

MODELO PREDICTIVO PARA LA DETECCIÓN PRECOZ DE INFARTOS EN LA POBLACIÓN DEL NORTE DE PERÚ

PREDICTIVE MODEL FOR THE EARLY DETECTION OF HEART ATTACKS IN THE POPULATION OF NORTHERN PERU

Henry Silva-Marchan ^{1*}, José Pardo Garces ¹, Raúl Sánchez Ancajima ¹, Gabriela Barrera Rea ²

¹Universidad Nacional de Tumbes, Tumbes, 24001, Perú

²Universidad Estatal de Milagro, Milagro, 091701, Ecuador

*Autor de Correspondencia: hsilvam@untumbes.edu.pe

Historial del artículo:

Recibido: 10.06.2025

Revisado: 27.06.2025

Aceptado: 10.07.2025

En línea: 15.07.2025

Palabras clave:

Redes neuronales artificiales

Modelo predictivo

Infarto de miocardio

Inteligencia artificial

Salud pública

Keywords:

Artificial neural networks

Predictive model

Myocardial infarction

Artificial intelligence

Public health

Cómo citar:

Silva-Marchan, H., Pardo H. Silva-Marchan, J. Pardo Garces, R. Sánchez Ancajima, y G. Barrera Rea, «Modelo predictivo para la detección precoz de infartos en la población del norte de Perú», *International Journal of Computational Innovations, Intelligent Systems and AI*, vol. 1, n.º 1, pp. 62–77, jul. 2025. doi: [10.64439/cisai.v1i1.9](https://doi.org/10.64439/cisai.v1i1.9)

Este es un artículo de acceso abierto bajo licencia CC-BY-NC-ND



Resumen:

Las enfermedades cardiovasculares representan una de las principales causas de mortalidad a nivel global, y los infartos de miocardio constituyen una de las manifestaciones más críticas de esta problemática. En la región norte del Perú, la limitada infraestructura hospitalaria y el acceso insuficiente a diagnósticos especializados incrementan la vulnerabilidad de la población. El presente estudio propone el desarrollo de un modelo predictivo basado en Redes Neuronales Artificiales (RNA) para la detección temprana de infartos. Se trabajó con datos secundarios de la encuesta demográfica y de salud familiar del periodo 2005–2024, con un total de 573,080 registros que incluyen factores de riesgo como hipertensión, diabetes, obesidad, tabaquismo y consumo de alcohol. La metodología contempló la limpieza, integración y normalización de datos, seguida de la implementación de un modelo de RNA en Python, validado mediante métricas de precisión, sensibilidad y especificidad. Los resultados preliminares evidencian una precisión predictiva superior al 90%, lo que respalda la viabilidad del uso de inteligencia artificial en la medicina preventiva, particularmente en escenarios caracterizados por una alta demanda asistencial.

Abstract:

Cardiovascular diseases represent one of the leading causes of mortality worldwide, and myocardial infarctions constitute one of the most critical manifestations of this problem. In the northern region of Peru, limited hospital infrastructure and insufficient access to specialized diagnostics increase the population's vulnerability. This study proposes the development of a predictive model based on Artificial Neural Networks (ANN) for the early detection of heart attacks. Secondary data from the demographic and family health survey for the period 2005–2024 were used, comprising a total of 573,080 records that include risk factors such as hypertension, diabetes, obesity, smoking, and alcohol consumption. The methodology involved data cleaning, integration, and normalization, followed by the implementation of an ANN model in Python, validated using accuracy, sensitivity, and specificity metrics. Preliminary results show a predictive accuracy above 90%, supporting the feasibility of applying artificial intelligence in preventive medicine, particularly in scenarios characterized by high healthcare demand.

I. INTRODUCCIÓN

Las enfermedades cardiovasculares constituyen uno de los principales desafíos para la salud pública mundial, siendo responsables de aproximadamente 17,9 millones de muertes anuales, lo que equivale a un 32 % del total de fallecimientos [1]. Entre ellas, los infartos de miocardio representan una de las manifestaciones clínicas más críticas debido a su elevada letalidad y a las secuelas que comprometen de manera permanente la calidad de vida de los pacientes [2]. Según Vázquez-Troche et al. [3] el infarto agudo de miocardio representa la mayor proporción de muertes por enfermedades isquémicas del corazón en Perú, con un 88.16% de las muertes atribuidas entre 2005 y 2017. Esta situación se intensifica cada vez más en la zona norte del Perú, donde la poca inversión en infraestructura hospitalaria, el déficit de personal especializado y las limitaciones en cuanto al equipamiento tecnológico disminuyen la capacidad de detección temprana, generando diagnósticos tardíos y por ende elevadas tasas de mortalidad.

Los métodos comúnmente utilizados en la práctica clínica, como los electrocardiogramas o determinación de biomarcadores en la sangre, constituyen herramientas fundamentales, pero nos muestran limitaciones en su capacidad de predicción [4][5]. Estos procedimientos, permiten identificar el evento una vez desencadenado, lo que limita la capacidad de realizar una intervención temprana y aumenta el riesgo de daño irreversible en el tejido cardíaco. En este escenario, la incorporación de tecnologías basadas en Inteligencia Artificial (IA), con el uso en particular de las Redes Neuronales Artificiales (RNA), se presentan como una alternativa tecnológica capaz de procesar grandes volúmenes de datos clínicos y epidemiológicos para poder identificar patrones complejos y no lineales que se asocian al riesgo de infarto [6].

En la actualidad, los modelos predictivos basados en RNA han demostrado una notable eficacia en la detección temprana de eventos cardiovasculares. Irsyad et al. [7] señalan que la aplicación de técnicas de Deep Learning (DL) mejora de manera significativa la precisión de los diagnósticos en comparación con los métodos tradicionales. En particular, los modelos de aprendizaje profundo, como las redes neuronales convolucionales (CNN), han mostrado un rendimiento superior en el análisis de electrocardiogramas. Por ejemplo, un modelo de CNN alcanzó una precisión del 98,29 %, superando ampliamente a enfoques convencionales como la regresión logística y las máquinas de vectores de soporte [7]. Asimismo, se ha evidenciado que arquitecturas avanzadas de Deep Learning, como Modified DenseNet201 y redes residuales profundas mejoradas, alcanzan precisiones de hasta el 99,12 % [8].

1.1 Problemática

Las enfermedades cardiovasculares, en particular los infartos de miocardio constituyen un factor de amenaza preocupante para la salud pública en el mundo, no solo por su alto índice de mortalidad sino también por los factores de secuelas físicas, psicológicas y sociales que generan en los sobrevivientes. En el Perú, el Ministerio de Salud viene reportando que, de una cuarta parte de las muertes anuales en su población, se atribuyen a las enfermedades cardiovasculares, situándose como la primera causa de mortalidad [9]. Esta realidad se ha intensificado en la zona norte del Perú, donde la precaria infraestructura hospitalaria y la brecha existente de especialistas disponibles en cardiología disminuyen gravemente la capacidad de respuesta ante episodios agudos. En las ciudades como Piura, Tumbes y Chiclayo, los servicios de salud son precarios y sobre todo limitados al encontrarse saturados constantemente y dependen en gran medida de realizar sólo electrocardiogramas o análisis bioquímicos, que casi siempre no permiten una detección temprana del riesgo. Teniendo como consecuencia que numerosos pacientes sean diagnosticados en una fase ya avanzada, cuando las posibilidades de intervención se encuentra reducida y el pronóstico vital se torna incierto y no sería efectivo.

