

Modelos de identificación de enfermedades cardiovasculares implementando técnicas de aprendizaje máquina: una revisión sistemática de la literatura

Johan Mardini - Bovea¹, Dixon Salcedo¹, Emiro De-la-Hoz-Franco¹, Yadira Quiñonez², Luz Jimenez-Roa³, Víctor De la Hoz-Avila¹, Jhon González Sanchez¹

[jmardini@cuc.edu.co](mailto:jardini@cuc.edu.co); dsalcedo2@cuc.edu.co; edelahoz@cuc.edu.co; yadiraqui@uas.edu.mx; lijimenezr@sena.edu.co; vdelahoz10@cuc.edu.co; jgonzalez120@cuc.edu.co

¹ Departamento de Ciencias de la Computación y Electrónica, Universidad de la Costa–CUC, Barranquilla 080002, Colombia.

² Facultad de informática de Mazatlán-FIMAZ, Universidad Autónoma de Sinaloa, Mazatlán 82000, México.

³ Centro de Enseñanza Colombo Alemán, SENA Barranquilla 08004, Colombia.

DOI: [10.17013/risti.53.87-105](https://doi.org/10.17013/risti.53.87-105)

Resumen: El uso de técnicas de Aprendizaje Automático (AA) en el área de la salud, específicamente en la identificación de enfermedades cardiovasculares (IEC), ha tenido un impacto significativo debido a la capacidad para analizar grandes cantidades de datos y extraer información relevante que puede ser esencial para la toma de decisiones médicas. Sin embargo, antes de ponerlos a disposición de los usuarios finales (médicos), se debe evaluar su capacidad para detectar sintomatologías relacionadas con enfermedades cardíacas utilizando puntos de referencia de conjuntos de datos en escenarios experimentales. Por lo tanto, es complicado determinar qué características utilizar en el proceso de evaluación y qué técnicas de ML son más adecuadas para la predicción de IEC. Este artículo presenta una revisión sistemática de la literatura sobre el procesamiento de conjuntos de datos basados en pruebas clínicas de enfermedades cardiovasculares y técnicas de AA. En este sentido, se realizó un análisis de las diferentes variables extraídas de publicaciones de revistas indexadas en bases de datos especializadas como Scopus, Web of Science, Science Direct, Biomed y Pubmed.

Palabras-clave: Enfermedades cardiovasculares, técnicas de inteligencia artificial, conjunto de datos, síndrome coronario agudo.

Models of identification cardiovascular diseases implementing machine learning techniques: a systematic literature review

Abstract: The use of Machine Learning (ML) techniques in the health area, specifically in the identification of cardiovascular diseases (IEC), has had a significant impact due to the ability to analyze large amounts of data and extract

relevant information that can be essential for medical decision-making. However, before making them available to end users (doctors), their ability to detect heart disease-related symptomatology should be evaluated using benchmark data sets in experimental settings. Therefore, determining which features to use in the evaluation process and which ML techniques are most suitable for IEC prediction is complicated. This article presents a systematic literature review on processing cardiovascular disease clinical trial-based datasets and ML techniques. In this sense, the different variables were analyzed from journal publications indexed in specialized databases such as Scopus, Web of Science, Science Direct, Biomed, and Pubmed.

Keywords: Cardiovascular diseases, artificial intelligence techniques, data set, acute coronary syndrome.

1. Introducción

La cardiopatía isquémica es responsable del 42% de las muertes por enfermedades cardiovasculares a nivel mundial, mientras que las enfermedades cerebrovasculares son responsables del 34%. Se estima que 17,7 millones de personas murieron como resultado de esta causa en 2015, lo que representa el 31% de todas las muertes registradas en todo el mundo (Organización Panamericana de la Salud, s. f.). Es decir, 7,4 millones de estas muertes fueron causadas por enfermedades cardíacas. Más de tres cuartas partes de las defunciones por ECV se producen en los países de ingresos bajos y medios. De los 17 millones de muertes de personas menores de 70 años causadas por enfermedades no transmisibles, el 82% se registra en países de ingresos bajos y medios, y el 37% se debe a enfermedades cardiovasculares (ECV), (Organización Mundial de la Salud, s. f.). De acuerdo con informes de la Organización Mundial de la Salud (OMS) (Organización Mundial de la Salud, s. f.) y la Organización Panamericana de la Salud (OPS) (Instituto Nacional de Salud, s. f.), las enfermedades cardiovasculares, junto con el cáncer, la diabetes y las enfermedades pulmonares crónicas, se identifican como enfermedades que están aumentando rápidamente y se clasifican como las principales causas de muerte en el mundo. Las enfermedades cardiovasculares son la causa más importante de muerte, superando en gran medida el cáncer, la diabetes mellitus y las enfermedades pulmonares crónicas.

En la actualidad, las instituciones de salud están demandando cada vez más la incorporación de la tecnología. Los sistemas de IEC identifican enfermedades cardíacas, lo que mejora la calidad de vida de los pacientes al permitir una identificación temprana de ECV. El área de investigación de ECV ha tenido un impacto significativo en la creación de soluciones pre-evento cardíaco, como el apoyo a los pacientes con SCA. La aplicación de múltiples soluciones en entornos clínicos monitoreados con datos de exámenes médicos demuestra esto (Yao, 2020). Basados en pruebas relacionadas con la enfermedad cardiovascular, los objetivos de los sistemas de IEC son los siguientes: 1) la creación de modelos predictivos que permitan identificar de forma bi-clase la tendencia a padecer enfermedad cardiovascular; 2) proporcionar al equipo médico las herramientas necesarias basadas en software para identificar alguna enfermedad cardiovascular; y 3) estratificar el riesgo de los pacientes mediante la combinación de puntuaciones de riesgo validadas.

Los datos recopilados de sensores en entornos controlados se almacenan en un conjunto de datos incluyendo pruebas de laboratorio, exámenes diagnóstico y pruebas de esfuerzo clínicas. Se han sugerido varias formas de recopilar datos, como biomarcadores (Idris, 2020), pruebas clínicas (Narayan, 2019), sensores biomédicos (Metsker, 2020; Mendes, 2018; González-Alcaide, 2020) y pruebas de riesgo cardiovascular (Xue, 2020; Rezaianzade, 2020; Brandberg, 2020; Konys, 2019). Luego, el conjunto de datos se utiliza para entrenar varias técnicas de inteligencia artificial, incluido el aprendizaje automático, con el objetivo de predecir las tendencias de los pacientes hacia una enfermedad específica de SCA. Esto daría una alerta al equipo médico adecuado y reduciría los riesgos relacionados con el deterioro de la salud de los pacientes analizados. Actualmente existen conjuntos de datos para ECV. En consecuencia, determinar qué características o evaluaciones médicas utilizar en el proceso de evaluación de un sistema de IEC y qué técnicas implementar en las fases de construcción del modelo (preprocesamiento, selección de características, clasificación y postclasificación) son las más adecuadas para mejorar los índices de identificación de algunas enfermedades relacionadas con SCA.

