

Ciência e Tecnologia

Para o Desenvolvimento
Ambiental, Cultural
e Socioeconômico

Xosé Somoza Medina
(organizador)

Ciência e Tecnologia

Para o Desenvolvimento
Ambiental, Cultural
e Socioeconômico

Xosé Somoza Medina
(organizador)

2025 by Editora Artemis
Copyright © Editora Artemis
Copyright do Texto © 2025 Os autores
Copyright da Edição © 2025 Editora Artemis



O conteúdo deste livro está licenciado sob uma Licença de Atribuição Creative Commons Atribuição-Não-Comercial NãoDerivativos 4.0 Internacional (CC BY-NC-ND 4.0). Direitos para esta edição cedidos à Editora Artemis pelos autores. Permitido o download da obra e o compartilhamento, desde que sejam atribuídos créditos aos autores, e sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

A responsabilidade pelo conteúdo dos artigos e seus dados, em sua forma, correção e confiabilidade é exclusiva dos autores. A Editora Artemis, em seu compromisso de manter e aperfeiçoar a qualidade e confiabilidade dos trabalhos que publica, conduz a avaliação cega pelos pares de todos manuscritos publicados, com base em critérios de neutralidade e imparcialidade acadêmica.

Editora Chefe	Profª Drª Antonella Carvalho de Oliveira
Editora Executiva	M.ª Viviane Carvalho Mocellin
Direção de Arte	M.ª Bruna Bejarano
Diagramação	Elisangela Abreu
Organizador	Prof. Dr. Xosé Somoza Medina
Imagen da Capa	peacestock/123RF
Bibliotecário	Maurício Amormino Júnior – CRB6/2422

Conselho Editorial

- Prof.º Dr.º Ada Esther Portero Ricol, *Universidad Tecnológica de La Habana "José Antonio Echeverría"*, Cuba
Prof. Dr. Adalberto de Paula Paranhos, Universidade Federal de Uberlândia, Brasil
Prof. Dr. Agustín Olmos Cruz, *Universidad Autónoma del Estado de México*, México
Prof.º Dr.º Amanda Ramalho de Freitas Brito, Universidade Federal da Paraíba, Brasil
Prof.º Dr.º Ana Clara Monteverde, *Universidad de Buenos Aires*, Argentina
Prof.º Dr.º Ana Júlia Viamonte, Instituto Superior de Engenharia do Porto (ISEP), Portugal
Prof. Dr. Ángel Mujica Sánchez, *Universidad Nacional del Altiplano*, Peru
Prof.º Dr.º Angela Ester Mallmann Centenaro, Universidade do Estado de Mato Grosso, Brasil
Prof.º Dr.º Begoña Blandón González, *Universidad de Sevilla*, Espanha
Prof.º Dr.º Carmen Pimentel, Universidade Federal Rural do Rio de Janeiro, Brasil
Prof.º Dr.º Catarina Castro, Universidade Nova de Lisboa, Portugal
Prof.º Dr.º Cirila Cervera Delgado, *Universidad de Guanajuato*, México
Prof.º Dr.º Cláudia Neves, Universidade Aberta de Portugal
Prof.º Dr.º Cláudia Padovesi Fonseca, Universidade de Brasília-DF, Brasil
Prof. Dr. Cleberton Correia Santos, Universidade Federal da Grande Dourados, Brasil
Dr. Cristo Ernesto Yáñez León – New Jersey Institute of Technology, Newark, NJ, Estados Unidos
Prof. Dr. David García-Martul, *Universidad Rey Juan Carlos de Madrid*, Espanha
Prof.º Dr.º Deuzimar Costa Serra, Universidade Estadual do Maranhão, Brasil
Prof.º Dr.º Dina Maria Martins Ferreira, Universidade Estadual do Ceará, Brasil
Prof.º Dr.º Edith Luévano-Hipólito, *Universidad Autónoma de Nuevo León*, México
Prof.º Dr.º Eduarda Maria Rocha Teles de Castro Coelho, Universidade de Trás-os-Montes e Alto Douro, Portugal
Prof. Dr. Eduardo Eugênio Spers, Universidade de São Paulo (USP), Brasil
Prof. Dr. Elio Martins Senhoras, Universidade Federal de Roraima, Brasil
Prof.º Dr.º Elvira Laura Hernández Carballedo, *Universidad Autónoma del Estado de Hidalgo*, México
Prof.º Dr.º Emilas Darlene Carmen Lebus, *Universidad Nacional del Nordeste / Universidad Tecnológica Nacional*, Argentina

