

Implementación de redes neuronales para la predicción del síndrome metabólico: un estudio con datos multinacionales

Alberto Guevara-Tirado 

Facultad de Medicina Humana, Universidad Científica del Sur, Lima, Perú

Resumen

Antecedentes: El síndrome metabólico (SM) está constituido por anomalías como la obesidad central, la resistencia a la insulina, la hipertensión y la dislipidemia. **Objetivo:** Implementar una red neuronal para predecir el SM a partir del colesterol, los triglicéridos, el colesterol unido a lipoproteínas de alta densidad (HDL), la obesidad y la hipertensión. **Material y métodos:** Estudio analítico y transversal con 1878 pacientes de bases de datos de Venezuela, Tailandia e Indonesia. Se incluyeron las variables SM, hipertensión, obesidad, HDL, triglicéridos y colesterol total. Se usaron redes neuronales tipo perceptrón multicapa, evaluadas con tablas de clasificación, área bajo la curva (AUC) y métricas de desempeño (sensibilidad, especificidad y valores predictivos positivo y negativo). **Resultados:** La red neuronal mostró una alta capacidad para predecir el SM, con un bajo porcentaje de pronósticos incorrectos tanto en el conjunto de entrenamiento (15.80%) como en el de prueba (18.20%). En el entrenamiento, la precisión global fue del 84.20%, con mayor precisión para casos sin SM (88.30%) que para casos con SM (77.80%). En las pruebas, la precisión global fue del 81.80%, también con mayor precisión para casos sin SM (86.60%) que con SM (74.80%). El AUC fue 0.911, indicando una capacidad predictiva sobresaliente. Respecto al desempeño del modelo, la sensibilidad fue del 81.25% en el entrenamiento y del 79.26% en la prueba, mientras que la especificidad alcanzó el 85.92% y el 83.33%, respectivamente. El valor predictivo positivo fue del 77.80% en entrenamiento y del 74.78% en prueba, y el valor predictivo negativo del 88.30% y el 86.57%, respectivamente. **Conclusiones:** La red neuronal tipo perceptrón multicapa es una herramienta eficaz para predecir el SM, mostrando una capacidad predictiva sobresaliente.

PALABRAS CLAVE: Síndrome metabólico. Registros médicos. Lípidos. Redes neurales de la computación. Toma de decisiones asistida por computador.

Implementation of neural networks for predicting metabolic syndrome: a multinational data study

Abstract

Background: Metabolic syndrome (MS) is composed of abnormalities such as central obesity, insulin resistance, hypertension and dyslipidemia. **Objective:** To implement a neural network to predict MS from cholesterol, triglycerides, high density lipoproteins (HDL), obesity and hypertension. **Material and methods:** Analytical and cross-sectional study with 1878 patients from databases in Venezuela, Thailand and Indonesia. Variables such as MS, hypertension, obesity, HDL, triglycerides and total cholesterol were included. Multilayer perceptron neural networks were used, evaluated with classification tables, area under the curve (AUC) and performance metrics (sensitivity, specificity, positive and negative predictive values). **Results:** The neural network showed a high capacity to predict MS, with a low percentage of incorrect predictions both in the training set (15.80%) and in the test set (18.20%). In training, the overall accuracy was 84.20%, with higher accuracy for cases without MS (88.30%)

Correspondencia:

Alberto Guevara-Tirado

E-mail: albertoguevara1986@gmail.com

0016-3813/© 2025 Academia Nacional de Medicina de México, A.C. Publicado por Permanyer. Este es un artículo open access bajo la licencia CC BY-NC-ND (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

Fecha de recepción: 29-11-2024

Fecha de aceptación: 26-02-2025

DOI: 10.24875/GMM.24000411

Gac Med Mex. 2025;161:334-341

Disponible en PubMed

www.gacetamedicademexico.com

than for cases with MS (77.80%). In testing, the overall accuracy was 81.80%, also with higher accuracy for cases without MS (86.60%) than for cases with MS (74.80%). The AUC was 0.911, indicating an outstanding predictive capacity. Regarding the model performance, the sensitivity was 81.25% in training and 79.26% in testing, while the specificity reached 85.92% and 83.33%, respectively. The positive predictive value was 77.80% in training and 74.78% in testing, and the negative predictive value was 88.30% and 86.57%, respectively. **Conclusions:** The multilayer perceptron neural network is an effective tool to predict MS, showing an outstanding predictive capacity.

KEYWORDS: Metabolic syndrome. Medical records. Lipids. Neural networks computer. Decision making by computer-assisted.