Los sistemas IEC se derivan de los avances tecnológicos en los sensores biomédicos y en los resultados de pruebas médicas en tiempos extremadamente cortos, o incluso instantáneos. Como se detalla en (Chu, 2020; Magoev, 2018; Metsker, 2017; Si, 2021), hay muchas aplicaciones de IEC, pero se destacan la detección de riesgo cardiovascular (Kutyrev, 2019), la detección del Síndrome Coronario Agudo con Elevación del Segmento ST (SCACEST) (Xie, 2020), la identificación de angina inestable (AI) (Sun, 2020) y el Síndrome Coronario Agudo sin Elevación del Segmento ST (SCASEST) (Iannatton, 2020), entre otros. Por lo tanto, es extremadamente difícil evaluar la confiabilidad de los IEC en términos de su capacidad para predecir una variedad de enfermedades cardíacas, especialmente las del SCA recopiladas en el conjunto de datos. Con el desarrollo de nuevas IEC, la recolección de datos se ha vuelto más difícil. Esto se debe a que la recolección de datos registrados en Historias Clínicas Electrónicas (HCE) o Registro Médico Electrónico (RME) no incluye una variedad de pruebas médicas y exámenes diagnósticos para evaluar los IEC con suficiente rigor. Una variedad de puntos de referencia del conjunto de datos ha surgido como resultado de esta situación. Estos incluyen el registro de Coronary Heart Attack Ireland Register (CHAIR), Global Registry of Acute Coronary Event (GRACE), Myocardial Infarction National Audit Project (MINAP), Spanish Register form Non-St Elevation MI (DESCARTES), Acute Coronary Syndrome Registry International (ACOSRI), entre otros.

Teniendo en cuenta lo anterior, este trabajo tiene como objetivo que los investigadores desarrollen IEC y los coloquen a prueba en el contexto de pruebas de laboratorio, para enviar sus propuestas a evaluación utilizando distintos puntos de referencia de conjunto de datos, y finalmente, validar sus desarrollos con una muestra con pacientes y así contrastar efectividad de los mismos. La aplicación de tecnologías emergentes como las técnicas de AA, que todavía están en proceso de investigación y evolución, es indudable que el sector de la salud es uno de los muchos contextos en los que se pueden utilizar, porque aún existe un largo camino por recorrer en este ámbito de aplicación, y el uso de estas herramientas presenta múltiples retos y desafíos. Los siguientes son algunos de los principales hallazgos que este documento ha hecho para el campo de estudio:

1. Encontrar modelos de enfermedades cardiovasculares basados en AA.
2. Encontrar modelos basados en AA para encontrar enfermedades de SCA.
3. Identificación y evaluación del riesgo cardiovascular, incluidas sus aplicaciones y avances.

Debido a que no existe un punto de referencia que haya identificado los conjuntos de datos más relevantes en términos de los sistemas de IEC, el enfoque propuesto en este documento es una contribución original. Este artículo proporciona una caracterización detallada y posterior análisis de las diferentes técnicas de clasificación utilizadas para enfermedades cardiovasculares, incluidas las técnicas de distribución del conjunto de datos utilizado para entrenamiento y pruebas, así como las métricas de calidad (precisión, exactitud) utilizadas en la evaluación de IEC. Un volumen considerable de artículos a revisar se compone de la condensación, documentación y análisis de las variables mencionadas para cada conjunto de datos y estudios realizados entre 2019 a 2024.

Todas estas propuestas se basan en herramientas basadas en las Tecnologías de la Información y las Comunicaciones (TIC), como el Internet de las cosas (IoT), los sistemas independientes (para la captación de datos y la gestión de los procesos de comunicación entre sensores y repositorios de datos), las plataformas de computación en la nube (para el almacenamiento y procesamiento en la nube de grandes volúmenes de datos capturados en tiempo real), y la construcción de modelos predictivos basados en inteligencia artificial. Específicamente, utilizando técnicas de AA (para encontrar relaciones no triviales entre varias variables cardiovasculares y predecir comportamientos futuros a partir de la información histórica de los pacientes).

Este artículo científico está organizado de la siguiente manera: la investigación relacionada se describe en la sección 2. La metodología para realizar una revisión exhaustiva de la literatura se presenta en la Sección 3. El análisis cuantitativo se encuentra en la Sección 4. Los análisis técnicos se encuentran en la sección 5, se incluye una caracterización de cada conjunto de datos identificado, así como análisis de resultados correspondientes. Finalmente, la conclusión y el trabajo futuro se presentan en la sección 6.

2. Investigación relacionada

Al revisar literatura relacionada, se ha observado que algunas revisiones de la literatura analizan enfoques tecnológicamente muy específicos (técnicas de AA e hibridación (Manterola, 2013; García-Pérez, 2013), estudios comparativos de técnicas de riesgo cardiovascular (Merlano-Porras, 2013), (Henriquez, 2022), sistema de monitoreo cardiovascular (Kitchenham, 2009)). Aunque no son completas, otras revisiones se limitan a evaluar la precisión de las técnicas de clasificación para predecir enfermedades cardiovasculares a partir de conjuntos de datos (Grams, 2011; Comas-Gonzalez, 2023; Sanchez, 2016). Sin embargo, no hay pruebas que sugieran una revisión exhaustiva de la literatura que permita caracterizar conjuntos de datos de ECV enfocados en el SCA. La siguiente es una descripción detallada de los estudios mencionados anteriormente:

Kitchenham et al. (Kitchenham, 2009), usan técnicas asistidas por computadora para proporcionar herramientas rápidas y precisas para identificar las señales de

electrocardiograma (ECG) de un paciente. También resumen y comparan métodos de diagnóstico utilizando preprocesamiento de datos, ingeniería de características, clasificación y aplicación. Además, se lleva a cabo un modelo extremo a extremo que integra la extracción y clasificación de características en algoritmos de aprendizaje en (Manterola, 2013). Esto hace que el proceso de análisis de datos sea mucho más sencillo y muestra una precisión evaluada de alta calidad. En (García-Pérez, 2013), los autores examinaron dispositivos portátiles que permiten a los usuarios monitorear su estado cardiovascular en cualquier momento. Estos dispositivos ofrecen nuevos escenarios y plantean desafíos para la aplicación de algoritmos para identificar enfermedades utilizando electrocardiogramas. De manera similar, en (Merlano-Porras, 2013) se implementó la extracción de características clínicas multivariadas de pacientes con SCA registrados en una base de datos, y se crearon varios modelos que predecían con éxito un paro cardíaco en pacientes con SCA utilizando diferentes algoritmos de AA.

En (Henriquez, 2022) proponen el desarrollo de un modelo predictivo basado en hibridación para la predicción de enfermedades cardiovasculares utilizando el conjunto de datos Cleveland Heart Disease mediante técnicas de selección de características estadísticas con aprendizaje automático y redes neuronales artificiales (SOM Y GHSOM). Los autores en (Comas-Gonzalez, 2023) realizaron un estudio comparativo de métodos de inteligencia artificial no supervisados para encontrar el modelo de predicción cardiovascular más efectivo. En (Sanchez, 2016), se comparó la predicción del aprendizaje automático con modelos de predicción de riesgos cardiovascular existentes para pacientes con SCA, como Registry of Acute Coronary Events (GRACE), National Early Warning Score (NEWS) y Modified Early Warning Score (MEWS).