Prof.^a Dr.^a Erla Mariela Morales Morgado, *Universidad de Salamanca*, Espanha
Prof. Dr. Ernesto Cristina, *Universidad de la República*, Uruguay
Prof. Dr. Ernesto Ramírez-Briones, *Universidad de Guadalajara*, México
Prof. Dr. Fernando Hitt, *Université du Québec à Montréal*, Canadá
Prof. Dr. Gabriel Diaz Cobos, *Universitat de Barcelona*, Espanha
Prof.^a Dr.^a Gabriela Gonçalves, Instituto Superior de Engenharia do Porto (ISEP), Portugal
Prof.^a Dr.^a Galina Gumovskaya – Higher School of Economics, Moscow, Russia
Prof. Dr. Geoffroy Roger Pointer Malpass, Universidade Federal do Triângulo Mineiro, Brasil
Prof.^a Dr.^a Gladys Esther Leoz, *Universidad Nacional de San Luis*, Argentina
Prof.^a Dr.^a Glória Beatriz Álvarez, *Universidad de Buenos Aires*, Argentina
Prof. Dr. Gonçalo Poeta Fernandes, Instituto Politécnico da Guarda, Portugal
Prof. Dr. Gustavo Adolfo Juarez, *Universidad Nacional de Catamarca*, Argentina
Prof. Dr. Guillermo Julián González-Pérez, *Universidad de Guadalajara*, México
Prof. Dr. Håkan Karlsson, *University of Gothenburg*, Suécia
Prof.^a Dr.^a Iara Lúcia Tescarollo Dias, Universidade São Francisco, Brasil
Prof.^a Dr.^a Isabel del Rosario Chiyon Carrasco, *Universidad de Piura*, Peru
Prof.^a Dr.^a Isabel Yohena, *Universidad de Buenos Aires*, Argentina
Prof. Dr. Ivan Amaro, Universidade do Estado do Rio de Janeiro, Brasil
Prof. Dr. Iván Ramón Sánchez Soto, *Universidad del Bío-Bío*, Chile
Prof.^a Dr.^a Ivânia Maria Carneiro Vieira, Universidade Federal do Amazonas, Brasil
Prof. Me. Javier Antonio Albornoz, *University of Miami and Miami Dade College*, Estados Unidos
Prof. Dr. Jesús Montero Martínez, *Universidad de Castilla - La Mancha*, Espanha
Prof. Dr. João Manuel Pereira Ramalho Serrano, Universidade de Évora, Portugal
Prof. Dr. Joaquim Júlio Almeida Júnior, UniFIMES - Centro Universitário de Mineiros, Brasil
Prof. Dr. Jorge Ernesto Bartolucci, *Universidad Nacional Autónoma de México*, México
Prof. Dr. José Cortez Godinez, Universidad Autónoma de Baja California, México
Prof. Dr. Juan Carlos Cancino Diaz, Instituto Politécnico Nacional, México
Prof. Dr. Juan Carlos Mosquera Feijoo, *Universidad Politécnica de Madrid*, Espanha
Prof. Dr. Juan Diego Parra Valencia, *Instituto Tecnológico Metropolitano de Medellín*, Colômbia
Prof. Dr. Juan Manuel Sánchez-Yáñez, *Universidad Michoacana de San Nicolás de Hidalgo*, México
Prof. Dr. Juan Porras Pulido, *Universidad Nacional Autónoma de México*, México
Prof. Dr. Júlio César Ribeiro, Universidade Federal Rural do Rio de Janeiro, Brasil
Prof. Dr. Leinig Antonio Perazolli, Universidade Estadual Paulista (UNESP), Brasil
Prof.^a Dr.^a Lívia do Carmo, Universidade Federal de Goiás, Brasil
Prof.^a Dr.^a Luciane Spanhol Bordignon, Universidade de Passo Fundo, Brasil
Prof. Dr. Luis Fernando González Beltrán, *Universidad Nacional Autónoma de México*, México
Prof. Dr. Luis Vicente Amador Muñoz, *Universidad Pablo de Olavide*, Espanha
Prof.^a Dr.^a Macarena Esteban Ibáñez, *Universidad Pablo de Olavide*, Espanha
Prof. Dr. Manuel Ramiro Rodriguez, *Universidad Santiago de Compostela*, Espanha
Prof. Dr. Manuel Simões, Faculdade de Engenharia da Universidade do Porto, Portugal
Prof.^a Dr.^a Márcia de Souza Luz Freitas, Universidade Federal de Itajubá, Brasil
Prof. Dr. Marcos Augusto de Lima Nobre, Universidade Estadual Paulista (UNESP), Brasil
Prof. Dr. Marcos Vinicius Meiado, Universidade Federal de Sergipe, Brasil
Prof.^a Dr.^a Mar Garrido Román, *Universidad de Granada*, Espanha
Prof.^a Dr.^a Margarida Márcia Fernandes Lima, Universidade Federal de Ouro Preto, Brasil
Prof.^a Dr.^a María Alejandra Arecco, *Universidad de Buenos Aires*, Argentina
Prof.^a Dr.^a Maria Aparecida José de Oliveira, Universidade Federal da Bahia, Brasil
Prof.^a Dr.^a Maria Carmen Pastor, *Universitat Jaume I*, Espanha
Prof.^a Dr.^a Maria da Luz Vale Dias – Universidade de Coimbra, Portugal
Prof.^a Dr.^a Maria do Céu Caetano, Universidade Nova de Lisboa, Portugal