Introducción

El síndrome metabólico (SM) es un conjunto de alteraciones metabólicas caracterizadas por obesidad central, resistencia a la insulina, hipertensión y dislipidemia.¹ Su origen involucra una combinación de factores genéticos, ambientales y de estilo de vida, como el sedentarismo y una alimentación poco saludable.² La acumulación de tejido adiposo desempeña un papel fundamental en su desarrollo, al promover resistencia a la insulina y la liberación de citocinas proinflamatorias, como factor de necrosis tumoral alfa, leptina, adiponectina e inhibidor del activador del plasminógeno.³ Su alta prevalencia, que oscila entre el 12.50% y el 31.40% según los criterios diagnósticos aplicados, lo convierte en un desafío global para la salud pública.^{4,5}

El impacto económico del SM es significativo, ya que incrementa los costos asociados con el tratamiento de la diabetes tipo 2, las enfermedades cardiovasculares y los accidentes cerebrovasculares.⁶ Los sistemas de salud enfrentan una creciente presión debido a la detección tardía y la falta de estrategias preventivas eficaces, lo que genera un aumento en los gastos sanitarios derivados de hospitalizaciones y tratamientos a largo plazo. A esto se suma la falta de consenso en su diagnóstico, debido a la coexistencia de múltiples criterios establecidos por organismos como la Organización Mundial de la Salud, el Panel de Tratamiento de Adultos III y la Federación Internacional de Diabetes.⁷ Además, la variabilidad en los parámetros de obesidad abdominal entre poblaciones dificulta la estandarización del diagnóstico, limitando el acceso temprano a intervenciones preventivas.

En este contexto, la inteligencia artificial ha surgido como una herramienta clave en la medicina preventiva, permitiendo desarrollar modelos predictivos que mejoran la detección temprana del SM mediante el análisis de múltiples factores de riesgo. Entre estos enfoques, las redes neuronales artificiales,

específicamente el perceptrón multicapa (MLP, *multi-layer perceptron*), han demostrado ser altamente efectivas en la identificación de patrones complejos dentro de grandes volúmenes de datos clínicos. Su aplicación podría reducir la dependencia de pruebas de laboratorio costosas y facilitar un diagnóstico más accesible y preciso.

Por ello, el objetivo de este estudio fue implementar un modelo basado en redes neuronales artificiales para predecir el SM, utilizando variables clínicas de bajo costo, como el colesterol total (CT), los triglicéridos (TG), el colesterol unido a lipoproteínas de alta densidad (HDL), la obesidad y la hipertensión. Con ello, se busca contribuir a la detección temprana del SM y la optimización de estrategias preventivas, reduciendo la carga económica y clínica del SM.

Material y métodos

Diseño y población

Estudio analítico y transversal de 1878 pacientes pertenecientes a tres bases de datos secundarias de Venezuela, Tailandia e Indonesia. Los participantes tuvieron varias condiciones heterogéneas y comorbilidad, incluyendo enfermedades cardiovasculares y diabetes tipo 2, ya que, aunque el SM y la diabetes están relacionados y comparten varios factores de riesgo, no siempre los diabéticos tienen SM, pues la diabetes se centra principalmente en la regulación glucémica, mientras que el SM incluye una combinación de factores como obesidad abdominal, hipertensión, dislipidemia y resistencia a la insulina.⁸

Es fundamental aclarar cómo se estandarizaron los criterios diagnósticos entre las distintas bases de datos. Las tres bases de datos comparten la característica de contener diagnósticos de SM, aunque provienen de estudios con objetivos y enfoques diferentes. La base de datos de Venezuela se originó a partir del Estudio de Prevalencia del Síndrome Metabólico en la Ciudad de Maracaibo. La base de datos de Tailandia

fue recolectada en un estudio que investigó los patrones alimentarios y su asociación con los componentes del SM en adultos del noreste del país. Por su parte, la base de datos de Indonesia proviene de un estudio centrado en evaluar el desempeño de los perfiles y las ratios lipídicas como predictores de rigidez arterial, determinada mediante la velocidad de la onda de pulso braquial-tobillo en pacientes con diabetes tipo 2. En este análisis se incluyeron todos los pacientes registrados en las bases de datos secundarias que disponían de las variables necesarias para la investigación, sin realizar muestreo alguno. Las características de cada base de datos se detallan en la tabla 1. Debido a las diferencias inherentes entre las bases, algunos datos clínicos no están disponibles de manera homogénea.

Variables y mediciones

Las variables seleccionadas para entrenar la red neuronal fueron aquellas presentes en las tres bases de datos, garantizando su comparabilidad y maximizando la cantidad de datos disponibles para el modelo. La variable dependiente fue la presencia de SM, dicotomizada en sí o no, conforme a los criterios diagnósticos específicos de cada base de datos de origen.