Este retraso diagnóstico produce elevados costos significativos para el sistema de salud peruano, al incrementarse la demanda de cuidados intensivos y resalta las desigualdades de oportunidades de acceso a un servicio de salud eficiente y eficaz en las poblaciones vulnerables. Esta problemática no se encuentra restringida únicamente a una situación de recursos físicos o humanos, sino que también refleja una brecha tecnológica, en plena era digital. Mientras que países de primer mundo ya incorporan modelos predictivos basados en IA que permiten identificar patrones de riesgo con una elevada precisión. En el Perú estas herramientas aún no han sido implementadas de forma sistemática, las RNA, en particular, han demostrado ser capaces de procesar grandes volúmenes de datos clínicos y epidemiológicos, evidenciando las relaciones no lineales que no pueden detectarse con los métodos tradicionales [10]. Sin embargo, el no uso en los sistemas hospitalarios de la región norte del Perú, mantiene este círculo vicioso que se refleja en la demora de diagnósticos, incremento en la mortalidad y falta de respuesta institucional.

Ante este escenario, surge la necesidad inmediata de plantear e implementar soluciones innovadoras que utilicen la IA en la práctica médica. El uso de RNA, aplicada a bases de datos nacionales, como en la Encuesta Demográfica y de Salud Familiar (ENDES), representa una ventaja estratégica para obtener modelos predictivos contextualizados, los cuales sean capaces de anticipar los riesgos de infarto antes de que ocurran. Lo cual no solo podría salvar más vidas mediante intervenciones tempranas, sino contribuir en la optimización de la distribución de recursos en un sistema de salud históricamente limitado.

Ante este panorama, se hace evidente la necesidad inmediata de diseñar e implementar soluciones innovadoras que integren la inteligencia artificial en la práctica médica. El uso de redes neuronales artificiales aplicadas a bases de datos nacionales, como la ENDES, constituye una ventaja estratégica para el desarrollo de modelos predictivos contextualizados a la realidad peruana. Dichos modelos no solo tendrían la capacidad de anticipar el riesgo de infarto antes de su ocurrencia, posibilitando intervenciones tempranas y más efectivas, sino que también contribuirían a optimizar la asignación de recursos en un sistema de salud caracterizado por limitaciones históricas en infraestructura y cobertura. En este sentido, la incorporación de herramientas de IA en la salud pública no solo fortalecería la capacidad diagnóstica, sino que también representa un paso hacia la consolidación de políticas de prevención más eficientes y sostenibles, orientadas a reducir la carga de las enfermedades cardiovasculares en la población.

1.2 Trabajos relacionados

En esta subsección se expone una revisión de investigaciones recientes centradas en el desarrollo de modelos predictivos para la detección precoz del infarto agudo de miocardio, con énfasis en su aplicabilidad al contexto del norte del Perú. Se incluyen estudios que utilizan enfoques estadísticos y de inteligencia artificial, así como investigaciones que consideran variables clínicas, sociodemográficas y geográficas. La selección prioriza artículos con evidencia empírica en regiones andinas, rurales o de baja cobertura médica, similares a las del norte peruano. Esta revisión permite contextualizar los avances existentes y fundamentar la pertinencia de nuevas propuestas adaptadas a poblaciones locales.

En el ámbito del electrocardiograma (ECG), Acharya et al. [11] propusieron uno de los primeros modelos de aprendizaje profundo para detectar infarto de miocardio sin recurrir a ingeniería de características manual. Su red neuronal convolucional aprendió patrones morfológicos directamente de señales crudas de ECG, demostrando que la clasificación entre trazos normales y con infarto puede automatizarse con buen rendimiento en bases públicas. Sin embargo, este enfoque se centra en la fase aguda, cuando el evento ya ha ocurrido, lo que limita su utilidad en la prevención poblacional. A diferencia de este trabajo, nuestro modelo busca adelantarse al episodio crítico mediante la identificación temprana de individuos en riesgo, incorporando factores sociodemográficos y clínicos desde bases de datos nacionales como ENDES.

La investigación de Gustafsson et al. [12] avanzó hacia la validación externa de modelos basados en ECG, integrando variables como edad y sexo en la predicción de infartos en pacientes que llegaban a emergencias. Su valor radica en agilizar el triaje clínico en entornos hospitalarios de alta presión, donde la decisión debe tomarse en minutos. No obstante, este tipo de aproximaciones sigue dependiendo de la presencia de síntomas y del acceso inmediato a equipamiento especializado. En contraste, nuestra propuesta no se restringe al contexto hospitalario, sino que aprovecha encuestas poblacionales masivas, permitiendo anticipar riesgos incluso en sujetos que aún no han tenido un contacto clínico directo.

En diagnóstico bioquímico, Doudehis et al. [13] introdujeron el algoritmo CoDE-ACS, que integra niveles de troponina y tiempo de síntomas para estimar la probabilidad de infarto. Su fortaleza es mejorar la estratificación en pacientes de admisión temprana, evitando pruebas repetidas. Sin embargo, al igual que en los modelos basados en ECG, depende de biomarcadores de laboratorio, cuya disponibilidad es limitada en regiones periféricas como el norte del Perú. Nuestra investigación se diferencia al prescindir de pruebas invasivas o costosas, orientándose a predicciones basadas en factores de riesgo reportados en encuestas nacionales, lo que aumenta la escalabilidad y la aplicabilidad en contextos de recursos restringidos.

Toprak et al. [14] evaluaron un algoritmo de aprendizaje automático aplicado a troponina I ultrasensible en dispositivos “point-of-care”, que permite una exclusión rápida de infarto en áreas sin laboratorios centrales. Esta innovación traslada el poder de la IA a la cabecera del paciente, aunque sigue requiriendo equipamiento específico y consumibles regulares. Frente a esta dependencia tecnológica, nuestro modelo representa una alternativa preventiva más amplia, pues se entrena sobre millones de registros poblacionales con variables de riesgo acumuladas a lo largo del tiempo. Esto abre la posibilidad de políticas de salud preventiva a nivel regional, en lugar de limitarse al diagnóstico reactivo en salas de emergencia.

Khera et al. [15] compararon modelos de aprendizaje automático con técnicas tradicionales para predecir mortalidad post-infarto, concluyendo que ML ofrece mejor resolución individual y recalibración dinámica, aunque no siempre supera a los métodos clásicos en métricas globales. Este hallazgo resalta la importancia de construir sistemas robustos, interpretables y validados externamente. Nuestro modelo se distancia de este enfoque porque no se enfoca únicamente en pacientes ya diagnosticados, sino en la población general en riesgo, lo que lo convierte en una herramienta de medicina preventiva. Al integrar variables epidemiológicas, factores conductuales y biomarcadores indirectos, buscamos anticipar la ocurrencia del infarto antes de que el daño miocárdico se produzca, ampliando el horizonte de acción en comparación con las investigaciones previas.

En síntesis, los estudios revisados evidencian que los avances en modelos predictivos para infarto se han orientado principalmente a escenarios clínicos agudos, ya sea mediante el análisis de ECG, el uso de biomarcadores como la troponina o la predicción de riesgo tras el evento. Si bien estos enfoques han demostrado eficacia en entornos hospitalarios, presentan limitaciones en su alcance poblacional y en la prevención temprana en contextos de recursos restringidos como el norte del Perú. La presente investigación se diferencia al aprovechar una base de datos poblacional amplia como la ENDES y aplicar redes neuronales artificiales para identificar de manera anticipada a los individuos con mayor riesgo, contribuyendo así a la medicina preventiva. En la siguiente subsección se presentarán la motivación y la contribución central de este estudio.

1.3 Motivación y contribución

La alta incidencia de infartos en el norte del Perú no solo representa una estadística alarmante, sino también una realidad dolorosa que afecta a miles de familias cada año. Detrás de cada número se encuentran historias de vidas interrumpidas de manera prematura, hogares que enfrentan la ausencia de un padre, una madre o un hijo, y comunidades que pierden valioso capital humano. Esta situación genera una reflexión profunda sobre la urgencia de actuar, no únicamente desde la medicina curativa, sino desde la prevención apoyada en la ciencia y la innovación tecnológica. La motivación de este estudio surge, precisamente, de esa necesidad de transformar la forma en que se enfrenta la salud cardiovascular. Más allá de la búsqueda de resultados académicos, existe el compromiso de aportar con una herramienta que pueda tener un impacto real en la vida de las personas, especialmente en una región históricamente marcada por limitaciones en infraestructura de salud, desigualdad en el acceso a servicios médicos y diagnósticos que suelen llegar tarde. En este sentido, la inteligencia artificial a través de la RNA se presenta como una oportunidad para cerrar brechas, democratizar el acceso a la prevención y brindar a la población norteña una posibilidad real de anticiparse al riesgo de infarto.