Prof.ª Dr.ª Maria do Socorro Saraiva Pinheiro, Universidade Federal do Maranhão, Brasil
Prof.ª Dr.ª M³ Graça Pereira, Universidade do Minho, Portugal
Prof.ª Dr.ª Maria Gracinda Carvalho Teixeira, Universidade Federal Rural do Rio de Janeiro, Brasil
Prof.ª Dr.ª María Guadalupe Vega-López, Universidad de Guadalajara, México
Prof.ª Dr.ª Maria Lúcia Pato, Instituto Politécnico de Viseu, Portugal
Prof.ª Dr.ª Maritza González Moreno, Universidad Tecnológica de La Habana, Cuba
Prof.ª Dr.ª Mauriceia Silva de Paula Vieira, Universidade Federal de Lavras, Brasil
Prof. Dr. Melchor Gómez Pérez, Universidad del País Vasco, Espanha
Prof.ª Dr.ª Ninfa María Rosas-García, Centro de Biotecnología Genómica-Instituto Politécnico Nacional, México
Prof.ª Dr.ª Odara Horta Boscolo, Universidade Federal Fluminense, Brasil
Prof. Dr. Osbaldo Turpo-Gebera, Universidad Nacional de San Agustín de Arequipa, Peru
Prof.ª Dr.ª Patrícia Vasconcelos Almeida, Universidade Federal de Lavras, Brasil
Prof.ª Dr.ª Paula Arcoverde Cavalcanti, Universidade do Estado da Bahia, Brasil
Prof. Dr. Rodrigo Marques de Almeida Guerra, Universidade Federal do Pará, Brasil
Prof. Dr. Saulo Cerqueira de Aguiar Soares, Universidade Federal do Piauí, Brasil
Prof. Dr. Sergio Bitencourt Araújo Barros, Universidade Federal do Piauí, Brasil
Prof. Dr. Sérgio Luiz do Amaral Moretti, Universidade Federal de Uberlândia, Brasil
Prof.ª Dr.ª Silvia Inés del Valle Navarro, Universidad Nacional de Catamarca, Argentina
Prof.ª Dr.ª Solange Kazumi Sakata, Instituto de Pesquisas Energéticas e Nucleares (IPEN)- USP, Brasil
Prof.ª Dr.ª Stanislava Kashtanova, Saint Petersburg State University, Russia
Prof.ª Dr.ª Susana Álvarez Otero – Universidad de Oviedo, Espanha
Prof.ª Dr.ª Teresa Cardoso, Universidade Aberta de Portugal
Prof.ª Dr.ª Teresa Monteiro Seixas, Universidade do Porto, Portugal
Prof. Dr. Valter Machado da Fonseca, Universidade Federal de Viçosa, Brasil
Prof.ª Dr.ª Vanessa Bordin Viera, Universidade Federal de Campina Grande, Brasil
Prof.ª Dr.ª Vera Lúcia Vasilévski dos Santos Araújo, Universidade Tecnológica Federal do Paraná, Brasil
Prof. Dr. Wilson Noé Garcés Aguilar, Corporación Universitaria Autónoma del Cauca, Colômbia
Prof. Dr. Xosé Somoza Medina, Universidad de León, Espanha

**Dados Internacionais de Catalogação na Publicação (CIP)
(eDOC BRASIL, Belo Horizonte/MG)**

C569 Ciência e tecnologia para o desenvolvimento ambiental, cultural e socioeconômico VII [livro eletrônico] / Organizador Xosé Somoza Medina. – Curitiba, PR: Artemis, 2025.

Formato: PDF

Requisitos de sistema: Adobe Acrobat Reader

Modo de acesso: World Wide Web

Inclui bibliografia

Edição bilíngue

ISBN 978-65-81701-76-5

DOI 10.37572/EdArt_091225765

1. Desenvolvimento sustentável. 2. Tecnologia – Aspectos ambientais. I. Somoza Medina, Xosé.

CDD 363.7

Elaborado por Maurício Amormino Júnior – CRB6/2422



PRÓLOGO

El presente volumen de **Ciencia y Tecnología para el Desarrollo Ambiental, Cultural y Socioeconómico** reúne investigaciones que, aunque diversas en métodos, contextos y objetos de estudio, convergen en un propósito común: comprender y transformar las realidades socioambientales mediante la ciencia aplicada, la innovación tecnológica y el diálogo interdisciplinario. Para orientar esta pluralidad, como en ediciones anteriores, los capítulos se organizan en dos grandes ejes temáticos: Ciencia y Tecnología. No obstante, que nadie piense que los trabajos del primer bloque carecen de visión y aplicaciones tecnológicas o que los del segundo no nacen del más sincero rigor científico, es simplemente una forma de organizar los saberes, como en la antigüedad clásica que diferenciaban entre Artes Mayores y Menores.