3. Metodología

La Revisión Sistemática de la Literatura (RSL) es una pieza clave de la investigación que permite establecer las bases para investigaciones futuras. Kitchenham et al (Kitchenham, 2009) proponen una metodología basada en la definición de preguntas de investigación, proceso de búsqueda, criterios de inclusión y exclusión, evaluación de la calidad, recolección de datos, análisis de datos y desviaciones del protocolo. La gran mayoría de las investigaciones de este tipo se realizan en el campo de la salud, para lo cual se analizaron las metodologías propuestas en (Manterola, 2013; García-Pérez, 2013; Comas-Gonzalez, 2023). En el campo de la ingeniería, los autores (Merlano-Porras, 2013; Henriquez, 2022; Comas-Gonzalez, 2023; Sanchez, 2016; Grams, 2011; Crimi, 2020) proponen sus respectivas metodologías. Para el desarrollo de esta RSL se basará en la implementación de (Knowles, 2018), lo cual se organiza en tres etapas: la definición de parámetros de búsqueda (objetivo, hipótesis e índice de búsqueda), identificación y depuración en bases de datos (selección de cadenas de búsqueda cuyos resultados se profundizarán) y la propuesta de respuestas a la hipótesis (de la información obtenida, categorización y análisis de los artículos más relevantes).

En particular, en (Kandasamy, 2018) se propone un método para validar la Enfermedad Renal Crónica (ERC) y otras afecciones relacionadas en conjuntos de datos existentes, incluidos los registros de enfermedades y los conjuntos de datos administrativos. Cada uno de estos estudios contribuyó al método RSL utilizado en la investigación, que se organizó en tres etapas. La primera etapa comienza con la determinación de los objetivos

de la revisión y consiste en identificar la siguiente hipótesis: “¿Qué conjunto de datos de SCA y técnicas de AA ha tenido un mayor impacto en el desarrollo de investigaciones asociadas a IEC?”. Posteriormente, se procedió a localizar los temas de las bases de datos en las que se centraría la búsqueda. Estas bases de datos incluyen Scopus, Web of Science, Science Direct, Biomed y Pubmed. Finalmente, se identificaron y validaron las palabras clave a utilizar y descartar debido al ruido que produce en los resultados.

El esquema utilizado para crear los términos de búsqueda se muestra en la Figura 1 (a). Se ha eliminado el término “imagen” porque el procesamiento de datos de este tipo de investigación es muy diferente al de otros mecanismos de retroalimentación, como biomarcadores, pruebas invasivas y no invasivas, así como exámenes diagnósticos cardiovasculares. Adicionalmente, se estableció una ventana de búsqueda de 2019 a 2024 para determinar el alcance de las publicaciones a analizar.

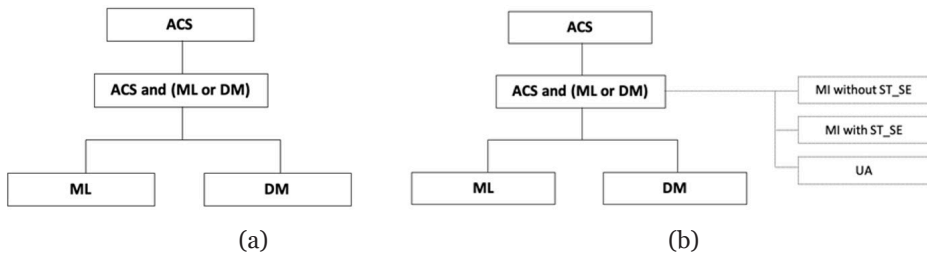


Figura 1 – (a) Cadena general de búsqueda (b) Cadena de específica de búsqueda

En la segunda etapa, se identificaron y filtraron datos de bases de datos especializadas y se analizaron los resultados de varias cadenas de búsqueda. Los datos se han representado sintéticamente en una variedad de arreglos y los índices de búsqueda se han organizado jerárquicamente, eliminando las combinaciones que no dieron ningún resultado. Además, se descubrieron los términos que produjeron una gran cantidad de resultados, y se agregaron criterios más precisos para limitar las búsquedas al tema de la investigación. Se seleccionaron las cadenas de búsqueda que produjeron resultados al identificar artículos que coinciden con la hipótesis sugerida. La Figura 1(b) muestra las palabras clave utilizadas para crear las cadenas de búsqueda, cuya estructura es la siguiente: (“Acute Coronary Syndrome”) OR (“Myocardial Infarction without ST-segment Elevation”) OR “Myocardial Infarction with ST-segment Elevation” OR “Angina Inestable”) AND (Machine Learning) AND NOT (“Image”).

Se registraron varias variables cuantitativas para cada artículo de investigación, que incluían el año de publicación, la revista, la tipología del documento, el cuartil de la revista, el país de publicación, el país de producción y la entidad o universidad que presentó el producto. Además, se registraron múltiples variables técnicas para describir el tipo de conjunto de datos utilizado en los varios trabajos de investigación consultados. Finalmente, los clasificadores a los que se hace referencia en los artículos, se incluye su enfoque, segmentación, representación, selección de características, equilibrio y adición de instancias, categoría y subcategoría. La última etapa consiste en presentar y analizar los resultados.

4. Análisis cuantitativo

Después de realizar el análisis de las variables cuantitativas de las 245 publicaciones que se encuentran en bases de datos de revistas indexadas de alto impacto, se utilizaron diferentes métodos para medir las publicaciones. Estos métodos incluyeron el año de publicación, el número de artículos publicados en cada base de datos, la tipología de los artículos y el cuartil de la revista, congreso o libro en la que se publicaron. Además, se examinaron los países que reciben y producen un mayor flujo de trabajos, así como las revistas y universidades que han avanzado más en este campo de investigación.

La Figura 2 muestran una tendencia creciente en el número de publicaciones relacionadas con enfermedades cardiovasculares en términos de: análisis de síndrome coronario agudo, descripción general de enfermedades cardiovasculares, uso de tecnología para identificar cardiopatías y sistemas de evaluación de riesgo cardiovascular. La Figura 2 (a) muestra cómo el número de publicaciones en este campo de investigación ha aumentado cada año, alcanzando su punto máximo en 2022 con 71 documentos. La figura 2 (b) indica que Science Direct es la base de datos científica con el mayor número de productos de investigación registrados entre 2019 y 2023.

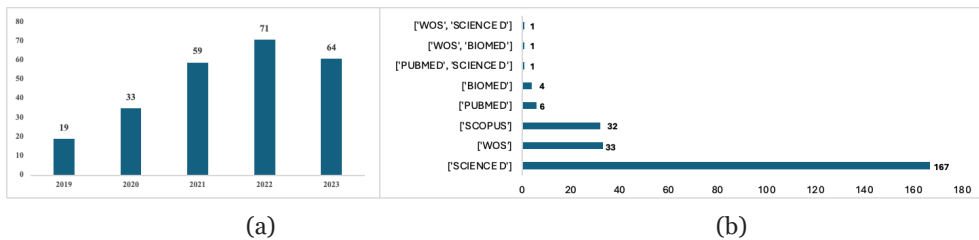


Figura 2 – (a) Cantidad de trabajos en el área de enfermedades cardiovasculares publicados en marco de tiempo, (b) Cantidad de trabajos en el área enfermedades cardiovasculares por

El 68,16% de las publicaciones en esta área de estudio se encuentran en revistas, las cuales se pueden acceder a través de la base de datos Science Direct, que es especializada en este tema. El 85% de las publicaciones accesibles desde diversas bases de datos especializadas son revistas, mientras que el 15% son libros, capítulos de libro y actas. El mayor porcentaje de publicaciones se realiza en recursos categorizados en primer cuartil (55%), segundo cuartil (20%), tercer cuartil (8%), cuarto cuartil (1%) y productos no categorizados (16%).