En la base de datos de Indonesia, el diagnóstico de SM se estableció por la presencia de hipertensión arterial, alteración del metabolismo de la glucosa y dislipidemia aterogénica (niveles elevados de colesterol no HDL).⁹ En la base de datos de Tailandia se aplicaron los criterios armonizados establecidos por seis organizaciones internacionales, los cuales incluyen circunferencia de la cintura ≥ 90 cm en hombres o ≥ 80 cm en mujeres, glucemia basal ≥ 100 mg/dl o diagnóstico previo de diabetes tratada, TG ≥ 150 mg/dl, HDL < 40 mg/dl en hombres y < 50 mg/dl en mujeres, y presión arterial sistólica ≥ 130 mmHg o diastólica ≥ 85 mmHg, o hipertensión tratada.¹⁰ En la base de datos de Venezuela, el diagnóstico de SM se basó en los criterios del consenso de armonización del síndrome metabólico del año 2009.¹¹

Dado que los criterios diagnósticos de SM diferían entre las bases de datos, se adoptó un enfoque basado en la estandarización de las variables predictoras. Se seleccionaron como predictoras aquellas variables que estaban presentes de manera homogénea en las tres fuentes de datos y que, además, han sido ampliamente documentadas en la literatura como

Tabla 1. Características de las poblaciones de las bases de datos de Venezuela, Indonesia y Tailandia utilizadas para la investigación

Base de datos de Venezuela (n = 1226): SM = 483, HTA = 201, obesidad = 624				
	Mínimo	Máximo	Media	DE
IMC (kg/m ²)	14.22	68.80	28.70	7.69
Insulina (U/ml)	1.20	100.20	15.09	10.28
Glucosa basal (mg/dl)	62.50	350	94.70	20.31
CT (mg/dl)	89	677	187.64	45.53
TG (mg/dl)	21	1039	123.68	95.32
HDL (mg/dl)	20	114	44.43	11.94
VLDL (mg/dl)	4.20	207.80	24.64	18.80
LDL (mg/dl)	2.40	273.88	118.09	36.93
PAS (mmHg)	90	200	118.97	17.13
PAD (mmHg)	50	140	76.98	11.71
Base de datos de Indonesia (n = 184): SM = 82, HTA = 104, obesidad = 51				
	Mínimo	Máximo	Media	DE
Edad (años)	18	77	57.16	9.71
HbA1c (%)	5.20	12.50	7.97	1.46
CT (mg/dl)	100	350	201.96	44.01
TG (mg/dl)	38	412	158.77	66.99
HDL (mg/dl)	28	157	49.37	16.02
LDL (mg/dl)	43	275	134.05	39.25
Base de datos de Tailandia (n = 468): SM = 183, HTA = 199, obesidad = 48				
	Mínimo	Máximo	Media	DE
Edad (años)	32	65	49.13	6.36
IMC (kg/m ²)	15.30	42.60	25.04	3.91
PAS (mmHg)	89.50	189.50	127.91	17.41
PAD (mmHg)	56	111	75.98	9.81
Glucosa basal (mg/dl)	58.20	298.60	90.43	25.40
CT (mg/dl)	72.30	391	197.82	50.38
TG (mg/dl)	43.70	617.10	155.10	90.40
HDL (mg/dl)	15	88.70	43.85	12.39

CT: colesterol total; DE: desviación estándar; HbA1c: hemoglobina glucosilada; HDL: lipoproteínas de alta densidad; HTA: hipertensión arterial; IMC: índice de masa corporal; LDL: lipoproteínas de baja densidad; PAD: presión arterial diastólica; PAS: presión arterial sistólica; SM: síndrome metabólico; TG: triglicéridos; VLDL: lipoproteínas de muy baja densidad.

factores determinantes del SM. Estas incluyeron la presencia de hipertensión arterial y obesidad (ambas tratadas como variables dicotómicas), así como los

niveles de colesterol HDL, TG y CT (como variables continuas).

La inclusión de estas variables no solo garantizó la comparabilidad entre las bases de datos, sino que también permitió que el modelo capturara los principales componentes fisiopatológicos del SM, independientemente de las diferencias en los criterios diagnósticos utilizados en cada país. Los datos clínicos fueron extraídos de las historias clínicas correspondientes, y luego las bases de datos se integraron en un único archivo en SPSS Statistics 25™, asegurando la compatibilidad y la consistencia de las variables en común junto con los diagnósticos de SM según los criterios específicos de cada base de datos.

Análisis estadístico

Se utilizaron tablas con frecuencias absolutas y relativas para describir las características de la muestra. Para la predicción del SM se empleó una red neuronal de tipo MLP, utilizando como variables predictoras los niveles de HDL, TG y CT, y la presencia de hipertensión arterial y obesidad, de acuerdo con los criterios adoptados por cada base de datos.