La primera parte, **Ciencia**, agrupa trabajos que examinan procesos históricos, sociales y productivos vinculados al desarrollo regional, en los que está presente, en todos ellos, las características propias de los lugares como elemento fundamental. Incluye la caracterización de tecnologías constructivas en sitios arqueológicos del Periodo Formativo andino, el análisis del bovino criollo mixteco como recurso zoogenético estratégico y motor potencial de desarrollo comunitario, la caracterización del sector frutícola en provincias clave de Argentina y la aplicación de técnicas geoestadísticas combinadas con el concepto C4 para el mapeo objetivo de biomasa que ayude a mitigar la plaga de los incendios forestales. Aunque provenientes de campos distintos, estos trabajos comparten una visión amplia de sostenibilidad, en la que convergen la preservación cultural, la innovación social, la valorización de recursos locales y la aplicación de metodologías experimentales con impacto directo en la gestión ambiental.

La segunda parte, **Tecnología**, reúne estudios que abordan el uso de la inteligencia artificial en diagnósticos médicos, la seguridad alimentaria, fenómenos fisicoquímicos y sistemas inteligentes. Los capítulos aquí incluidos muestran cómo la ingeniería, la física aplicada, la química y las ciencias de los materiales pueden ofrecer soluciones concretas a problemas actuales. Este eje integra estudios sobre la evaluación comparativa de arquitecturas de inteligencia artificial aplicadas al diagnóstico temprano de diabetes tipo 2, el encogimiento del chícharo en secadores de lecho fluidizado para mejorar su conservación, el uso de aplicadores de microondas para un más eficiente tratamiento térmico de suelos y el diseño de sistemas domóticos alimentados con energía fotovoltaica. Se trata de contribuciones que destacan la relevancia de la investigación aplicada para avanzar en la eficiencia energética, la mejora de la salud y el desarrollo de prácticas sostenibles de alta precisión.

De este modo, el volumen VII reafirma el espíritu interdisciplinario de la colección *Ciencia y Tecnología para el Desarrollo Ambiental, Cultural y Socioeconómico*. Aquí, la tecnología dialoga con el territorio, la modelización científica se entrelaza con saberes tradicionales y la innovación se presenta como puente entre los desafíos ambientales y el bienestar colectivo. Confiamos en que esta obra inspire nuevas investigaciones, motive lecturas críticas y contribuya al fortalecimiento de prácticas orientadas hacia un futuro más sostenible, justo e integrado.

Xosé Somoza Medina

SUMÁRIO

CIÊNCIA

CAPÍTULO 1.....1

TECNOLOGÍA CONSTRUCTIVA EN EL SITIO ARQUEOLÓGICO BUENAVISTA (LA PAMPA, CORONGO – ÁNCASH) DURANTE EL PERÍODO FORMATIVO

Efraín Vidal Espinoza

Elsa Celina Sánchez Cacha

Cesar Augusto Serna Lamas

Fredy Osvaldo Loli Natividad

Leopoldo Enrique Neglia Valderrama

Victor Alberto Pocoy Yauri

José Yovera Saldarriaga

 https://doi.org/10.37572/EdArt_0912257651

CAPÍTULO 2.....17

REVALORIZACIÓN Y CONSERVACIÓN GENÉTICA ANTE LAS DEMANDAS ECO-SALUDABLES Y ESTRATEGIAS DE VALOR AGREGADO COMUNITARIO

Martín Reyes García

 https://doi.org/10.37572/EdArt_0912257652

CAPÍTULO 3.....27

CARACTERIZACIÓN DEL SECTOR FRUTÍCOLA EN LAS PROVINCIAS DE BUENOS AIRES, SANTA FE Y ENTRE RÍOS (REPÚBLICA ARGENTINA)

Victorina Mariana Rivera Rúa

Violeta Riolfo

Rubén Marcelo Coniglio

 https://doi.org/10.37572/EdArt_0912257653

CAPÍTULO 4.....40

CONCEPTO CONDICIONAL DE CARGAS COMBUSTIBLES -C4- PARA LA SELECCIÓN DE MODELOS DE COMBUSTIBLES FORESTALES

José German Flores-Garnica

 https://doi.org/10.37572/EdArt_0912257654

TECNOLOGIA

CAPÍTULO 5.....	54
-----------------	----

COMPARACIÓN DE ARQUITECTURAS DE APRENDIZAJE PROFUNDO GENERADAS POR IA (COPILOT VS. DEEPSEEK) EN LA CLASIFICACIÓN DEL CONJUNTO DE DATOS PIMA INDIANS DIABETES