Con relación al número de publicaciones recibidas y generadas por país entre 2019 y 2024. El primer valor se obtuvo dividiendo las publicaciones por país de origen o edición de la revista, procedimiento o libro. De acuerdo con esto, se ha establecido que los países con el mayor número de publicaciones recibidas con relación al alcance del SCA son Canadá, Estados Unidos, Reino Unido y Holanda, entre otros. Para registrar las publicaciones según el país de origen, se determinó la universidad, centro de investigación u organización a la que están afiliados los autores. El país de generación se determinó principalmente utilizando el criterio más común entre todos los autores del

trabajo y, en algunos casos, el lugar de origen del primer autor. Como resultado, hemos descubierto que Canadá, los Estados Unidos, China y Japón son los países con la mayor cantidad de publicaciones generadas en base a SCA. La tabla 1 se muestran los recursos donde se publicaron más investigaciones sobre enfermedades cardiovasculares, en particular las relacionadas con el síndrome coronario agudo (SCA).

Revista	Publicaciones	País	ISSN	SJR	JCR
<i>Canadian Journal of Cardiology</i>	65	Canada	1916-7075	Q1	Q1
<i>American Journal of Cardiology</i>	16	United States	1879-1913	Q1	Q2
<i>IEEE Access</i>	9	United States	2169-3536	Q1	Q1
<i>BMC BIOINFORMATICS</i>	5	United Kingdom	1471-2105	Q1	NO
<i>Artificial Intelligence in Medicine</i>	5	Netherlands	1873-2860	Q1	Q1
<i>Trends in Cardiovascular Medicine</i>	5	United States	1050-1738	Q1	Q1
<i>Journal of Biomedical Informatics</i>	4	United States	1532-0464	Q1	Q2
<i>Journal of the American College of Cardiology</i>	4	United States	0735-1097	Q1	Q1

Tabla 1 – Revistas con más publicaciones

5. Análisis técnico

5.1. Categorías de análisis técnico

Las variables que influyen directamente en la caracterización temprana de biomarcadores y las tendencias en el uso de herramientas tic para la identificación, evaluación y predicción del riesgo cardiovascular, se identificaron al revisar los artículos seleccionados y tener en cuenta los marcadores de investigación. En este sentido, se han identificado 245 artículos científicos para el desarrollo de la investigación, con relación a la RSL y cadenas de búsqueda aplicadas, por lo que se han agrupado en cinco parámetros técnicos para su análisis: Caracterización del síndrome coronario agudo (SCA), enfermedad cardiovascular distinta del síndrome agudo de la enfermedad de las arterias coronarias (OCD), uso de tecnología para la identificación de enfermedades cardiovasculares (AITC) y Sistema de Evaluación de Riesgo Cardiovascular (RTCS). Se encontraron 62 artículos científicos para la caracterización del síndrome coronario agudo con una proporción del 25% del total de artículos identificados, de los 46 artículos científicos encontrados para enfermedades cardiovasculares diferentes al síndrome coronario agudo, con una proporción del 19%, 95 artículos científicos para aplicación de tecnología en la identificación de enfermedades del corazón con una proporción del 39% del total de artículos identificados y para el sistema de evaluación de riesgo cardiovascular se encontrar 42 artículos con una proporción del 17% del total de artículos identificados.

5.1.1. Caracterización del síndrome coronario agudo

Se encontraron 62 artículos para la caracterización del síndrome coronario agudo (SCA) y se dividieron en tres categorías: Identificación y Caracterización de patrones asociados al Síndrome Coronario Agudo (ICSCA), Clasificación del síndrome coronario agudo y patologías clínicas asociadas (CSCA), Historia clínica electrónica-EMR y Conjunto de datos clínicos tipo historia clínica electrónica (KDD). Se identificaron 27 artículos para ICSCA con una proporción del 44 %, 23 artículos para CSCA con una proporción del 37% y 12 para KDD con una proporción del 19%. En el análisis de doce comorbilidades de muestras de pacientes con tendencia médica al infarto de miocardio (Gulea, 2021; Kramer, 2019; Ordovas, 2021; Retnakaran, 2021), se examinó la viabilidad de utilizar métodos analíticos como los análisis de correlación PCA (Citro, 2020), LCA (Leonardi, 2021), Pearson (Harhash, 2021) e ICA (Khan, 2019) para clasificar la tendencia a sufrir infarto de miocardio solo basándose en las comorbilidades previas al evento cardíaco. En contraste, un estudio cuyo objetivo fue investigar en el ensayo MATRIX (Minimizing Adverse Hemorrhagic Events by Transradial Access Site and Systemic Implementation of Angiox) la incidencia, los predictores y las implicaciones pronósticas de la caída y los cambios de hemoglobina en pacientes con SCA manejados invasivamente estratificados por la presencia de sangrado intrahospitalario.

Además, en (Knowles, 2018) identifican los datos clínicos iniciales utilizados tradicionalmente para sospechar un SCA en pacientes con DTN en el servicio de urgencias; todos identifican a los pacientes con mayor riesgo de ser clasificados inicial y finalmente como SCA; algunos sobrestiman y otros subestiman inicialmente el riesgo final. De manera similar, en (Al Ghorani, 2021; Andreu, 2020; Lu, 2021; Palojoki, 2021) encuentran bioindicadores prometedores para enfermedades cardiovasculares agudas. Se ha demostrado en numerosos estudios clínicos que existe una correlación significativa entre los valores de sST₂ y los resultados del paciente. Es importante destacar que se ha encontrado una correlación entre los niveles más altos de sST₂ y un mayor riesgo de remodelado ventricular izquierdo adverso. Como resultado, sST₂ podría ser una herramienta útil para la estratificación de riesgo y el diagnóstico de pacientes que ingresan al servicio de urgencias.

5.1.2. Identificación de enfermedades cardiovasculares distintas del síndrome coronario agudo (SCA)

Los artículos de la investigación se dividieron en dos pilares o categorías: Descripción de enfermedades cardiovasculares distintas a la SCA (DOSCAD) y Criterios para la Identificación de Patologías (CIP). Se encontraron 24 artículos para Descripción de enfermedades cardiovasculares diferentes a la SCA (DOSCAD) con una relación del 52%, mientras que 22 artículos para criterios para la identificación de patologías (CIP) con una relación del 48%. En (Leslie, 2020) se realizó un estudio observacional prospectivo (ARTEMIS) en pacientes con síntomas lipídicos anormales para identificar y comparar la incidencia de muerte súbita cardíaca por diabetes y pacientes no diabéticos con infarto de miocardio. El mapeo electrofisiológico de alta densidad también se utilizó en (Brun, 2020; Hong, 2019) para detectar y caracterizar fenotípicamente la fibrilación ventricular idiopática en pacientes de edad avanzada.