La contribución de esta investigación se materializa en el diseño de un modelo predictivo basado en RNA, entrenado con una base robusta de 573,080 registros de la Encuesta Demográfica y de Salud Familiar correspondientes al periodo 2005–2024. Este modelo tiene como objetivo principal anticipar el riesgo de infarto en la población del norte del Perú, a partir de la identificación de patrones complejos en factores de riesgo como hipertensión, diabetes, obesidad, tabaquismo y consumo de alcohol. A diferencia de enfoques tradicionales, la propuesta busca ofrecer predicciones contextualizadas a la realidad peruana, lo que incrementa la posibilidad de implementar intervenciones tempranas y efectivas. El impacto esperado de este trabajo trasciende la dimensión técnica. Su verdadero valor radica en el potencial de salvar vidas mediante la detección precoz de eventos cardiovasculares, contribuir a la reducción de la carga económica que generan en un sistema de salud históricamente limitado y proporcionar evidencia científica que sirva de base para políticas públicas más efectivas y sostenibles en materia de prevención cardiovascular.

II. MÉTODOS

En esta subsección se describen los métodos aplicados para la construcción y validación del modelo predictivo, asegurando su solidez científica y su aplicabilidad práctica en la detección temprana de infartos en la población del norte del Perú. La investigación se estructuró con el objetivo de superar las limitaciones de estudios previos enfocados en contextos hospitalarios o en pruebas diagnósticas de alto costo, adoptando un enfoque preventivo fundamentado en datos poblacionales masivos de la ENDES. Este diseño permite anticipar riesgos en comunidades con limitado acceso a servicios clínicos especializados y fortalece la perspectiva de la medicina preventiva. Para cumplir este propósito se desarrollaron tres etapas metodológicas.

- Se efectuó la recolección, limpieza y normalización de más de medio millón de registros relacionados con factores de riesgo como hipertensión, diabetes, obesidad, tabaquismo y consumo de alcohol.
- Se diseñó una arquitectura de RNA optimizada para procesar datos heterogéneos de naturaleza clínica y sociodemográfica
- Se implementó el entrenamiento y la validación del modelo mediante métricas robustas como precisión, sensibilidad y especificidad.

2.1. Recolección y preparación de datos

En esta primera fase se procedió a la identificación y selección de la fuente de información que sustentaría el modelo predictivo. Se empleó la ENDES, correspondiente al periodo 2005–2024, que ofrece 573,080 registros individuales y garantiza una base representativa para el análisis poblacional. La elección de la ENDES se justifica porque recoge información sociodemográfica y clínica de relevancia, permitiendo trabajar con factores de riesgo reconocidos en la literatura científica. Se incorporaron variables ampliamente asociadas con la ocurrencia de infartos (véase Tabla 1). Entre ellas destacan el diagnóstico de hipertensión arterial y las mediciones de presión arterial sistólica y diastólica, que constituyen predictores fundamentales de riesgo cardiovascular. Asimismo, se consideraron características sociodemográficas como la edad, el sexo, la región geográfica, el área de residencia y el tamaño de la localidad, que permiten establecer perfiles diferenciales de vulnerabilidad. También se integraron variables de hábitos de vida como el tabaquismo (años fumando, número de cigarrillos por día y consumo en los últimos 12 meses) y el consumo de alcohol (edad de inicio y frecuencia). En cuanto a las medidas antropométricas, se incluyeron el peso corporal, la talla y el perímetro abdominal, que reflejan el estado nutricional y la obesidad abdominal, ampliamente vinculada con enfermedades cardiovasculares. Finalmente, se incorporaron patrones de alimentación saludable, como la frecuencia semanal de consumo de jugos de frutas, ensaladas de frutas y ensaladas de verduras, variables que permiten valorar la calidad de la dieta.

Tabla 1. Variables que conforman el conjunto de datos

Variable	Descripción	Tipo	Valores posibles / Rango
Year	Año de observación	Numérico	2005 – 2024
QS182	Diagnóstico de hipertensión	Categórica	1 = Sí, 2 = No
QS23	Edad de la persona	Numérico	15 – 89 años
QSEXO	Sexo de la persona	Categórica	1 = Hombre, 2 = Mujer
QS208	Fumó cigarrillos en los últimos 12 meses	Categórica	1 = Sí, 2 = No
QS284C	Años que ha fumado	Numérico	0 – 39 años
QS205C	Número de cigarrillos que fuma por día	Numérico	0 – 39 cigarrillos
QS286	Alguna vez ha consumido bebidas alcohólicas	Categórica	1 = Sí, 2 = No
QS207C	Edad en que empezó a consumir alcohol	Numérico	10 – 59 años
QS084C	Peso corporal	Numérico	~40 – 140 kg (normal con desviación)
QS085	Talla corporal	Numérico	~140 – 190 cm (normal con desviación)
QS087	Perímetro abdominal	Numérico	~60 – 130 cm (normal con desviación)
QS832	Presión arterial sistólica (1ra medición)	Numérico	100 – 179 mmHg
QS833	Presión arterial diastólica (1ra medición)	Numérico	60 – 109 mmHg
QS835	Presión arterial sistólica (2da medición)	Numérico	100 – 179 mmHg
QS836	Presión arterial diastólica (2da medición)	Numérico	60 – 109 mmHg
HV824	Región geográfica	Categórica	6 = Cajamarca, 13 = La Libertad, 14 = Lambayeque, 20 = Piura, 24 = Tumbes
HV025	Área de residencia	Categórica	1 = Urbana, 2 = Rural
HV026	Tamaño de localidad	Categórica	0 = Capital, 1 = Ciudad, 2 = Pueblo, 3 = Campo
Q5215C	Días por semana que tomó jugo de frutas	Numérico	0 – 6 días
Q5215D	Vasos de jugo consumidos al día	Numérico	0 – 4 vasos
Q5217C	Días por semana que comió ensalada de frutas	Numérico	0 – 6 días
Q5219C	Días por semana que comió ensalada de verduras	Numérico	0 – 6 días

Teniendo en cuenta las variables antes expuestas, se aplicaron procedimientos sistemáticos de limpieza y preparación de datos para asegurar la calidad de la información antes de ser utilizada por la RNA. En esta etapa, se eliminaron los registros incompletos o con inconsistencias, con el fin de reducir el impacto de valores atípicos que podrían distorsionar el aprendizaje del modelo. Luego, los datos faltantes fueron imputados mediante métodos de interpolación estadística, evitando la reducción drástica del tamaño muestral y preservando la validez de los patrones poblacionales. Posteriormente, se normalizaron las variables cuantitativas, de manera que los diferentes factores de riesgo compartieran una escala común, mejorando la estabilidad y la velocidad de convergencia de la RNA durante el entrenamiento. Con este procedimiento, cada valor numérico fue ajustado en rangos comparables, lo cual previene que variables con magnitudes mayores dominen en el proceso de aprendizaje.

2.2. Diseño del modelo de RNA

En la segunda fase de la investigación se procedió al diseño de la arquitectura de la red neuronal artificial, orientada a la construcción de un sistema predictivo capaz de identificar patrones complejos asociados al riesgo de infarto en la población del norte del Perú. Se optó por una RNA multicapa debido a su reconocida capacidad para procesar variables heterogéneas y modelar relaciones no lineales entre los factores de riesgo, lo que resulta especialmente pertinente en un contexto donde intervienen datos clínicos, antropométricos, sociodemográficos y de hábitos de vida [16]. La capa de entrada estuvo conformada por las 23 variables seleccionadas en la fase de preparación de datos, garantizando que cada dimensión relevante participara en el proceso de predicción y que la información se integrara de manera coherente (Figura 3). Posteriormente, se incorporaron dos capas ocultas densas, con un rango de entre 64 y 128 neuronas cada una, activadas mediante la función ReLU (Rectified Linear Unit), que facilita la convergencia eficiente y contribuye a mantener estabilidad numérica durante el entrenamiento [17]. Por último, la capa de salida se configuró en función del problema de clasificación binaria presencia o ausencia de riesgo de infarto, empleando la activación sigmoide para obtener una probabilidad ajustada al rango [0,1].