Se utilizó el MLP, un modelo de red neuronal artificial utilizado en aprendizaje automático que está compuesto por una capa de entrada, una o más capas ocultas, y una capa de salida. A través del algoritmo de retropropagación ajusta los pesos de las conexiones para minimizar el error entre las predicciones y los valores reales, permitiendo identificar patrones complejos en los datos.¹²

La elección del MLP sobre otros modelos más complejos, como las redes neuronales convolucionales o los modelos de aprendizaje profundo más sofisticados, se debió a varios factores. En primer lugar, el MLP es suficiente para capturar relaciones no lineales entre las variables clínicas sin requerir una cantidad excesiva de datos para el entrenamiento, lo que le hace adecuado para conjuntos de datos clínicos con tamaños moderados. Además, su arquitectura relativamente simple facilita la interpretabilidad de los resultados y reduce el riesgo de sobreajuste en comparación con modelos más complejos, que requieren una mayor capacidad computacional y volúmenes de datos más extensos para su correcto entrenamiento.

Este método se ejecutó en el *software* SPSS Statistics 25™ mediante la ruta Analizar → Redes neuronales → Perceptrón multicapa. Se estableció la variable SM como dependiente y las variables HDL, TG, CT, hipertensión y obesidad como covariables.

Glosario de términos

- Perceptrón multicapa: modelo de red neuronal artificial que procesa información a través de múltiples capas para identificar patrones en los datos.
- Retropropagación: algoritmo que ajusta los pesos de una red neuronal para mejorar su precisión en la predicción.
- Funciones de activación: transformaciones matemáticas que permiten a la red neuronal modelar relaciones no lineales en los datos. Ejemplos: tangente hiperbólica y softmax.
- Covariables: variables utilizadas como predictores en un modelo estadístico o de aprendizaje automático.

Disponibilidad de bases de datos

Las bases de datos utilizadas se encuentran en el repositorio F1000Research, de libre acceso. F1000Research requiere que todos los datos asociados con los artículos se compartan con una licencia abierta, como CC0 o CC-BY, que permite su reutilización con mínimas restricciones, con el propósito de facilitar la transparencia y la reproducibilidad en la investigación.¹³ Por tanto, al ser de acceso abierto, no se precisaron autorizaciones. Las bases de datos, junto a sus descripciones y metadatos, se encuentran en los siguientes enlaces:

- Base de datos de Venezuela: <https://f1000research.com/articles/7-230#ref-18>.
- Base de datos de Indonesia: <https://f1000research.com/articles/11-1582#ref26>.
- Base de datos de Tailandia: <https://f1000research.com/articles/7-905>.

Resultados

Se presentó un resumen del modelo de redes neuronales tipo MLP utilizado para la predicción del SM a partir de las variables CT, TG, HDL, presencia de obesidad e hipertensión, a través de un modelo de retropropagación hacia atrás. El modelo fue entrenado con bases de datos secundarias provenientes de tres países. Durante la fase de entrenamiento (n = 1313), el modelo obtuvo un error de entropía cruzada de 473.954, con un 15.80% de pronósticos incorrectos. La regla de parada utilizada fue una iteración sin reducción del error, y el tiempo total de

entrenamiento fue de 0.12 segundos. En la fase de prueba ($n = 565$), el error de entropía cruzada fue de 214.619 y el porcentaje de pronósticos incorrectos ascendió al 18.20% (Tabla 2).

El modelo de red neuronal tipo MLP utilizado para la predicción del SM se compone de una estructura con una capa de entrada, una capa oculta y una capa de salida. La capa de entrada incluye cinco covariables (CT, TG, HDL, presencia de hipertensión y obesidad), las cuales fueron escaladas mediante estandarización. La capa oculta está compuesta por seis neuronas y su función de activación es la tangente hiperbólica, permitiendo representar relaciones no lineales entre las variables de entrada y la salida del modelo. Finalmente, la capa de salida tiene como variable dependiente el SM, con dos unidades de salida que representan la clasificación del síndrome. Para esta capa se utilizó la función de activación softmax, que asigna probabilidades a cada categoría, y la función de error de entropía cruzada, la cual es adecuada para problemas de clasificación (Figura 1).

El área bajo la curva fue de 0.911, lo que indicó que el modelo predictivo tuvo una capacidad predictiva sobresaliente (Figura 2).

La precisión global del modelo en el conjunto de entrenamiento fue del 84.20%, con una mayor precisión en la predicción de los casos sin SM (88.30%) en comparación con los casos con SM (77.80%). En el conjunto de pruebas, la precisión global fue del 81.80%, manteniendo una tendencia similar con una mayor precisión para los casos sin SM (86.60%) en comparación con los casos con SM (74.80%) (Tabla 3).