Moisés García Villanueva

Salvador Ramírez Zavala

 https://doi.org/10.37572/EdArt_0912257655

CAPÍTULO 6.....	72
-----------------	----

INFLUENCIA DE LA VELOCIDAD DE AIRE Y TEMPERATURA EN EL ENCOGIMIENTO DEL CHÍCHARO POR SECADO

Alejandro Díaz Medina

Alexis Marco Antonio Romero Pérez

Miguel Angel Pérez Fajardo

Noe Jardon Romero Pérez

Ivan Vladimir Molina Jaramillo

Francisco Vidal Caballero Dominguez

 https://doi.org/10.37572/EdArt_0912257656

CAPÍTULO 7.....	82
-----------------	----

EXPERIMENTAL STUDY OF MICROWAVE SLOW WAVE COMB AND CERAMIC APPLICATORS FOR SOIL TREATMENT AT FREQUENCY 2.45 GHZ

Grigory Torgovnikov

Graham Brodie

 https://doi.org/10.37572/EdArt_0912257657

CAPÍTULO 8.....	93
-----------------	----

SISTEMAS DOMÓTICOS ALIMENTADOS CON ENERGÍA FOTOVOLTAICA

Guadalupe García Toribio

Heriberto Hernández Rodríguez

José Luis Gutiérrez García

Román Gutiérrez Marcos

Israel Ildefonso García

Luisa Angelica Viñas Meza

 https://doi.org/10.37572/EdArt_0912257658

SOBRE O ORGANIZADOR.....	109
--------------------------	-----

ÍNDICE REMISSIVO	110
------------------------	-----

CAPÍTULO 5

COMPARACIÓN DE ARQUITECTURAS DE APRENDIZAJE PROFUNDO GENERADAS POR IA (COPILOT VS. DEEPSEEK) EN LA CLASIFICACIÓN DEL CONJUNTO DE DATOS PIMA INDIANS DIABETES

Data de submissão: 01/11/2025

Data de aceite: 18/11/2025

Moisés García Villanueva

Facultad de Ingeniería Eléctrica
Universidad Michoacana de
San Nicolás de Hidalgo
Morelia, Michoacán, México

<https://orcid.org/0000-0002-8567-309X>

Salvador Ramírez Zavala

Facultad de Ingeniería Eléctrica
Universidad Michoacana de
San Nicolás de Hidalgo
Morelia, Michoacán, México

<https://orcid.org/0000-0001-5480-6593>

RESUMEN: La Diabetes Mellitus Tipo 2 (DM2) es una enfermedad crónica que afecta a millones de personas en el mundo, su diagnóstico temprano es crucial para prevenir complicaciones. Existe la necesidad de desarrollar sistemas que apoyen el diagnóstico a tiempo de esta

enfermedad. En este trabajo se propone evaluar los modelos de aprendizaje profundo que es posible generar a través de una herramienta de Inteligencia Artificial (IA), se estableció utilizar plataformas que permiten su uso gratuito (Copilot y DeepSeek), y para el estudio se emplea el conjunto de datos Pima Indians Diabetes, un conjunto que se cita en cualquier trabajo de aprendizaje de máquina que involucra la DM2. El objetivo es determinar cuál propuesta algorítmica ofrece mayor precisión, robustez y generalización. Con la evidencia empírica obtenida, un desarrollador justifica el uso de dichas arquitecturas en el desarrollo de aplicaciones médicas que requieren implementar el diagnóstico preliminar de pacientes con riesgo de DM2. Se evaluaron arquitecturas de redes neuronales profundas, generadas a partir de diferentes frases que contextualizan los requerimientos en el prompt de las herramientas de IA; los resultados obtenidos indican que las arquitecturas logran obtener una precisión en el mejor de los casos superior al 89%, mientras que la sensibilidad superó el 91% en otro de los modelos. Los resultados demuestran que DeepSeek puede ser una herramienta más eficaz para desarrollar modelos que permitan la creación de aplicaciones para el diagnóstico médico.

PALABRAS CLAVE: IA en salud; LLMs; DeepSeek; Copilot; Redes Neuronales Convolucionales; Diabetes Mellitus tipo 2.