5.1.3. Identificación de enfermedades cardiovasculares distintas del síndrome coronario agudo (SCA)

Se encontraron 95 artículos que se dividieron en cuatro categorías para el desarrollo de aplicaciones tecnológicas enfocadas en la identificación de enfermedades cardiovasculares: implementación de técnicas de aprendizaje automático para la predicción de enfermedades cardiovasculares diferentes a la SCA a nivel de pre-evento cardíaco (MLTSCA), de técnicas de aprendizaje automático para la predicción de SCA a nivel de pre-evento cardíaco (MLTSCABE), Problemas y limitaciones en conjuntos de datos (PLKDD) y Propuestas basadas en modelos estadísticos (OM). Se identificaron 32 artículos para la implementación de técnicas de aprendizaje automático para la predicción de enfermedades cardiovasculares diferentes a la SCA a nivel de pre-evento cardíaco (MLTSCA), que representaban el 34% de los artículos encontrados, mientras que 26 artículos para la implementación de técnicas de aprendizaje automático para la predicción de SCA a nivel de pre-evento cardíaco (MLTSCABE) con una proporción del 27% de los artículos encontrados. Para problemas y limitaciones en conjuntos de datos (PLKDD) se encontró un total de 18 artículos con una proporción del 19%, y para las propuestas basadas en modelos estadísticos (OM), un total de 19 artículos científicos con una proporción de 20%.

Algunas investigaciones relacionadas con la identificación de patrones (Blagova, 2019; Haïssaguerre, 2020), el preprocesamiento de datos (Dinesh, 2018), las técnicas de eliminación de datos ruidosos (Al'Aref, 2019), la eliminación de datos faltantes y la clasificación de atributos para la predicción (Mezzatesta, 2019) y la toma de decisiones como herramientas basadas en tecnologías de la información enfocadas en inteligencia artificial (Dinh, 2019; Karunathilake; 2018; Seetharam, 2019) y la construcción de modelos para el diagnóstico oportuno de enfermedades asociadas al síndrome coronario agudo son el método más relevante y utilizado para la identificación temprana de cardiopatías con especial atención a las enfermedades asociadas al síndrome coronario agudo (Ramalingam, 2018).

5.1.4. Sistema de evaluación del riesgo cardiovascular

Se encontraron 42 artículos para encontrar sistemas de evaluación del riesgo cardiovascular, de los cuales 3 pertenecían a tres categorías: Modelos enfocados a la identificación del riesgo cardiovascular (mundial) (MCVG), Modelos enfocados a la identificación del riesgo cardiovascular (Colombia) (MCVC) y Análisis de datos para la identificación (DAI). Se identificaron 21 artículos para los modelos enfocados a la identificación del riesgo cardiovascular (mundial) (MCVG), que representaron el 50% del total categorizado; 10 artículos para los modelos enfocados a la identificación del riesgo cardiovascular en Colombia (MCVC), que representaron el 24% del total encontrado; y 11 artículos para el Análisis de datos para la identificación (DAI), que representaron el 26% del total de artículos.

Algunas investigaciones relacionadas con los sistemas de evaluación del riesgo cardiovascular se centran en el análisis estadístico de la identificación de enfermedades utilizando diferentes métodos como SCORE (Everett, 2019), TIMI (Kao, 2020) y FRAMINGHAM (Komiyama, 2018) como métodos más efectivos para la identificación de dichas cardiopatías a nivel pre-evento cardíaco. De manera similar, en (Kao, 2020)

se explican diferentes herramientas para medir el riesgo cardiovascular donde existen marcadores de riesgo que aportan poco a la estratificación y las más utilizadas son las tablas de estratificación de riesgo. Además, es importante mencionar que las tablas de riesgo pueden evaluar y controlar preventivamente la hipertensión arterial y la dislipidemia (Chotechuang, 2020; Gernaat, 2018; Komiyama, 2018).

5.2. Conjuntos de datos y técnicas de minería de datos aplicadas para la predicción de enfermedades cardiovasculares

La minería de datos es el proceso de encontrar patrones y tendencias previamente desconocidos en las bases de datos y usar esa información para construir modelos predictivos (Murphy, 2020). En salud, la minería de datos es un campo de gran importancia y se ha vuelto cada vez más eficiente e imprescindible. En la actualidad, la industria de la salud genera grandes cantidades de datos complejos sobre pacientes, recursos hospitalarios, diagnósticos de enfermedades, registros electrónicos de pacientes, dispositivos, etc. La gran cantidad de datos es un recurso clave para ser procesado y analizado para la extracción de conocimiento y permitir el ahorro de costos, y toma de decisiones. La minería de datos proporciona un conjunto de herramientas y técnicas que se pueden aplicar a los datos para lograr estos objetivos (Lloyd-Jones, 2019).

Los conjuntos de datos que tratan los mismos problemas médicos, como la enfermedad de las arterias coronarias (CAD) (Alić, 2017), pueden mostrar resultados diferentes cuando se aplica la misma técnica de aprendizaje automático. Los resultados de precisión de la clasificación y las características importantes seleccionadas se basan principalmente en la eficiencia del diagnóstico y análisis médico. Estos conjuntos de datos brindan información complementaria al diagnóstico médico y, por lo tanto, pueden mejorar las tasas de precisión de la clasificación (Fatima, 2017; Pouriyeh, 2017; Nilashi, 2017; Ramalingam, 2018). El diagnóstico de los pacientes se basa en la precisión de los datos médicos recopilados. Las características comunes entre estos conjuntos de datos se extraen y utilizan en análisis posteriores para la misma enfermedad en cualquier conjunto de datos.

En (Abdar, 2019; Dwivedi, 2018), se denotan los conjuntos de datos con mayor implementación en áreas de construcción de modelos de inteligencia artificial, los cuales son elegidos por su efectividad, tales como: Cleveland Clinic Foundation, Instituto Húngaro de Cardiología, V.A. Medical Center, Long Beach, CA y registro de estadísticas del proyecto. El número de instancias para cada conjunto de datos es el siguiente: Cleveland: 303, Húngary: 294, Long Beach VA: 200 y Project Statlog: 270. Estos conjuntos de datos contienen 14 atributos que se extrajeron de un conjunto más grande de 75 atributos. Estos conjuntos de datos públicos se pueden descargar del repositorio de UCI para su implementación (The UCI Machine Learning Repository, n. d.).

6. Conclusión y trabajos futuros

6.1. Conclusión

El propósito de este artículo ha sido proponer una serie de recomendaciones para la investigación centrada en las enfermedades cardiovasculares en relación con la

identificación del conjunto de datos más adecuado según el tipo de investigación. El desarrollo en esta área de investigación en los últimos 5 años ha crecido mucho desde la presentación de 19 artículos de investigación en 2019 a 69 artículos publicados en 2023. La base de datos especializada más representativa, con relación al número de publicaciones científicas en SCA es Science Direct con un 68,16% de las publicaciones en total (167) superando la suma de los resultados obtenidos al consultar las bases de datos Wos, PubMed y BioMed.

