using non-activity-related clinical data,” *PLoS One*, vol. 16, no. 8, August, 2021, doi: 10.1371/journal.pone.0255977.
- [17] Q. Sun et al., “A prediction model for major adverse cardiovascular events in patients with heart failure based on high-throughput echocardiographic data,” *Front Cardiovasc Med*, vol. 9, 2022, doi: 10.3389/fcvm.2022.1022658.
- [18] S. Mehrang et al., “Classification of Atrial Fibrillation and Acute Decompensated Heart Failure Using Smartphone Mechanocardiography: A Multilabel Learning Approach,” *IEEE Sens J*, vol. 20, no. 14, pp. 7957–7968, Jul. 2020, doi: 10.1109/JSEN.2020.2981334.
- [19] D. Chicco, C. A. Lovejoy, and L. Oneto, “A Machine Learning Analysis of Health Records of Patients With Chronic Kidney Disease at Risk of Cardiovascular Disease,” *IEEE Access*, vol. 9, pp. 165132–165144, May 2021, doi: 10.1109/ACCESS.2021.3133700.
- [20] L. Wang, S. Tian, and R. Zhu, “A new method of continuous blood pressure monitoring using multichannel sensing signals on the wrist,” *Microsyst Nanoeng*, vol. 9, no. 1, 2023, doi: 10.1038/s41378-023-00590-4.
- [21] J. Souza, J. V. Santos, V. B. Canedo, A. Betanzos, D. Alves, and A. Freitas, “Importance of coding co-morbidities for APR-DRG assignment: Focus on cardiovascular and respiratory diseases,” *Health Information Management Journal*, vol. 49, no. 1, pp. 47–57, 2020, doi: 10.1177/1833358319840575.
- [22] J. Raffort et al., “Artificial intelligence in abdominal aortic aneurysm,” *J Vasc Surg*, vol. 72, no. 1, pp. 321-333.e1, 2020, doi: 10.1016/j.jvs.2019.12.026. E. Wierda, D. van Veghel, A. Hirsch, and B. A. J. M. de Mol, “Heart teams in the Netherlands: From teamwork to data driven decision-making,” *Netherlands Heart Journal*, vol. 28, pp. 73–77, Aug. 2020, doi: 10.1007/S12471-020-01452-8.
- [23] J. Yang et al., “Analysis of the Radial Pulse Wave and its Clinical Applications: A Survey,” *IEEE Access*, vol. 9, pp. 157940–157959, May 2021, doi: 10.1109/ACCESS.2021.3128916.
- [24] G. Medic et al., “Evidence-based clinical decision support systems for the prediction and detection of three disease states in critical care: A systematic literature review,” *F1000Res*, vol. 8, 2019, doi: 10.12688/f1000research.20498.1.
- [25] C. Boyd et al., “Machine learning quantitation of cardiovascular and cerebrovascular disease: A systematic review of clinical applications,” *Diagnostics*, vol. 11, no. 3, 2021, doi: 10.3390/diagnostics11030551.
- [26] I. U. Haq, K. Chhatwal, K. Sanaka, and B. Xu, “Artificial Intelligence in Cardiovascular Medicine: Current Insights and Future Prospects,” *Vasc Health Risk Manag*, vol. 18, pp. 517–528, 2022, doi: 10.2147/VHRM.S279337.