COMPARISON OF AI-GENERATED DEEP LEARNING ARCHITECTURES (COPILOT VS. DEEPSEEK) IN THE CLASSIFICATION OF THE PIMA INDIANS DIABETES DATASET

ABSTRACT: Type 2 Diabetes Mellitus (DM2) is a chronic disease affecting millions of people worldwide, and its early diagnosis is crucial for preventing complications. There is a need to develop systems that support the timely diagnosis of this disease. This work proposes to evaluate deep learning models that can be generated using an Artificial Intelligence (AI) tool. Platforms that allow free use (Copilot and DeepSeek) were used, and the study employed the Pima Indians Diabetes dataset, a dataset cited in any machine learning work involving DM2. The objective is to determine which algorithmic approach offers greater accuracy, robustness, and generalizability. Using the empirical evidence obtained, a developer justifies the use of these architectures in the development of medical applications that require the preliminary diagnosis of patients at risk of type 2 diabetes. Deep neural network architectures, generated from different phrases that contextualize the requirements in the AI tool prompt, were evaluated. The results indicate that the architectures achieve an accuracy exceeding 89% in the best-case scenario, while sensitivity surpassed 91% in another model. These results demonstrate that DeepSeek can be a more effective tool for developing models that enable the creation of medical diagnostic applications.

KEYWORDS: AI in health; LLMs; DeepSeek; Copilot; Convolutional Neural Networks; Type 2 Diabetes Mellitus.

1. INTRODUCCIÓN

Las enfermedades crónicas tienen una gran prevalencia alrededor del mundo e impactan significativamente la calidad de vida de las personas y sus familias. La Organización Mundial de la Salud clasifica la DM2 como una epidemia global, con un costo estimado de \$966 mil millones anuales en gastos sanitarios (WHO, 2023). Ante este escenario, soluciones basadas en IA pueden reducir el costo económico mediante la implementación de las siguientes estrategias:

- Automatización de diagnósticos preliminares: Útil en regiones con acceso limitado a médicos.

- Identificación de patrones de riesgo: Como la correlación entre resistencia a la insulina y obesidad (Zheng et al., 2022).

En esta nueva era de la Inteligencia Artificial (IA) se puede revolucionar y potenciar la investigación y la práctica de la medicina, además de muchas otras áreas. La IA puede ayudar a mejorar la exactitud del diagnóstico, debido a que permite analizar grandes conjuntos de datos con información de pacientes, logrando identificar patrones que son difíciles de detectar por los humanos (Brinker et al., 2019, Sharma et al., 2023). Los padecimientos crónico-degenerativos, como la DM2, una enfermedad caracterizada por la resistencia a la insulina y deficiencia relativa de su producción, es un caso en el que se debe considerar de suma importancia implementar políticas para un diagnóstico preliminar, esta acción juega un papel muy importante, su dimensión repercute en gran medida en las problemáticas identificadas y propiciadas por el padecimiento, que al ser implementado permite optimizar recursos y sobrellevar de la mejor forma las dificultades que se enfrentan por esta enfermedad.

La automatización de diagnósticos preliminares mediante aplicaciones que utilizan Inteligencia Artificial (IA) representa una solución transformadora para sistemas de salud en regiones con escasos recursos médicos, donde la relación de médicos por paciente puede ser extremadamente baja (WHO, 2023). En el caso de la Diabetes Mellitus Tipo 2, una enfermedad que puede prevenirse, siempre y cuando exista el compromiso por parte del paciente, con altas tasas de subdiagnóstico en países en desarrollo, los algoritmos de Aprendizaje Profundo (DL por sus siglas en Inglés de Deep Learning) pueden analizar datos clínicos básicos (glucosa en ayunas, IMC (Índice de Masa Corporal), historial familiar) y generar evaluaciones de riesgo tempranas, facilitando con ello contribuir en las siguientes problemáticas:

- La detección temprana: Modelos entrenados con conjuntos de datos públicos (como Pima Indians Diabetes) pueden identificar pacientes en estado prediabético (HbA1c entre 5.7% y 6.4%) con una precisión superior al 85% (Agarwal et al., 2022).
- La reducción de costos: Un estudio en India demostró que el uso de la IA para pruebas de detección temprana o tamizajes médicos preliminares disminuyó un 30% las asignaciones o atenciones innecesarias por endocrinólogos (Patel et al., 2023).

- La telemedicina integrada con IA: Aplicaciones móviles que usan modelos ligeros (ej. Random Forest optimizado) permiten a trabajadores comunitarios realizar pruebas rápidas de datos clínicos o químicos en aldeas remotas, enviando alertas a centros urbanos solo para casos de alto riesgo (UNICEF, 2023).

Para lograrlo, se hace necesario la creación de modelos de IA que puedan implementarse de forma rápida, con códigos robustos y reproducibles. Además de la colaboración de desarrolladores de software y médicos para validar los resultados. En la actualidad existen múltiples herramientas de IA que permiten la generación de código que implementa algoritmos de clasificación de datos y este hecho facilita la implementación de aplicaciones para diagnósticos clínicos. Surgen entonces varias preguntas en relación al resultado obtenido (código generado) por las IA, tales como: ¿Cuál es la IA con el código más eficiente?, ¿Es posible lograr la eficiencia que reportan los estudios previos o son más eficientes que las soluciones publicadas en el área?, ¿Qué enfoque es el más completo?, etcétera. La respuesta a estos cuestionamientos, entre otras cosas, justifica la comparación de las arquitecturas proporcionadas por las herramientas de IA. El documentar el análisis de resultados, proporciona a los desarrolladores de software evidencias empíricas de la eficacia obtenida por los modelos generados por la IA para documentar su desarrollo e implementación y además, permite apoyar alguna de las metodologías de desarrollo de software.