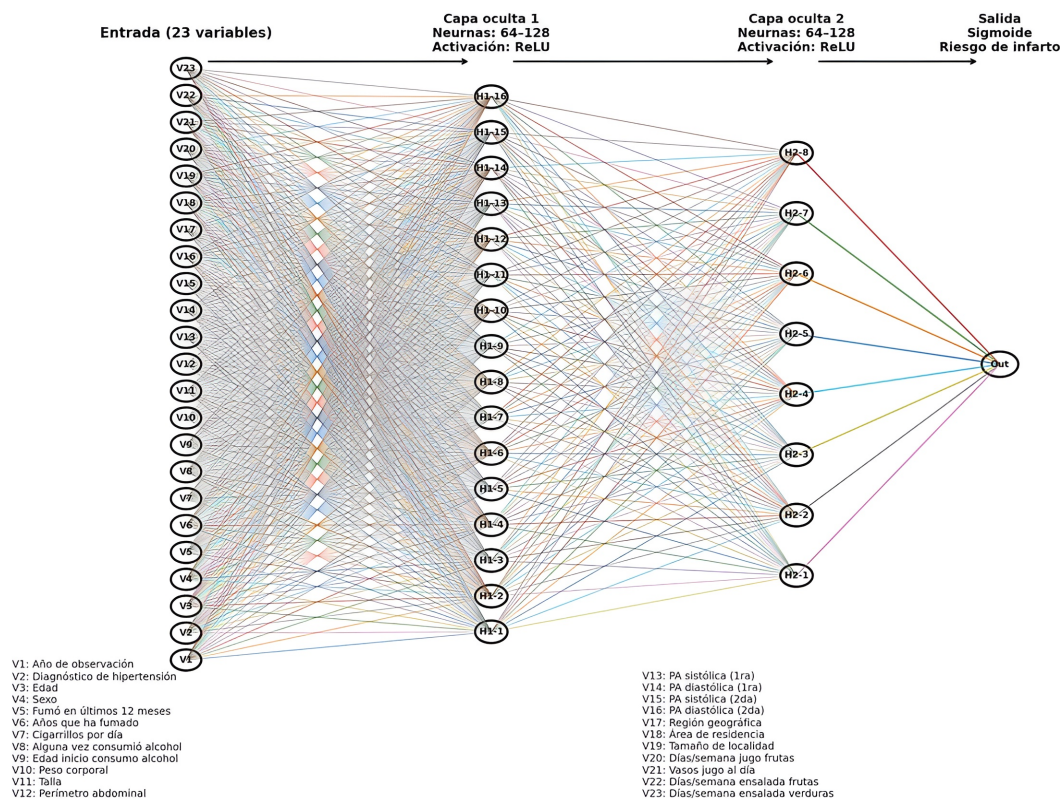


Figura 1. Esquema de la arquitectura del modelo de RNA para predicción de infartos.

Esta configuración fue seleccionada tras un análisis exploratorio de distintos tamaños de red, asegurando un equilibrio entre complejidad computacional y capacidad predictiva. Por último, la capa de salida se constituyó por una única neurona con función de activación sigmoide, adecuada para problemas de clasificación binaria, donde el valor 1 correspondió a alto riesgo de infarto y el valor 0 a bajo riesgo. La estructura general de este modelo se presenta en la Figura 1, que muestra el flujo de información desde las variables de entrada hasta la salida final, ilustrando cómo cada capa procesa la información y permite la transformación progresiva de los datos en una predicción clínica significativa. Para el entrenamiento de la RNA se empleó el algoritmo de retropropagación del error (backpropagation), ampliamente validado en la literatura, optimizado con el método Adam, que ofrece un equilibrio entre velocidad de convergencia y capacidad de adaptación en escenarios con múltiples variables. Se estableció una tasa de aprendizaje inicial de 0.001, ajustada experimentalmente para garantizar un rendimiento estable. Con el propósito de evitar sesgos derivados del sobreajuste, se aplicó validación cruzada k-fold con k=10, lo cual permitió evaluar la consistencia del modelo en diferentes particiones de los datos y obtener estimaciones más robustas de desempeño. Este diseño asegura que la RNA pueda no solo reconocer los factores de riesgo más relevantes, sino también integrar la compleja interacción entre ellos, dotando al modelo de un mayor poder predictivo en comparación con técnicas estadísticas tradicionales. En consecuencia, el diseño de esta arquitectura constituye un paso esencial en la construcción del modelo propuesto, garantizando una base sólida para las etapas de entrenamiento, validación y posterior análisis de desempeño.

2.3. Entrenamiento y validación

En esta fase se llevó a cabo el proceso de entrenamiento y validación del modelo de RNA etapa fundamental para garantizar que la arquitectura diseñada no solo aprendiera patrones relevantes, sino que también pudiera generalizarlos a nuevos datos sin perder precisión. Para lograrlo, el conjunto de información previamente depurado se dividió en tres subconjuntos independientes 70 % destinado al entrenamiento del modelo, 15 % asignado a la validación durante el proceso de ajuste de parámetros y 15 % reservado para las pruebas finales. Esta división equilibrada permitió maximizar el uso de los datos disponibles y, al mismo tiempo, obtener una evaluación objetiva de su desempeño. Durante el entrenamiento se aplicaron técnicas avanzadas de regularización orientadas a prevenir el sobreajuste, fenómeno común cuando la red neuronal memoriza los datos de entrenamiento en lugar de aprender patrones generalizables. En este sentido, se implementó el método Dropout con una tasa de $p=0.3$, lo cual implica que en cada iteración de entrenamiento el 30 % de las neuronas se desactiva de manera aleatoria, forzando a la red a no depender de combinaciones específicas y mejorando su capacidad de generalización (Figura 2).