Es importante resaltar que la mayor parte de las publicaciones en este campo de investigación son revistas con 208 artículos científicos con una relación total del 84,89% de las publicaciones encontradas. donde un porcentaje muy importante se realiza en revistas del primer (55% con 135 artículos) y segundo cuartil (20% con 49 artículos). Canadá y Estados Unidos se destacan como los países con mayor número de publicaciones aceptadas en esta área de investigación, con una participación de 26.94% para Canadá y 24.49% para Estados Unidos con relación al total de publicaciones recibidas de nivel mundial.

Aunque los medios en los que se publican los resultados de la investigación en esta área del conocimiento son muy diversos, los más destacados son: Canadian Journal of Cardiology (Canadá) con un 26,54% y American Journal of Cardiology (EE.UU.) con un 6,53%. También, se logró identificar que las instituciones con mayor experiencia en investigación en este campo son: Montreal Behavioral Medicine Centre, CIUSSS-NIM (Canadá) con una participación del 12,25% y Department of Health, Kinesiology, and Applied Physiology (HKAP), Concordia University (CANADÁ) con una participación del 8,57% respectivamente.

6.2. Trabajos futuros

Las tecnologías de la información y su contribución a la identificación de enfermedades cardiovasculares mediante la implementación de técnicas de inteligencia artificial es una de las soluciones que ha despegado en estos tiempos. En este sentido, se postulan los siguientes trabajos futuros: conjunto de datos clínicos basados en muestras médicas no invasivas, construcción de modelos de inteligencia artificial basados en redes neuronales artificiales, desarrollo de un sistema basado en el método wrapper donde se hibridan técnicas de selección de características. como chi cuadrado, ganancia de información, relación de ganancia con técnicas de soporte de máquina vectorial para crear un sistema con un alto nivel de precisión enfocado a la predicción de enfermedades cardiovasculares

Referencias

- Abdar, M., Książek, W., Acharya, U. R., Tan, R. S., Makarenkov, V., & Pławiak, P. (2019). A new machine learning technique for an accurate diagnosis of coronary artery disease. *Computer Methods and Programs in Biomedicine*, 179, 104992.
- Al'Aref, S. J., Anouchche, K., Singh, G., Slomka, P. J., Kolli, K. K., Kumar, A., ... & Min, J. K. (2019). Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *European Heart Journal*, 40(24), 1975-1986.

- Al Ghorani, H., Schwarz, V., Vukadinovic, D., Fries, P., Scheller, B., & Mahfoud, F. (2021). Myocardial infarction in a patient with single coronary artery-rare but real. *Journal of Cardiology Cases*, 23(5), 246-249.
- Aleksova, A., Paldino, A., Beltrami, A. P., Padoan, L., Iacoviello, M., Sinagra, G., ... & Maisel, A. S. (2019). Cardiac biomarkers in the emergency department: the role of soluble ST2 (sST2) in acute heart failure and acute coronary syndrome—there is meat on the bone. *Journal of Clinical Medicine*, 8(2), 270.
- Alić, B., Gurbeta, L., & Badnjević, A. (2017). Machine learning techniques for classification of diabetes and cardiovascular diseases. In 2017 6th Mediterranean Conference on Embedded Computing (MECO) (pp. 1-4). IEEE.
- Andreu, Ó. M., Nadal, G. M., Hernández, S. J., Gómez-Angelats, E., Alonso, J. R., Santaliestra, A. A., ... & Barbeito, B. L. (2020). Asociación entre los datos clínicos y electrocardiográficos iniciales en pacientes con dolor torácico no traumático y la sospecha inicial y el diagnóstico final de síndrome coronario agudo. *Emergencias: Revista de la Sociedad Española de Medicina de Urgencias y Emergencias*, 32(1), 9-18.
- Blagova, O., Osipova, Y., Nedostup, A., Kogan, E., Zaitsev, A., & Fomin, V. (2019). Diagnostic value of different noninvasive criteria of latent myocarditis in comparison with myocardial biopsy. *Cardiology*, 142(3), 167-174.
- Brandberg, H., Kahan, T., Spaak, J., Sundberg, K., Koch, S., Adeli, A., ... & Zakim, D. (2020). A prospective cohort study of self-reported computerised medical history taking for acute chest pain: protocol of the CLEOS-Chest Pain Danderyd Study (CLEOS-CPDS). *BMJ Open*, 10(1).
- Brun, F., Gigli, M., Graw, S. L., Judge, D. P., Merlo, M., Murray, B., ... & James, C. A. (2020). FLNC truncations cause arrhythmogenic right ventricular cardiomyopathy. *Journal of Medical Genetics*, 57(4), 254-257.
- Chotechuang, Y., Phrommintikul, A., Kuanprasert, S., Muenpa, R., Ruengorn, C., Patumanond, J., ... & Sukonthasarn, A. (2020). GRACE score and cardiovascular outcomes prediction among the delayed coronary intervention after post-fibrinolytic STEMI patients in a limited PCI-capable hospital. *Open Heart*, 7(1), e001133.
- Chu, D., Al Rifai, M., Virani, S. S., Brawner, C. A., Nasir, K., & Al-Mallah, M. H. (2020). The relationship between cardiorespiratory fitness, cardiovascular risk factors, and atherosclerosis. *Atherosclerosis*.
- Citro, R., Okura, H., Ghadri, J. R., Izumi, C., Meimoun, P., Izumo, M., ... & EACVI Scientific Documents Committee. (2020). Multimodality imaging in takotsubo syndrome: a joint consensus document of the European Association of Cardiovascular Imaging (EACVI) and the Japanese Society of Echocardiography (JSE). *European Heart Journal-Cardiovascular Imaging*, 21(11), 1184-1207. <https://doi.org/10.1093/ehjci>