Se evaluó el desempeño del modelo de redes neuronales en la predicción del síndrome metabólico. La sensibilidad (81.25% en entrenamiento y 79.26% en prueba) y la especificidad (85.92% y 83.33%, respectivamente) indican una buena capacidad del modelo para identificar correctamente tanto casos positivos como negativos. El valor predictivo positivo (77.80% en entrenamiento y 74.78% en prueba) sugiere una precisión aceptable en la clasificación de casos positivos, mientras que el valor predictivo negativo (88.30% y 86.57%) refleja una alta confiabilidad en la identificación de negativos. En general, el modelo muestra un buen equilibrio entre ambos conjuntos, aunque presenta una leve disminución en el rendimiento en la prueba, lo que podría mejorarse con ajustes en el entrenamiento o mayor cantidad de datos (Tabla 4).

Tabla 2. Resumen del modelo de redes neuronales tipo perceptrón multicapa para predicción de síndrome metabólico a partir del colesterol total, los triglicéridos, el colesterol unido a lipoproteínas de alta densidad y la presencia de obesidad e hipertensión, a partir de bases de datos secundarios de tres países

Entrenamiento ($n = 1313$)	Error de entropía cruzada	473.954
	Porcentaje de pronósticos incorrectos	15.80%
	Regla de parada utilizada	Un paso consecutivo sin disminución del error
	Tiempo de entrenamiento	0:00:00.12
Prueba ($n = 565$)	Error de entropía cruzada	214.619
	Porcentaje de pronósticos incorrectos	18.20%

Tabla 3. Clasificación de los pronósticos correctos de síndrome metabólico predichos mediante redes neuronales tipo perceptrón multicapa

SM a partir de HDL, CT, TG, obesidad e hipertensión	Presencia de SM		Porcentaje correcto
	Sí	No	
Entrenamiento			
Sí	403	115	77.80%
No	93	702	88.30%
Porcentaje global	37.80%	62.20%	84.20%
Pruebas			
Sí	172	58	74.80%
No	45	290	86.60%
Porcentaje global	38.40%	61.60%	81.80%

CT: colesterol total; HDL: lipoproteínas de alta densidad; SM: síndrome metabólico; TG: triglicéridos.

Discusión

Los hallazgos obtenidos resaltan el potencial del modelo de red neuronal tipo perceptrón multicapa (MLP) en la predicción del síndrome metabólico, evidenciando una alta capacidad discriminativa. Esto sugiere un desempeño prometedor del modelo, aunque también pone de manifiesto la necesidad de optimizar su capacidad para identificar correctamente los casos positivos. En este contexto, es crucial comparar estos resultados con estudios previos y analizar sus implicancias en la práctica clínica y la investigación epidemiológica. A continuación, se presentan los principales hallazgos en relación con la literatura

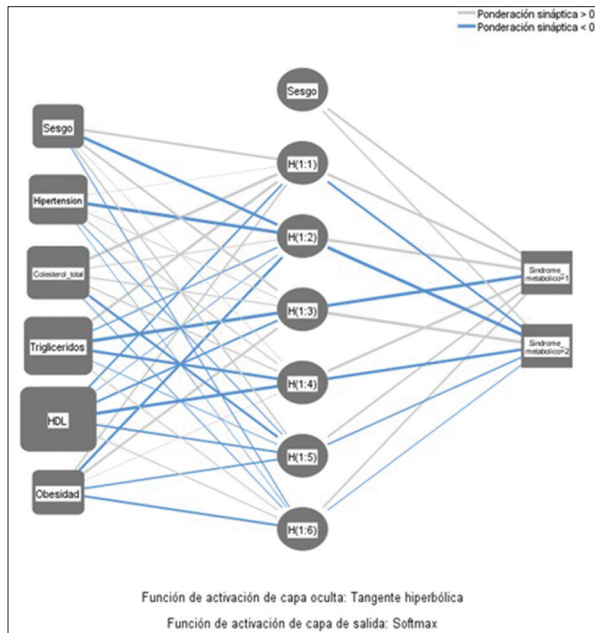


Figura 1. Estructura de red neuronal para predicción de síndrome metabólico a partir del colesterol, los triglicéridos, las lipoproteínas de alta densidad (HDL) y la presencia de obesidad e hipertensión.