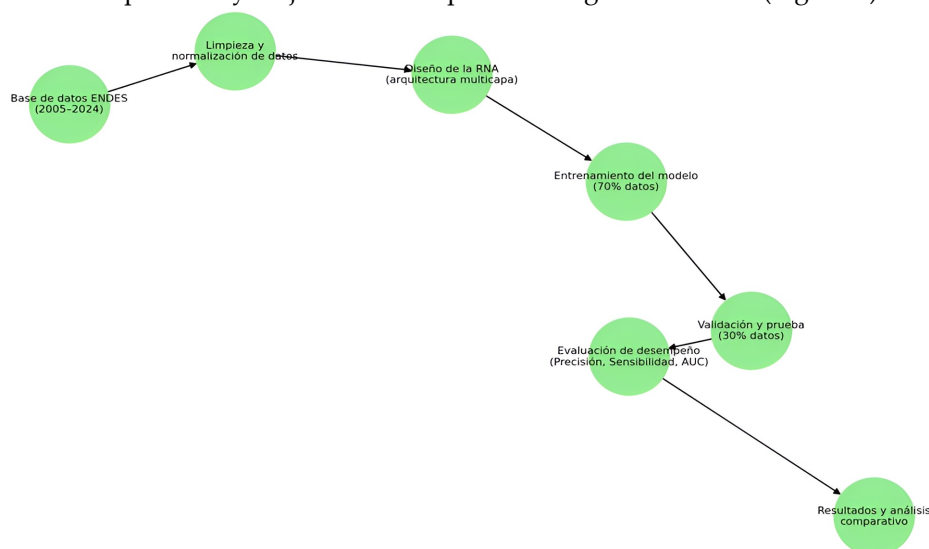


Figura 2. Recopilación de requerimientos

Asimismo, se aplicó la técnica de Early Stopping, que consiste en detener el entrenamiento automáticamente cuando el error en el conjunto de validación deja de disminuir durante varias iteraciones consecutivas, evitando que el modelo continúe ajustándose en exceso a los datos de entrenamiento [18]. Con estas estrategias se garantizó un modelo más robusto y confiable, capaz de adaptarse a la heterogeneidad propia de los registros poblacionales. La evaluación del desempeño del modelo se realizó utilizando un conjunto de métricas ampliamente aceptadas en la investigación biomédica, con el propósito de obtener una valoración integral de su capacidad predictiva y su pertinencia clínica. Se consideró la precisión, que refleja la proporción de aciertos globales, junto con la sensibilidad, que mide la habilidad del modelo para detectar correctamente a los individuos en alto riesgo de infarto, aspecto crucial para evitar omisiones de pacientes vulnerables. De manera complementaria, se evaluó la especificidad, que permite identificar a quienes no presentan riesgo, reduciendo falsas alarmas y optimizando el uso de recursos médicos. Se calculó además el F1-score, indicador que armoniza precisión y sensibilidad, lo cual resulta fundamental en contextos con clases desbalanceadas propios de la epidemiología poblacional [19]. Por último, se incluyó el área bajo la curva ROC (AUC), métrica que resume la capacidad discriminativa del modelo frente a diferentes umbrales de decisión, siendo un estándar de referencia para valorar su eficacia clínica [20]. El análisis conjunto de estas métricas permitió no solo valorar la exactitud estadística del modelo, sino también su relevancia práctica en la predicción temprana del infarto, asegurando una herramienta robusta, confiable y aplicable en programas de prevención en salud pública.

III. RESULTADOS

En esta sección se exponen los resultados del estudio. Se inicia presentando la evaluación del modelo base de referencia implementado mediante regresión logística, el cual se utilizó como línea de comparación frente a arquitecturas más avanzadas como las redes neuronales artificiales, que constituyen el eje central de la investigación. Este enfoque escalonado permite observar el rendimiento de un método estadístico clásico antes de valorar los aportes de técnicas más complejas. En el análisis se integraron los factores de riesgo definidos en la metodología dentro de un enfoque multivariado orientado a estimar la probabilidad de infarto en la población estudiada. Tal como se muestra en la Figura 3, el modelo alcanzó un accuracy general de 0.82, con una precisión de 0.84 en la clase negativa y de 0.79 en la clase positiva. La sensibilidad fue de 0.87 en los no casos y de 0.75 en los casos positivos, mientras que el F1-score registró valores de 0.86 y 0.77 respectivamente. Estos resultados evidencian un rendimiento aceptable en la clasificación global, aunque con limitaciones en la identificación precisa de individuos en alto riesgo.

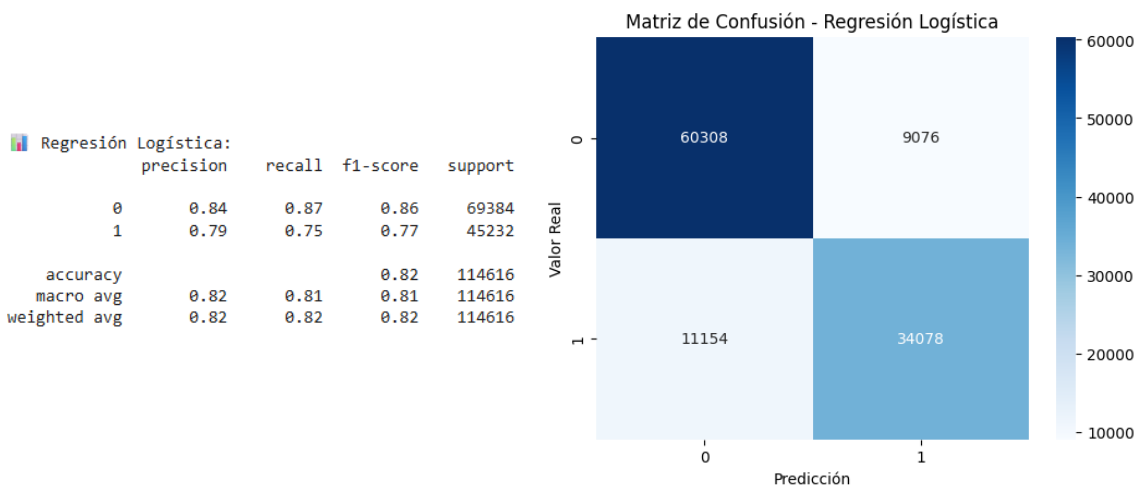


Figura 3. Regresión Logística (Modelo 1)

La matriz de confusión de la Figura 3 complementa la interpretación de los hallazgos, mostrando que se registraron 60,308 clasificaciones correctas como negativas y 34,078 como positivas, lo cual confirma un comportamiento general adecuado. Sin embargo, se observaron 9,076 falsos positivos y 11,154 falsos negativos, siendo estos últimos los más relevantes desde la perspectiva clínica, al representar individuos en riesgo que no fueron detectados. Este resultado explica la sensibilidad moderada de la clase positiva (0.75), que reduce la capacidad del modelo para anticipar tempranamente todos los casos de riesgo. Por ello, si bien la regresión logística constituye una técnica sólida y de fácil interpretación, sus limitaciones frente a interacciones no lineales refuerzan la necesidad de utilizar modelos más sofisticados, como las RNA, que serán evaluadas a continuación. De este modo, la progresión metodológica de los resultados permite establecer una comparación justa y evidenciar los beneficios adicionales de las arquitecturas neuronales.

Por contraste con el modelo de regresión logística, la Figura 4. RNA Multicapa (MLP) muestra los resultados obtenidos mediante la arquitectura MLP implementada en esta investigación. Este modelo alcanzó un accuracy global del 96 %, superando de manera notable al desempeño del modelo base. Las métricas obtenidas evidencian un comportamiento más equilibrado, con una precisión de 0.97, una sensibilidad de 0.95 y un F1-score de 0.96, lo cual indica que la red neuronal logra mantener un balance adecuado entre la detección de casos positivos y la correcta identificación de los negativos. La matriz de confusión refleja que fueron clasificados correctamente 69,251 casos como negativos y 41,204 como positivos, mientras que los errores se redujeron a 133 falsos positivos y 4,028 falsos negativos. Este último aspecto es de especial relevancia clínica, ya que la disminución de falsos negativos incrementa la confiabilidad del modelo al minimizar el riesgo de subdiagnóstico en pacientes vulnerables.

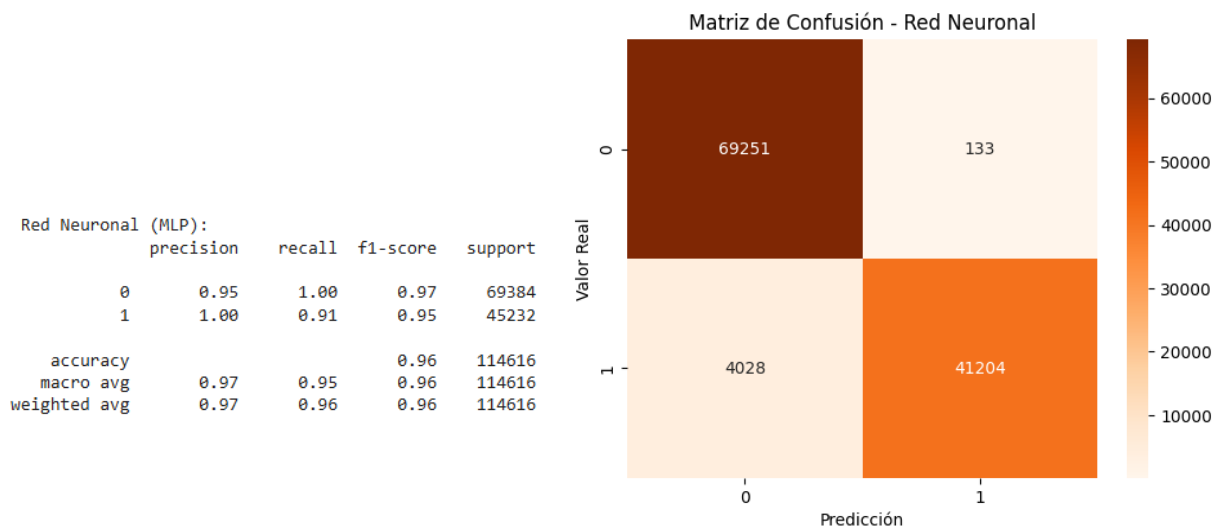


Figura 4. Red Neuronal Multicapa (Modelo 2)