Generación de código con Modelos del Lenguaje Grande

El rápido surgimiento de las tecnologías de inteligencia artificial (IA) ha impulsado un auge en la investigación centrada en la evaluación de las capacidades de generación de código. Los Modelos de Lenguaje Grande (LLMs por sus siglas en Inglés de Large Language Models) son cada vez más fiables para una amplia gama de usuarios, incluidos programadores principiantes y desarrolladores experimentados. Esta accesibilidad representa una importante oportunidad para ampliar la educación informática y el desarrollo de software, transformando potencialmente tanto los entornos de aprendizaje autodidacta como los de educación superior formal. Los LLMs ofrecen aplicaciones prometedoras en la automatización de la generación de tareas y proporcionan retroalimentación personalizada, además de apoyar la comprensión conceptual en los cursos de programación (Shakia, et. al., 2025).

La integración de la IA en la programación ha transformado el panorama del aprendizaje en este campo, ofreciendo así beneficios a estudiantes, educadores y desarrolladores profesionales. En particular, los modelos de lenguaje natural (MLN) se han consolidado como potentes herramientas de IA en programación, aprovechando redes neuronales profundas entrenadas con grandes cantidades de datos textuales. Cuando estos modelos se entranan con extensas bases de código, adquieren la capacidad de generar código similar al escrito por humanos a partir de indicaciones en lenguaje natural (Yadav, et. al., 2025; Tian, et. al., 2023). Estudios recientes destacan que Modelos de Lenguajes Grandes como DeepSeek y Copilot superan a alternativas generalistas en tareas de bioinformática, gracias a su entrenamiento en la literatura médica (Wagner et al., 2024).

Este trabajo aporta evidencia empírica sobre las herramientas de IA que permiten obtener código para generar los modelos que clasifican datos específicos que se han utilizado en estudios previos en la predicción de enfermedades crónicas, nuestro caso de estudio comprende el conjunto de datos Pima Indians Diabetes, un conjunto clásico en la literatura para los estudios realizados de la DM2. Es además el primer trabajo que realiza una comparación de dos IA que reducen la brecha en el uso de la tecnología para generar código y que permite implementar sistemas para la clasificación de datos en la enfermedad DM2. Los resultados obtenidos indican que las arquitecturas proporcionadas por las IAs tienen una mejor eficacia que otras técnicas de aprendizaje de máquina utilizando el mismo conjunto de datos y se cumple que con la generación de código automático es posible reducir el desarrollo de sistemas y aplicaciones de software para el diagnóstico médico o en general en otras áreas.

2. MATERIALES Y MÉTODO

2.1. CONJUNTO DE DATOS PIMA INDIANS DIABETES

El conjunto de datos Pima Indians Diabetes es un recurso fundamental en la investigación médica y la ciencia de datos aplicada a la salud, particularmente en el estudio de la DM2. Su relevancia radica en los siguientes aspectos (Rajeswari y Suthendran, 2022):

1. Población de Alto Riesgo: Contiene registros de mujeres de la tribu Pima de Arizona (EE.UU.), una población con una de las mayores prevalencias

- de diabetes en el mundo debido a factores genéticos y cambios en el estilo de vida. Proporciona información valiosa para estudiar disparidades en salud y enfermedades metabólicas en grupos étnicos específicos.
2. Comparativo de referencia en Aprendizaje de Máquina Médico: Es ampliamente utilizado para evaluar modelos predictivos de diabetes debido a su accesibilidad y características clínicamente relevantes. Ha servido como base para investigaciones en diagnóstico temprano, clasificación de riesgo y medicina preventiva.
 3. Variables Clínicas Relevantes: Consta de las variables clave como glucosa en sangre, índice de masa corporal, nivel de insulina y antecedentes familiares (pedigrí de diabetes), que son esenciales para estudios epidemiológicos.
 4. Impacto en la Inteligencia Artificial en Salud: Ha sido utilizado en cientos de estudios para desarrollar sistemas de alerta temprana, modelos de riesgo y herramientas de apoyo a la decisión clínica.