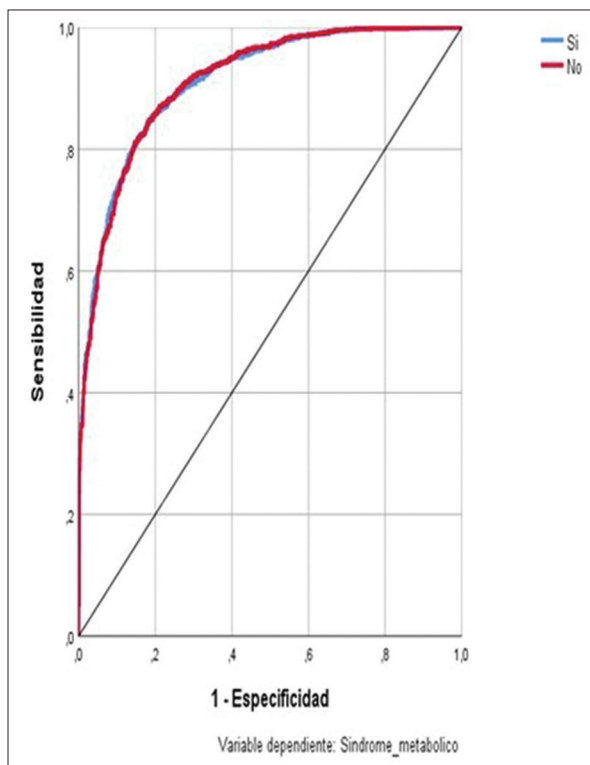


Figura 2. Área bajo la curva del modelo de predicción de síndrome metabólico a partir del colesterol, los triglicéridos, las lipoproteínas de alta densidad y la presencia de obesidad e hipertensión.

existente y se examinan las posibles aplicaciones del modelo en la detección temprana del síndrome metabólico.

Tabla 4. Desempeño del modelo de redes neuronales en la predicción del síndrome metabólico

Métrica	Entrenamiento	Prueba
Sensibilidad	81.25%	79.26%
Especificidad	85.92%	83.33%
Valor predictivo positivo	77.80%	74.78%
Valor predictivo negativo	88.30%	86.57%

El diagnóstico del síndrome metabólico se ha basado tradicionalmente en la identificación de múltiples factores de riesgo, entre los que se incluyen los evaluados en esta investigación, como niveles elevados de TG y CT, bajos niveles de colesterol HDL, obesidad e hipertensión. No obstante, el uso de modelos predictivos avanzados, como las redes neuronales tipo perceptrón multicapa, permite integrar estos factores y mejorar la detección temprana y precisa de la condición.

Diversos estudios han explorado el uso de redes neuronales en la predicción del síndrome metabólico, con diferencias en el tamaño de las muestras, enfoques metodológicos y tipos de redes utilizadas. Hirose et al.¹⁴ emplearon redes neuronales y regresión logística binaria en 400 adultos japoneses, incorporando variables como IMC, edad, presión arterial diastólica, HDL, LDL, HOMA-IR y adiponectina sérica. Por su parte, Eyvazlou et al.¹⁵ aplicaron redes neuronales en 468 trabajadores de una refinería en Irán, considerando predictores como sexo biológico, edad, ejercicio, tabaquismo, riesgo de apnea del sueño y factores de estrés laboral. Júnior et al.¹⁶ desarrollaron un modelo neuronal para predecir el síndrome metabólico en 2064 jóvenes brasileños de 18 y 19 años, basándose en edad, sexo, circunferencia de cintura, peso y presión arterial. Shin¹⁷, en un estudio con 5440 pacientes, incluyó variables como sexo biológico, edad, consumo de alcohol, ingesta calórica, estado civil, nivel educativo, ingresos, tabaquismo, consumo de algas y puntuación Z de riesgo poligénico. En este sentido, aún son escasas las investigaciones que han empleado redes neuronales con datos clínicos fácilmente obtenibles y provenientes de distintas regiones, lo que podría contribuir a mejorar la validez y generalización de los resultados.

En los últimos años, la inteligencia artificial, particularmente las redes neuronales artificiales (RNA), ha sido utilizada para desarrollar modelos predictivos con el objetivo de optimizar la detección temprana del

síndrome metabólico y mejorar la toma de decisiones clínicas. Júnior et al.¹⁶ implementaron una red neuronal de propagación hacia adelante para predecir el síndrome metabólico en adolescentes brasileños de 18 y 19 años, utilizando variables demográficas, antropométricas y clínicas no invasivas. Su modelo alcanzó una alta sensibilidad (89.8%) y una precisión del 86.8%, resaltando el potencial de estas herramientas para la detección del síndrome metabólico en poblaciones jóvenes sin necesidad de pruebas invasivas. Gutiérrez Esparza et al.¹⁸ desarrollaron un modelo de aprendizaje profundo para clasificar individuos con síndrome metabólico sin requerir análisis de sangre, utilizando datos de la cohorte Tlalpan 2020 en México, que incluían medidas antropométricas, hábitos alimenticios y calidad del sueño. Este modelo logró un valor predictivo positivo del 84.7% en mujeres y del 92.29% en hombres, demostrando la importancia de considerar factores del estilo de vida en la clasificación del SM.