Al comparar ambos enfoques, se confirma la utilidad de la regresión logística como referencia interpretativa, pero al mismo tiempo se destacan las ventajas de las redes neuronales artificiales en términos de capacidad predictiva y pertinencia clínica. Mientras que los modelos estadísticos tradicionales ofrecen claridad en la interpretación de la contribución individual de cada variable, las RNA muestran superioridad en la detección temprana, debido a su capacidad para capturar relaciones no lineales y multivariadas presentes en los datos poblacionales. En este sentido, la regresión logística permite validar estadísticamente la coherencia de los factores de riesgo, mientras que la red neuronal constituye una herramienta robusta para la identificación temprana de infartos en escenarios de salud pública, especialmente en regiones con recursos limitados. En conjunto, los resultados sugieren que la combinación de enfoques interpretativos y predictivos amplía las posibilidades de intervención temprana, ofreciendo un soporte más completo a las estrategias de prevención y gestión clínica en la población estudiada.

Los resultados comparativos entre la regresión logística y la red neuronal multicapa permiten establecer un análisis más profundo sobre el rendimiento y la pertinencia clínica de ambos modelos. Tal como se muestra en la Tabla 2, el modelo de regresión logística alcanzó un accuracy de 0.82, con valores moderados en la sensibilidad de la clase positiva (0.75) y un F1-score de 0.77 para los casos de alto riesgo. Si bien estos valores reflejan un desempeño aceptable, la presencia de más de 11,000 falsos negativos pone en evidencia una limitación crítica, ya que representa pacientes en riesgo que el modelo no logra detectar. Desde una perspectiva de salud pública, esta situación es particularmente sensible, puesto que los infartos no diagnosticados a tiempo incrementan la mortalidad y los costos asistenciales. En contraste, la red neuronal multicapa presentó un rendimiento notablemente superior, con un accuracy de 0.96 y un equilibrio mucho más consistente entre precisión, sensibilidad y F1-score. La reducción drástica de falsos negativos (4,028 casos) evidencia un avance sustancial en la confiabilidad del modelo y su utilidad clínica. Esta diferencia no solo representa una mejora estadística, sino también una ganancia significativa en términos de prevención temprana, al minimizar la posibilidad de omitir diagnósticos en individuos vulnerables.

Tabla 2. Comparación de desempeño entre Regresión Logística y Red Neuronal Multicapa

Métrica	Regresión Logística	Red Neuronal Multicapa	Observación comparativa
Accuracy	0.82	0.96	La RNA supera ampliamente la exactitud global.
Precisión clase 0	0.84	0.95	Mejora en la clasificación de no casos.
Precisión clase 1	0.79	1.00	La RNA elimina prácticamente los falsos positivos.
Sensibilidad clase 0	0.87	1.00	Incremento en la correcta detección de negativos.
Sensibilidad clase 1	0.75	0.91	Notable reducción de falsos negativos en la RNA.
F1-score clase 0	0.86	0.97	Equilibrio más alto en RNA.
F1-score clase 1	0.77	0.95	Mejora clara en la detección de positivos.
Falsos positivos	9,076	133	Reducción drástica con RNA.
Falsos negativos	11,154	4,028	Disminución significativa con RNA.

La comparación crítica entre ambos enfoques, también reflejada en la Tabla 1, permite extraer implicancias metodológicas y prácticas relevantes. La regresión logística constituye un punto de partida necesario, ya que ofrece interpretabilidad y permite validar la coherencia estadística de los factores de riesgo. Sin embargo, al trabajar con datos complejos y relaciones no lineales, como los que caracterizan a los determinantes de los infartos de miocardio, este modelo muestra limitaciones estructurales en su capacidad de predicción. Por el contrario, la red neuronal multicapa demuestra una mayor adaptabilidad a la complejidad de los datos y una mejor capacidad para identificar patrones ocultos que no son capturados por los métodos estadísticos tradicionales. Esta fortaleza se traduce en un impacto clínico directo, ya que incrementa la probabilidad de detectar de manera temprana a los pacientes en riesgo, incluso en escenarios con información heterogénea. No obstante, se debe reconocer que la mayor capacidad predictiva de la RNA implica una pérdida relativa en la interpretabilidad, aspecto que sugiere la conveniencia de integrar ambos enfoques de manera complementaria. De este modo, la regresión logística se consolida como un modelo explicativo de referencia, mientras que la RNA emerge como una herramienta predictiva robusta, capaz de potenciar la prevención y gestión temprana de los infartos en contextos de recursos limitados.

IV. DISCUSIÓN

Después de exponer los resultados de la investigación, en los que se analizaron dos modelos predictivos, la regresión logística se mantuvo como referente interpretativo, aunque mostró limitaciones relevantes para la detección temprana, con un *accuracy* de 0.82 y un total de 11,154 falsos negativos. Este patrón guarda coherencia con los hallazgos de Niedziela et al. [21] quienes observaron que las redes neuronales artificiales superaron a los modelos de regresión logística en valor predictivo al analizar pacientes con primer STEMI (infarto agudo de miocardio con elevación del segmento), alcanzando un AUROC (área bajo la curva ROC) más elevado. En el presente estudio, aunque la regresión logística resultó útil para validar la coherencia epidemiológica de los factores de riesgo, se evidenció su insuficiencia en poblaciones heterogéneas, donde el costo clínico de omitir casos positivos es alto. Este contraste refuerza la necesidad de transitar hacia arquitecturas más expresivas, como las RNA, que logran capturar interacciones no lineales y aportar mayor sensibilidad en escenarios preventivos.

De igual manera, Silva Marchan et al. [22] trabajaron con la ENDES-2021 y aplicaron algoritmos de machine learning disponibles en la plataforma Orange para predecir infartos a partir de variables sociodemográficas y antecedentes familiares, confirmando la factibilidad de construir modelos preventivos desde datos poblacionales. No obstante, su enfoque presentó limitaciones en alcance temporal y en la sofisticación de la arquitectura utilizada, lo cual restringió la estabilidad de los resultados a un único corte transversal. El presente estudio avanza sobre dicha línea en tres direcciones principales. Primero, se amplía la escala temporal al considerar el periodo 2005–2024 con más de medio millón de registros, lo que fortalece la representatividad y la detección de patrones longitudinales. Segundo, se implementa una arquitectura de RNA multicapa optimizada, entrenada con procedimientos de regularización y validación cruzada, lo que permite superar los sesgos de sobreajuste reportados en trabajos previos. Tercero, se aporta evidencia clínica relevante al mostrar una reducción sustancial de falsos negativos frente al modelo base de regresión logística, aspecto crítico para la identificación temprana de sujetos en riesgo. Si bien persisten limitaciones derivadas del sesgo de autorreporte en conductas (consumo de alcohol y tabaco), de la ausencia de biomarcadores bioquímicos y de la falta de validación externa en contextos multicéntricos, la progresión metodológica y la mejora en métricas requiere que la integración de evidencia poblacional con modelos profundos puede constituir una herramienta estratégica para fortalecer la prevención de infartos en la región norte del Perú.