- Comas-Gonzalez, Z., Mardini-Bovea, J., Salcedo, D., & De-la-Hoz-Franco, E. (2023). Comparison of Supervised Techniques of Artificial Intelligence in the Prediction of Cardiovascular Diseases. In H. Degen, S. Ntoa, & A. Moallem (Eds.), *HCI International 2023 – Late Breaking Papers. HCII 2023. Lecture Notes in Computer Science*, vol 14059. Springer, Cham. https://doi.org/10.1007/978-3-031-48057-7_4
- Crimi, G., Mandurino-Mirizzi, A., Gritti, V., Scotti, V., Strozzi, C., De Silvestri, A., ... & Klersy, C. (2020). Percutaneous coronary intervention techniques for bifurcation disease: network meta-analysis reveals superiority of double-kissing crush. *Canadian Journal of Cardiology*, 36(6), 906-914.
- Dinesh, K. G., Arumugaraj, K., Santhosh, K. D., & Mareeswari, V. (2018). Prediction of cardiovascular disease using machine learning algorithms. In 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) (pp. 1-7). IEEE.
- Dinh, A., Miertschin, S., Young, A., & Mohanty, S. D. (2019). A data-driven approach to predicting diabetes and cardiovascular disease with machine learning. *BMC Medical Informatics and Decision Making*, 19(1), 1-15.
- Dwivedi, A. K. (2018). Performance evaluation of different machine learning techniques for prediction of heart disease. *Neural Computing and Applications*, 29(10), 685-693.
- Everett, C. C., Fox, K. A., Reynolds, C., Fernandez, C., Sharples, L., Stocken, D. D., ... & Gale, C. P. (2019). Evaluation of the impact of the GRACE risk score on the management and outcome of patients hospitalised with non-ST elevation acute coronary syndrome in the UK: protocol of the UKGRIS cluster-randomised registry-based trial. *BMJ Open*, 9(9), e032165.
- Fatima, M., & Pasha, M. (2017). Survey of machine learning algorithms for disease diagnostic. *Journal of Intelligent Learning Systems and Applications*, 9(01), 1.
- García-Pérez, L. (2013). Systematic review of health-related utilities in Spain: the case of mental health. *Gaceta Sanitaria*, 28(1), 77-83. <https://doi.org/10.1016/j.gaceta.2013.04.006>
- Gernaat, S. A., Boer, J. M., van den Bongard, D. H., Maas, A. H., van der Pol, C. C., Bijlsma, R. M., ... & Peeters, P. H. (2018). The risk of cardiovascular disease following breast cancer by Framingham risk score. *Breast Cancer Research and Treatment*, 170(1), 119-127.
- González-Alcaide, G., Llorente, P., & Ramos-Rincón, J. M. (2020). Systematic analysis of the scientific literature on population surveillance. *Heliyon*, 6(10), e05141.
- Grams, M. E. (2011). Validation of CKD and related conditions in existing data sets: A systematic review. *American Journal of Kidney Diseases*, 57(1), 44-54. <https://doi.org/10.1053/j.ajkd.2010.05.013>

- Gulea, C., Zakeri, R., & Quint, J. K. (2021). Model-based comorbidity clusters in patients with heart failure: association with clinical outcomes and healthcare utilization. *BMC Medicine*, 19, 9. <https://doi.org/10.1186/s12916-020-01881-7>
- Haïssaguerre, M., Duchateau, J., Dubois, R., Hocini, M., Cheniti, G., Sacher, F., ... & Bernus, O. (2020). Idiopathic ventricular fibrillation: role of Purkinje system and microstructural myocardial abnormalities. *Clinical Electrophysiology*, 6(6), 591-608.
- Harhash, A. A., May, T. L., Hsu, C. H., Agarwal, S., Seder, D. B., Mooney, M. R., ... & Kern, K. B. (2021). Risk stratification among survivors of cardiac arrest considered for coronary angiography. *Journal of the American College of Cardiology*, 77(4), 360-371.
- Helwani, M. A., Amin, A., Lavigne, P., Rao, S., Oesterreich, S., Samaha, E., ... & Nagele, P. (2018). Etiology of acute coronary syndrome after noncardiac surgery. *Anesthesiology*, 128(6), 1084-1091.
- Henriquez, C. (2022). Predictive Model of Cardiovascular Diseases Implementing Artificial Neural Networks. In K. Saeed & J. Dvorský (Eds.), *Computer Information Systems and Industrial Management. CISIM 2022. Lecture Notes in Computer Science*, vol 13293. Springer, Cham. https://doi.org/10.1007/978-3-031-10539-5_17
- Hong, N., Wen, A., Stone, D. J., Tsuji, S., Kingsbury, P. R., Rasmussen, L. V., ... & Jiang, G. (2019). Developing a FHIR-based EHR phenotyping framework: A case study for identification of patients with obesity and multiple comorbidities from discharge summaries. *Journal of Biomedical Informatics*, 99, 103310.
- Iannattone, P. A., Zhao, X., VanHouten, J., Garg, A., & Huynh, T. (2020). Artificial intelligence for diagnosis of acute coronary syndromes: a meta-analysis of machine learning approaches. *Canadian Journal of Cardiology*, 36(4), 577-583.
- Idris, N. M., Chiam, Y. K., Varathan, K. D., Ahmad, W. A. W., Chee, K. H., & Liew, Y. M. (2020). Feature selection and risk prediction for patients with coronary artery disease using data mining. *Medical & Biological Engineering & Computing*, 58(12), 3123-3140. <https://doi.org/10.1007/s11517-020-02228-9>
- Instituto Nacional de Salud. (2023). 29 de septiembre - Día Mundial del Corazón. Recuperado el 5 de abril de 2024, de <https://www.ins.gov.co/Noticias/Paginas/29-de-septiembre---Dia-Mundial-del-Corazon.aspx>
- Kandasamy, S., & Anand, S. S. (2018). Cardiovascular disease among women from vulnerable populations: a review. *Canadian Journal of Cardiology*, 34(4), 450-457.
- Kao, Y. T., Hsieh, Y. C., Hsu, C. Y., Huang, C. Y., Hsieh, M. H., Lin, Y. K., & Yeh, J. S. (2020). Comparison of the TIMI, GRACE, PAMI and CADILLAC risk scores for prediction of long-term cardiovascular outcomes in Taiwanese diabetic patients with ST-segment elevation myocardial infarction: From the registry of the Taiwan Society of Cardiology. *PLoS ONE*, 15(2), e0229186.

- Karunathilake, S. P., & Ganegoda, G. U. (2018). Secondary prevention of cardiovascular diseases and application of technology for early diagnosis. *BioMed Research International*, 2018.
- Khan, M. U., Aziz, S., Iqtidar, K., Zainab, A., & Saud, A. (2019). Prediction of Acute Coronary Syndrome Using Pulse Plethysmograph. In 2019 4th International Conference on Emerging Trends in Engineering, Sciences and Technology (ICEEST) (pp. 1-6). IEEE.
- Knowles, J. W., & Ashley, E. A. (2018). Cardiovascular disease: The rise of the genetic risk score. *PLoS Medicine*, 15(3), e1002546. <https://doi.org/10.1371/journal.pmed.1002546>
- Kitchenham, B., Brereton, O. P., Budgen, D., Turner, M., Bailey, J., & Linkman, S. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51(1), 7–15. <https://doi.org/10.1016/j.infsof.2008.09.009>
- Komiyama, K., Nakamura, M., Tanabe, K., Niikura, H., Fujimoto, H., Oikawa, K., ... & Tokyo CCU Network Scientific Committee. (2018). In-hospital mortality analysis of Japanese patients with acute coronary syndrome using the Tokyo CCU Network database: Applicability of the GRACE risk score. *Journal of Cardiology*, 71(3), 251-258.
- Konys, A. (2019). Knowledge repository of ontology learning tools from text. *Procedia Computer Science*, 159, 1614-1628.
- Kramer, A. I., Trinder, M., & Brunham, L. R. (2019). Estimating the prevalence of familial hypercholesterolemia in acute coronary syndrome: a systematic review and meta-analysis. *Canadian Journal of Cardiology*, 35(10), 1322-1331.
- Kutyrev, K., Yakovlev, A., & Metsker, O. (2019). Mortality Prediction Based on Echocardiographic Data and Machine Learning: CHF, CHD, Aneurism, ACS Cases. *Procedia Computer Science*, 156, 97-104.
- Leonardi, S., Gagnano, F., Carrara, G., Gargiulo, G., Frigoli, E., Vranckx, P., ... & Valgimigli, M. (2021). Prognostic implications of declining hemoglobin content in patients hospitalized with acute coronary syndromes. *Journal of the American College of Cardiology*, 77(4), 375-388.
- Leslie, H. (2020). openEHR archetype use and reuse within multilingual clinical data sets: case study. *Journal of Medical Internet Research*, 22(11), e23361.
- Lloyd-Jones, D. M., Braun, L. T., Ndumele, C. E., Smith Jr, S. C., Sperling, L. S., Virani, S. S., & Blumenthal, R. S. (2019). Use of risk assessment tools to guide decision-making in the primary prevention of atherosclerotic cardiovascular disease: a special report from the American Heart Association and American College of Cardiology. *Circulation*, 139(25), e1162-e1177.