En comparación con estas investigaciones, el presente estudio implementó una red neuronal tipo perceptrón multicapa para predecir el SM en una muestra diversa con datos de Venezuela, Tailandia e Indonesia. A diferencia de otros modelos que incorporan hábitos alimenticios o biomarcadores metabólicos avanzados, este modelo se basa en variables clínicas esenciales como CT, TG, HDL, obesidad e hipertensión. En este sentido, la red neuronal propuesta demostró una alta capacidad predictiva, alcanzando un AUC de 0.911, lo que respalda su utilidad en entornos clínicos con acceso limitado a pruebas diagnósticas sofisticadas.

El modelo desarrollado exhibió una precisión global del 84.20% en la fase de entrenamiento y del 81.80% en la fase de prueba, lo que refleja su solidez. Se observó un mejor desempeño en la predicción de casos sin síndrome metabólico (88.30% en entrenamiento y 86.60% en pruebas) en comparación con los casos positivos (77.80% en entrenamiento y 74.80% en pruebas). Esto podría deberse a la complejidad y variabilidad de los factores involucrados en el síndrome metabólico, lo que sugiere la necesidad de ajustes adicionales para mejorar la identificación de casos positivos.

La arquitectura del modelo incluyó una capa de entrada con cinco unidades (CT, TG, HDL, obesidad e hipertensión), una capa oculta con seis unidades y una capa de salida con dos unidades, lo que permitió manejar de manera eficiente la interacción entre múltiples factores de riesgo. Se emplearon funciones de activación tangente hiperbólica en las capas ocultas

y softmax en la capa de salida, optimizando la capacidad de clasificación del modelo. Además, el área bajo la curva (AUC) de 0.911 confirma su alto desempeño predictivo. Para evitar el sobreajuste y garantizar la generalización del modelo a nuevos datos, se implementaron validación cruzada y una regla de parada basada en la ausencia de disminución del error en un paso consecutivo.

La capacidad del modelo para integrar múltiples variables y predecir el SM de manera rápida y precisa tiene importantes implicancias clínicas. Su uso puede facilitar diagnósticos más tempranos y personalizados, permitiendo intervenciones preventivas y terapéuticas más efectivas. Asimismo, su implementación en estudios epidemiológicos contribuiría a una mejor estimación de la prevalencia real del SM. Desde una perspectiva económica, esta herramienta podría reducir los costos médicos para pacientes y sistemas de salud, tanto públicos como privados. Además, su empleo facilitaría la toma de decisiones clínicas, brindando a los médicos un algoritmo ágil y de fácil aplicación que complemente su juicio clínico, mejorando la identificación y la exclusión del SM en la práctica diaria.

No obstante, este estudio presenta algunas limitaciones. El uso de bases de datos secundarios puede introducir sesgos relacionados con la calidad y la homogeneidad de los datos recolectados en distintos países. Además, la mayor precisión en la predicción de casos sin SM en comparación con los casos positivos sugiere la necesidad de optimizar el modelo para mejorar su sensibilidad y equilibrio entre clases. A futuro, el modelo podría perfeccionarse para incrementar su capacidad de detección de casos positivos, y ser evaluado en estudios longitudinales que validen su efectividad a largo plazo, en particular cuando se logre un consenso unificado en los criterios diagnósticos del SM.

En conclusión, el uso de redes neuronales tipo MLP para la predicción del SM ha demostrado ser una herramienta eficaz, con un alto poder discriminativo. Su implementación en la práctica clínica podría revolucionar el enfoque diagnóstico y preventivo de esta condición multifactorial, contribuyendo significativamente a mejorar los resultados de salud en el corto y largo plazo.

Financiamiento

Financiado por el autor.

Conflicto de intereses

Sin conflicto de intereses.

Consideraciones éticas

Protección de personas y animales. Los autores declaran que para esta investigación no se han realizado experimentos en seres humanos ni en animales.

Confidencialidad, consentimiento informado y aprobación ética. El estudio no involucra datos personales de pacientes ni requiere aprobación ética. No se aplican las guías SAGER.

Declaración sobre el uso de inteligencia artificial. Los autores declaran que no utilizaron ningún tipo de inteligencia artificial generativa para la redacción de este manuscrito.