Casciaro et al. [23] llevaron a cabo un estudio longitudinal durante siete años en el que aplicaron algoritmos de machine learning para identificar predictores de infarto agudo de miocardio, destacando el valor de incorporar horizontes temporales extendidos y cohortes de seguimiento para aumentar la estabilidad de los factores de riesgo en el tiempo. Sus hallazgos coinciden con la relevancia de variables clásicas como hipertensión, diabetes y hábitos nocivos, lo que refuerza su validez clínica. Sin embargo, a diferencia de dicho estudio, que se centra en una cohorte más acotada, la presente investigación amplía considerablemente la escala al integrar datos de la ENDES 2005–2024 con más de medio millón de registros heterogéneos. Esto no solo aporta mayor diversidad poblacional, sino que, además, mediante una arquitectura RNA multicapa optimizada, se alcanzan métricas superiores y un potencial de escalabilidad aplicable a estrategias de prevención en salud pública en contextos internacionales.

En esta misma línea, el desempeño de la red neuronal multicapa en esta investigación logró un *accuracy* 0.96, precisión 0.97, sensibilidad 0.95 y F1-score 0.96. Estos resultados se alinean con hallazgos donde los MLP superan al modelo de regresión logística en predicción de infarto. Choi et al. [24] desarrollaron modelos para el estadio prehospitalario sobre una base nacional y reportaron que el MLP rindió mejor que la regresión logística (AUC de referencia para el modelo clásico ≈ 0.824), atribuyendo la ventaja a la captura de interacciones y no linealidades entre predictores clínicos. En nuestro análisis, la reducción de falsos negativos a 4,028 confirma el valor

clínico de aumentar recall sin sacrificar especificidad, algo difícil de lograr con modelos lineales. Aunque el MLP sacrifica interpretabilidad, su estabilidad en métricas balanceadas respalda su uso cuando el objetivo primario es no omitir individuos en riesgo antes del evento agudo.

Aunque los resultados de esta investigación muestran un avance importante en la predicción de infartos a partir de datos poblacionales, es necesario reconocer ciertas limitaciones. En primer lugar, el uso de la ENDES como base de datos principal implica trabajar con información autodeclarada, lo que puede generar sesgos en el reporte de conductas de riesgo como tabaquismo, consumo de alcohol o prácticas alimentarias. En segundo lugar, la ausencia de biomarcadores bioquímicos y de pruebas clínicas complementarias limita la capacidad del modelo para caracterizar de manera más precisa el estado fisiopatológico de los participantes. Por último, los hallazgos no han sido validados en cohortes multicéntricas o en contextos hospitalarios, lo que restringe la generalización de los resultados y exige futuras aplicaciones en diferentes escenarios. A pesar de estas limitaciones, los resultados alcanzados confirman que las redes neuronales multicapa representan un avance sustantivo frente a los modelos de regresión logística al capturar interacciones no lineales y ofrecer mayor sensibilidad en la detección de casos positivos. La reducción de falsos negativos observada no debe interpretarse únicamente como un logro metodológico, sino como una contribución clínica de alto impacto, pues significa identificar con mayor antelación a personas en riesgo de infarto. No obstante, esta ganancia predictiva plantea el desafío de equilibrar la precisión con la interpretabilidad, condición indispensable para la adopción de los modelos en los sistemas de salud. De este modo, el porvenir de la predicción cardiovascular requiere avanzar hacia modelos que combinen la precisión de las arquitecturas profundas con la transparencia necesaria para su aceptación en la práctica clínica.

V. CONCLUSIONES

Los resultados de esta investigación confirman la superioridad de las redes neuronales artificiales frente a los modelos estadísticos tradicionales en la predicción temprana de infartos de miocardio. Mientras que la regresión logística alcanzó un desempeño aceptable con un 82 % de exactitud, mostró limitaciones significativas al momento de identificar casos positivos, reflejadas en una alta proporción de falsos negativos. En contraste, el modelo neuronal multicapa logró una exactitud del 96 %, acompañado de métricas más equilibradas y un número considerablemente menor de errores de clasificación. Esta diferencia revela la capacidad de las redes neuronales para reconocer interacciones no lineales y multivariadas, lo que permite delinear de manera más precisa el perfil de riesgo cardiovascular. Estos hallazgos no deben interpretarse como una sustitución absoluta de la regresión logística, sino como una relación complementaria entre enfoques. La regresión logística conserva un valor indiscutible gracias a su interpretabilidad, pues facilita estimar el peso relativo de cada variable y sustenta la validación epidemiológica en el análisis de factores de riesgo. Sin embargo, la potencia predictiva alcanzada por las redes neuronales introduce una necesidad inaplazable de incorporar inteligencia artificial en la práctica médica. Este paso resulta especialmente relevante en el norte del Perú, donde las deficiencias en infraestructura hospitalaria, los diagnósticos tardíos y la escasez de especialistas limitan las oportunidades de intervención temprana.

En términos académicos, este estudio aporta evidencia sólida sobre la viabilidad de emplear bases de datos poblacionales extensas, como la Encuesta Demográfica y de Salud Familiar, en la construcción de modelos predictivos de alto rendimiento. Este enfoque se alinea con investigaciones internacionales que recomiendan entrenar algoritmos con información contextualizada a fin de incrementar su aplicabilidad en entornos específicos. De esta manera, la investigación contribuye al fortalecimiento de la agenda latinoamericana en inteligencia artificial aplicada a la salud, ofreciendo un modelo robusto, clínicamente pertinente y sensible a la realidad social peruana. En suma, los resultados refuerzan la relevancia de vincular ciencia de datos, salud

pública y prevención de enfermedades cardiovasculares. Como proyección futura, se plantea la necesidad de perfeccionar el modelo mediante la integración de variables adicionales relacionadas con genética, hábitos de vida y factores ambientales que influyen en la aparición de infartos. Asimismo, resulta pertinente explorar la implementación de sistemas predictivos en entornos hospitalarios y comunitarios en tiempo real, lo que permitiría generar alertas personalizadas para la población en riesgo. Un desafío complementario será garantizar la interpretabilidad de los modelos complejos, con el fin de facilitar su adopción por profesionales de la salud y responsables de políticas públicas. Estas líneas abren un horizonte de investigación que combina innovación tecnológica, pertinencia social y mejora de la calidad de vida.

Contribución de los autores:

La conceptualización del estudio fue realizada por H.S.M., G.B.R. y R.S.A.; el diseño metodológico y el análisis formal de los datos, por H.S.M., G.B.R. y R.S.A.; la recolección y curación de datos, por H.S.M. y J.P.G.; la redacción del borrador original estuvo a cargo de H.S.M. y R.S.A., quienes además elaboraron la visualización de los resultados. La revisión crítica y edición del manuscrito fue realizada por H.S.M. y J.P.G.; la supervisión general, así como la administración del proyecto, estuvieron a cargo de H.S.M. y G.B.R. Todos los autores han leído y aprobado la versión final del manuscrito, y se hacen responsables de su contenido.

Agradecimientos:

Se expresa un profundo reconocimiento a la Universidad Nacional de Tumbes por el respaldo institucional brindado, que hizo posible el acceso a sus plataformas académicas y bibliotecas especializadas, insumos indispensables para la consolidación de este trabajo de investigación. Asimismo, se agradece a International Journal of Computational Innovations, Intelligent Systems and AI por constituir un espacio de difusión académica de alto nivel, así como a los revisores anónimos por el valioso soporte brindado en la mejora del manuscrito. Sus observaciones y sugerencias contribuyeron de manera significativa a perfeccionar el rigor metodológico, la claridad argumentativa y la proyección científica de los resultados aquí expuestos.

Financiamiento:

El estudio ha sido autofinanciado por los autores.

Declaración de consentimiento informado:

No aplica.

Declaración de disponibilidad de datos:

Los datos que respaldan los resultados de este estudio están disponibles previa solicitud al autor de correspondencia.

Conflictos de intereses:

No existen conflictos de intereses.