- Lu, J., Wang, L., Bennamoun, M., Ward, I., An, S., Sohel, F., ... & Sanfilippo, F. M. (2021). Machine learning risk prediction model for acute coronary syndrome and death from use of non-steroidal anti-inflammatory drugs in administrative data. *Scientific Reports*, 11(1), 1-10.
- Magoev, K., Krzhizhanovskaya, V. V., & Kovalchuk, S. V. (2018). Application of clustering methods for detecting critical acute coronary syndrome patients. *Procedia Computer Science*, 136, 370-379.
- Manterola, C., Astudillo, P., Arias, E., & Claros, N. (2013). Systematic reviews of the literature: What should be known about them. *Cirugía Española*, 91(3), 149-155. <https://doi.org/10.1016/j.ciresp.2011.07.009>
- Merlano-Porras, C. A., & Gorbanev, L. (2013). Health system in Colombia: A systematic review of literature. *Revista Gerencia Políticas Salud*, 12(24), 74-86.
- Mendes, D., Paredes, S., Rocha, T., Carvalho, P., Henriques, J., & Morais, J. (2018). Integration of current clinical knowledge with a data driven approach: An innovative perspective. *International Journal of Information Technology & Decision Making*, 17(01), 133-153.
- Metsker, O., Efimov, E., Trofimov, E., Kopanitsa, G., Bolgova, E., & Yakovlev, A. (2020). Performance improvement algorithms in big data analysis. *Procedia Computer Science*, 178, 386-393.
- Metsker, O., Bolgova, E., Yakovlev, A., Funkner, A., & Kovalchuk, S. (2017). Pattern-based mining in electronic health records for complex clinical process analysis. *Procedia Computer Science*, 119, 197-206.
- Mezzatesta, S., Torino, C., De Meo, P., Fiumara, G., & Vilasi, A. (2019). A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients
- Murphy, S. P., McCarthy, C. P., Cohen, J. A., Rehman, S., Jones-O'Connor, M., Olshan, D. S., ... & Wasfy, J. H. (2020). Application of the GRACE, TIMI, and TARRACO risk scores in type 2 myocardial infarction. *Journal of the American College of Cardiology*, 75(3), 344-345.
- Narayan, S. M., Wang, P. J., & Daubert, J. P. (2019). New concepts in sudden cardiac arrest to address an intractable epidemic: JACC state-of-the-art review. *Journal of the American College of Cardiology*, 73(1), 70-88. <https://doi.org/10.1016/j.jacc.2018.09.068>
- Nilashi, M., bin Ibrahim, O., Ahmadi, H., & Shahmoradi, L. (2017). An analytical method for diseases prediction using machine learning techniques. *Computers & Chemical Engineering*, 106, 212-223.
- Ordovas, K. G., Baldassarre, L. A., Bucciarelli-Ducci, C. (2021). Cardiovascular magnetic resonance in women with cardiovascular disease: position statement from the Society for Cardiovascular Magnetic Resonance (SCMR). *Journal of Cardiovascular Magnetic Resonance*, 23, 52. <https://doi.org/10.1186/s12968-021-00746-z>

- Organización Mundial de la Salud. (2023). Enfermedades cardiovasculares (ECV). Recuperado el 5 de abril de 2024, de [https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/es/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
- Organización Panamericana de la Salud. (2023). Hipertensión. Recuperado el 5 de abril de 2024, de <https://www.paho.org/es/temas/hipertension>
- Palojoki, S., Saranto, K., Reponen, E., Skants, N., Vakkuri, A., & Vuokko, R. (2021). Classification of Electronic Health Record–Related Patient Safety Incidents: Development and Validation Study. *JMIR Medical Informatics*, 9(8), e30470.
- Pouriyeh, S., Vahid, S., Sannino, G., De Pietro, G., Arabnia, H., & Gutierrez, J. (2017). A comprehensive investigation and comparison of machine learning techniques in the domain of heart disease. In 2017 IEEE Symposium on Computers and Communications (ISCC) (pp. 204-207). IEEE.
- Ramalingam, V. V., Dandapath, A., & Raja, M. K. (2018). Heart disease prediction using machine learning techniques: a survey. *International Journal of Engineering & Technology*, 7(2), 684-687.
- Retnakaran, R. (2018). Novel biomarkers for predicting cardiovascular disease in patients with diabetes. *Canadian Journal of Cardiology*, 34(5), 624-631.
- Rezaianzadeh, A., Dastoorpoor, M., Sanaei, M., Salehnasab, C., Mohammadi, M. J., & Mousavizadeh, A. (2020). Predictors of length of stay in the coronary care unit in patient with acute coronary syndrome based on data mining methods. *Clinical Epidemiology and Global Health*, 8(2), 383-388.
- Sanchez, A., Neira, D., & Cabello, J. J. (2016). Frameworks applied in quality management—A systematic review. *Rev. Espacios*, 37(9), 17.
- Seetharam, K., Shrestha, S., & Sengupta, P. P. (2019). Artificial intelligence in cardiovascular medicine. *Current Treatment Options in Cardiovascular Medicine*, 21(5), 1-14.
- Si, Y., Bernstam, E. V., & Roberts, K. (2021). Generalized and transferable patient language representation for phenotyping with limited data. *Journal of Biomedical Informatics*, 116, 103726.
- Sun, Q., Wang, Q., Ji, B., Wu, W., Huang, W., & Wang, C. (2020). The Cardiodynamicsgram Based Early Detection of Myocardial Ischemia Using the Lempel-Ziv Complexity. *IEEE Access*, 8, 207894-207904.
- The UCI Machine Learning Repository. (n. d.). Heart Disease Data Set. Retrieved from <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>
- Yao, L., Liu, C., Li, P., Wang, J., Liu, Y., Li, W., ... & Zhang, H. (2020). Enhanced automated diagnosis of coronary artery disease using features extracted from QT interval time series and ST-T waveform. *IEEE Access*, 8, 129510-129524. <https://doi.org/10.1109/ACCESS.2020.3001787>

- Xie, L., Li, Z., Zhou, Y., He, Y., & Zhu, J. (2020). Computational Diagnostic Techniques for Electrocardiogram Signal Analysis. *Sensors*, 20(21), 6318.
- Xue, Y., Hu, Z., Jing, Y., Wu, H., Li, X., Wang, J., ... & Lv, Q. (2020). Efficacy assessment of ticagrelor versus clopidogrel in Chinese patients with acute coronary syndrome undergoing percutaneous coronary intervention by data mining and machine-learning decision tree approaches. *Journal of Clinical Pharmacy and Therapeutics*, 45(5), 1076-1086.