Bibliografía

1. Jha BK, Sherpa ML, Imran M, Mohammed Y, Jha LA, Paudel KR, et al. Progress in understanding metabolic syndrome and knowledge of its complex pathophysiology. *Diabetology*. 2023;4:134-59.
2. Mohamed SM, Shalaby MA, El-Shiekh RA, El-Banna HA, Emam SR, Bakr AF. Metabolic syndrome: risk factors, diagnosis, pathogenesis, and management with natural approaches. *Food Chemistry Advances*. 2023;3:100335.
3. Al-Mansoori L, Al-Jaber H, Prince MS, Elrayess MA. Role of inflammatory cytokines, growth factors and adipokines in adipogenesis and insulin resistance. *Inflammation*. 2022;45:31-44.
4. Guembe MJ, Fernandez-Lazaro CI, Sayon-Orea C, Toledo E, Moreno-Iribas C; RIVANA Study Investigators. Risk for cardiovascular disease associated with metabolic syndrome and its components: a 13-year prospective study in the RIVANA cohort. *Cardiovasc Diabetol*. 2020;19:195.
5. Noubiap JJ, Nansseu JR, Lontchi-Yimagou E, Nkeck JR, Nyaga UF, Ngouo AT, et al. Geographic distribution of metabolic syndrome and its components in the general adult population: a meta-analysis of global data from 28 million individuals. *Diabetes Res Clin Pract*. 2022;188:109924.
6. Echevarria-Castro N, Silva-Parra K, Polar-Trinidad M, Sánchez-Vicente JC, Salinas-Sedo G, Toro-Huamanchumo CJ. Concordance between different criteria for metabolic syndrome in Peruvian adults undergoing bariatric surgery. *J Clin Med*. 2022;11:4692.
7. Zong X, Bovet P, Xi B. A proposal to unify the definition of the metabolic syndrome in children and adolescents. *Front Endocrinol (Lausanne)*. 2022;13:925976.
8. Rottenkolber M, Gar C, Then C, Wanger L, Sacco V, Banning F, et al. A pathophysiology of type 2 diabetes unrelated to metabolic syndrome. *J Clin Endocrinol Metab*. 2021;106:1460-71.
9. Dobrowolski P, Prejbisz A, Kurylowicz A, Baska A, Burchardt P, Chlebus K, et al. Metabolic syndrome — a new definition and management guidelines. *Arch Med Sci*. 2022;18:1133-56.
10. Sauvageot N, Leite S, Alkerwi A, Sisanni L, Zannad F, Saverio S, et al. Association of empirically derived dietary patterns with cardiovascular risk factors: a comparison of PCA and RRR methods. *PLoS One*. 2016;11:e0161298.
11. Alberti KGMM, Eckel RH, Grundy SM, Zimmet PZ, Cleeman JI, Donato KA, et al. Harmonizing the metabolic syndrome: a joint interim statement of the International Diabetes Federation task force on epidemiology and prevention; National Heart, Lung, and Blood Institute; American Heart Association; World Heart Federation; International atherosclerosis society; and International Association for the Study of Obesity. *Circulation*. 2009;120:1640-5.
12. Radhakrishnan S, Nair SG, Isaac J. Multilayer perceptron neural network model development for mechanical ventilator parameters prediction by real time system learning. *Biomed Signal Process Control*. 2022;71:103170.
13. Data policies and repositories. F1000. 2023. (Consultado el 05-11-2024.) Disponible en: <https://www.f1000.com/resources-for-researchers/how-to-publish-your-research/data-policies-repositories/?form=MG0AV3>.
14. Hirose H, Takayama T, Hozawa S, Hibi T, Saito I. Prediction of metabolic syndrome using artificial neural network system based on clinical data including insulin resistance index and serum adiponectin. *Comput Biol Med*. 2011;41:1051-6.
15. Eyvazlou M, Hosseinpouri M, Mokarami H, Gharibi V, Jahangiri M, Cousins R, et al. Prediction of metabolic syndrome based on sleep and work-related risk factors using an artificial neural network. *BMC Endocr Disord*. 2020;20:169.
16. Júnior AC, França AK, Santos E dos, Silveira V, Santos A dos. Artificial neural networks to predict metabolic syndrome without invasive methods in adolescents. *J Clin Med*. 2024;13:5914.
17. Shin D. Prediction of metabolic syndrome using machine learning approaches based on genetic and nutritional factors: a 14-year prospective-based cohort study. *BMC Med Genomics*. 2024;17:224.
18. Gutiérrez-Esparza GO, Ramírez-del Real TA, Martínez-García M, Infante Vázquez O, Vallejo M, Hernández-Torruco J. Machine and deep learning applied to predict metabolic syndrome without a blood screening. *Appl Sci (Basel)*. 2021;11:4334.