Declaración de uso de IA:

Se utilizó ChatGPT como apoyo para mejorar la redacción, gramática y fluidez del manuscrito. El contenido intelectual, la interpretación de los datos y las conclusiones son responsabilidad exclusiva de los autores.

REFERENCIAS

- [1] O. S. Medina-Gómez, "Impacto de la pandemia de COVID-19 en las tendencias de mortalidad por enfermedades cardiovasculares en México, 2000-2022," *Medicina de Familia. SEMERGEN*, vol. 50, no. 3, p. 102170, Apr. 2024. <https://doi.org/10.1016/j.semerg.2023.102170>
- [2] H. G. Quezada-Pinedo et al., "A comprehensive analysis of cardiovascular mortality trends in Peru from 2017 to 2022: Insights from 183,386 deaths of the national death registry," *American Heart Journal Plus: Cardiology Research and Practice*, vol. 35, p. 100335, Nov. 2023. <https://doi.org/10.1016/j.ahjo.2023.100335>
- [3] J. A. Vázquez-Troche, V. García-Fernández, A. Hernández-Vásquez, R. Vargas-Fernández, and G. Bendezu-Quispe, "Trends in Mortality from Ischemic Heart Disease in Peru, 2005 to 2017," *IJERPH*, vol. 19, no. 12, p. 7047, Jun. 2022. <https://doi.org/10.3390/ijerph19127047>
- [4] M. A. Muzammil et al., "Artificial intelligence-enhanced electrocardiography for accurate diagnosis and management of cardiovascular diseases," *Journal of Electrocardiology*, vol. 83, pp. 30–40, Mar. 2024. <https://doi.org/10.1016/j.jelectrocard.2024.01.006>
- [5] Gian Luigi Nicolosi, "Perché l'intelligenza artificiale applicata all'elettrocardiogramma non è ancora routine clinica?," *Recenti Progressi in Medicina*, no. 2023Marzo, Mar. 2023. <https://doi.org/10.1701/3981.39635>
- [6] I. Rojek, P. Kotlarz, M. Kozielski, M. Jagodziński, and Z. Królikowski, "Development of AI-Based Prediction of Heart Attack Risk as an Element of Preventive Medicine," *Electronics*, vol. 13, no. 2, p. 272, Jan. 2024. <https://doi.org/10.3390/electronics13020272>
- [7] A. Irsyad, P. P. Widagdo, P. P. Widagdo, and R. Wardhana, "Deep learning Methods for ECG-Based Heart Disease Detection," *j.electron.electromedical.eng.med.inform*, vol. 6, no. 4, pp. 467–477, Sep. 2024. <https://doi.org/10.35882/jeeemi.v6i4.498>
- [8] M. Mandava and S. Reddy Vinta, "MDensNet201-IDRSRNet: Efficient cardiovascular disease prediction system using hybrid deep learning," *Biomedical Signal Processing and Control*, vol. 93, p. 106147, Jul. 2024. <https://doi.org/10.1016/j.bspc.2024.106147>
- [9] J. Moya-Salazar et al., "Geospatial analysis of cardiovascular mortality before and during the COVID-19 pandemic in Peru: analysis of the national death registry to support emergency management in Peru," *Front. Cardiovasc. Med.*, vol. 11, p. 1316192, Jun. 2024. <https://doi.org/10.3389/fcvm.2024.1316192>
- [10] L. M. Davila-Zamora et al., "Artificial Intelligence Adoption in Emerging Economies: Challenges, Opportunities, And Strategies for Peru's Business Transformation," *Pak.j.life.soc.sci.*, vol. 22, no. 2, 2024. <https://doi.org/10.57239/PJLSS-2024-22.2.000864>
- [11] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of deep convolutional neural network for automated detection of myocardial infarction using ECG signals," *Information Sciences*, vol. 415–416, pp. 190–198, Nov. 2017. <https://doi.org/10.1016/j.ins.2017.06.027>
- [12] S. Gustafsson et al., "Development and validation of deep learning ECG-based prediction of myocardial infarction in emergency department patients," *Sci Rep*, vol. 12, no. 1, p. 19615, Nov. 2022. <https://doi.org/10.1038/s41598-022-24254-x>
- [13] D. Doudesis et al., "Machine learning for diagnosis of myocardial infarction using cardiac troponin concentrations," *Nat Med*, vol. 29, no. 5, pp. 1201–1210, May 2023. <https://doi.org/10.1038/s41591-023-02325-4>
- [14] B. Toprak et al., "Diagnostic accuracy of a machine learning algorithm using point-of-care high-sensitivity cardiac troponin I for rapid rule-out of myocardial infarction: a retrospective study," *The Lancet Digital Health*, vol. 6, no. 10, pp. e729–e738, Oct. 2024. [https://doi.org/10.1016/S2589-7500\(24\)00191-2](https://doi.org/10.1016/S2589-7500(24)00191-2)
- [15] R. Khera et al., "Use of Machine Learning Models to Predict Death After Acute Myocardial Infarction," *JAMA Cardiol*, vol. 6, no. 6, p. 633, Jun. 2021. <https://doi.org/10.1001/jamacardio.2021.0122>
- [16] H. Yu, Y. Qi, and Y. Ding, "Deep Learning in RNA Structure Studies," *Front. Mol. Biosci.*, vol. 9, p. 869601, May 2022. <https://doi.org/10.3389/fmolb.2022.869601>
- [17] Z. Hu, Y. Li, and Z. Yang, "Improving Convolutional Neural Network Using Pseudo Derivative ReLU," in *2018 5th International Conference on Systems and Informatics (ICSAI)*, Nanjing: IEEE, Nov. 2018, pp. 283–287. <https://doi.org/10.1109/ICSAI.2018.8599372>
- [18] I. Lauriola, "On the Impact of Early Stopping in Multiple Kernel Learning," in *Proceedings of the Future Technologies Conference (FTC) 2020, Volume 1*, vol. 1288, K. Arai, S. Kapoor, and R. Bhatia, Eds., in

- Advances in Intelligent Systems and Computing, vol. 1288. , Cham: Springer International Publishing, 2021, pp. 205–215. https://doi.org/10.1007/978-3-030-63128-4_16
- [19] D. J. Hand, P. Christen, and N. Kirielle, “F*: an interpretable transformation of the F-measure,” *Mach Learn*, vol. 110, no. 3, pp. 451–456, Mar. 2021. <https://doi.org/10.1007/s10994-021-05964-1>
- [20] J. A. Martínez Pérez and P. S. Pérez Martín, “La curva ROC,” *Medicina de Familia. SEMERGEN*, vol. 49, no. 1, p. 101821, Jan. 2023. <https://doi.org/10.1016/j.semerg.2022.101821>
- [21] J. T. Niedziela et al., “Is neural network better than logistic regression in death prediction in patients after ST-segment elevation myocardial infarction?,” *Kardiol Pol*, vol. 79, no. 12, pp. 1353–1361, Dec. 2021. <https://doi.org/10.33963/KP.a2021.0142>
- [22] H. Silva Marchan, G. Ortiz Castro, O. J. M. Peña Cáceres, and M. A. More More, “Logistic regression: an example for the prediction of heart attacks,” in *Proceedings of the 21th LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI 2023): “Leadership in Education and Innovation in Engineering in the Framework of Global Transformations: Integration and Alliances for Integral Development,” Latin American and Caribbean Consortium of Engineering Institutions, 2023.* <https://doi.org/10.18687/LACCEI2023.1.1.720>
- [23] M. Casciaro et al., “Predictors of Acute Myocardial Infarction: A Machine Learning Analysis After a 7-Year Follow-Up,” *Clinics and Practice*, vol. 15, no. 4, p. 72, Mar. 2025. <https://doi.org/10.3390/clinpract15040072>
- [24] A. Choi et al., “Development of Prediction Models for Acute Myocardial Infarction at Prehospital Stage with Machine Learning Based on a Nationwide Database,” *JCDD*, vol. 9, no. 12, p. 430, Dec. 2022. <https://doi.org/10.3390/jcdd9120430>