Estadísticas Clave del Conjunto de Datos

Las estadísticas descriptivas esenciales para un análisis médico son:

1. Estructura del Conjunto de Datos: El conjunto de datos Pima Indians Diabetes disponible cuenta con: Número de registros: 768, Variables: 8 características clínicas; 1 variable con la clasificación verdadera (diabetes: 1 = sí, 0 = no).
2. Distribución de Clases (Diabetes vs No Diabetes): El conjunto consiste en dos clases, No diabéticos con 500 registros, lo que representa el 65.1% del total del conjunto; mientras que los registros para la clase diabéticos son 268, es decir, representa el 34.9% del conjunto. Estas cantidades establecen que existe un desequilibrio de las clases, por lo que debe considerarse este punto como relevante para el análisis de sensibilidad y el resto de métricas en los resultados.
3. Las estadísticas descriptivas de las variables numéricas que se consideran son: media, desviación estándar, mínimo y máximo de las variables involucradas: Embarazos, glucosa, presión arterial, espesor de piel, insulina, IMC, pedigree diabetes y la edad, como se muestra en la Tabla 1.

Tabla 1. Estadísticas descriptivas de las variables en el conjunto de datos.

Variable	Media	Desviación Estándar	Mínimo	Máximo
Embarazos	3.85	3.37	0	17
Glucosa(mg/dL)	120.89	31.97	0	199
Presión arterial (mmHg)	69.11	19.36	0	122
Espesor piel (mm)	20.54	15.95	0	99
Insulina (mU/mL)	79.80	115.24	0	846
IMC (kg/m^2)	31.99	7.88	0	67.1
Pedigree Diabetes	0.47	0.33	0.078	2.42
Edad	33.24	11.76	21	81

2.2. HERRAMIENTAS DE IA UTILIZADAS

Existe una gran cantidad de herramientas de IA disponibles, algunas tienen un costo y otras no se utilizan específicamente para programación. El objetivo de este trabajo no es comparar la totalidad de herramientas especializadas en generar código o modelos para clasificar datos, se pretende proporcionar evidencia empírica con herramientas gratuitas, que permitan generar código y de acuerdo a la experiencia de los autores de este trabajo proporcionan las mejores propuestas de código de programación.

2.2.1. Copilot y DeepSeek

GitHub Copilot es una herramienta de IA generativa de código desarrollada por GitHub en colaboración con OpenAI (Github, 2023). Funciona como un asistente de programación en tiempo real, integrado en entornos de desarrollo de software, en los que sugiere fragmentos de código completos basados en comentarios descriptivos o contexto del proyecto.

A principios de 2025, la startup tecnológica china DeepSeek presentó DeepSeek R1 (DeepSeek, 2023), que pronto tuvo un gran impacto, entre ellos la fuerte caída de varias importantes acciones tecnológicas estadounidenses.

DeepSeek R1, que demostró un rendimiento de razonamiento prometedor frente a OpenAI a un coste competitivo por millón de tokens de salida en la API (2 \$ frente a 60 \$), despertó un gran interés público. Uno de los factores que contribuyó a este interés fue el bajo coste de entrenamiento que, según se afirmaba, tenía el modelo (6 millones de dólares para entrenar DeepSeek VS 78 millones que gastó OpenAI). Dicha eficiencia es el resultado de una serie de estrategias de ingeniería sobre la base de las técnicas LLM existentes, que se han ido aplicando gradualmente desde la época del primer modelo DeepSeek. Se le atribuye también a DeepSeek ser pionero en la democratización del desarrollo de modelos LLMs (al convertir el modelo en código abierto) (Puspitasari, et. al., 2025). En nuestro contexto, DeepSeek es un modelo avanzado de IA diseñado para asistir en el desarrollo de aplicaciones, logra proporcionar código que es utilizado en el campo asistido por aprendizaje de máquina. Su capacidad para generar, optimizar y explicar código lo convierte en una herramienta valiosa para investigadores, desarrolladores y profesionales que buscan implementar soluciones basadas en datos (Wagner, et al., 2024).

2.3. METODOLOGÍA

Para comparar los códigos generados por cada una de las herramientas, se estableció que se escribieran en el prompt de cada una de ellas frases contextuales que se muestran en la Tabla 2. La intención del primer texto es que sea muy simple, indicando con esto la posibilidad de interacción de un principiante en el uso de la herramienta. Los textos para los modelos posteriores tienen la intención de contextualizar y proporcionar mayores detalles de lo que se pretende con el código generado, es decir, indicar el objetivo principal del clasificador.

Tabla 2.- Textos de prueba utilizados en las herramientas de IA para generar códigos de programación de computadoras para obtener un modelo que clasifica el conjunto de datos Pima Indians Diabetes.

No. de Modelo	Texto escrito en el prompt de la herramienta de IA
1	crear una red neuronal para clasificar pima indians diabetes
2	Proponer una versión mejorada de la arquitectura que se obtuvo previamente para clasificar el conjunto de datos pima indians diabetes, quiero superar el 82% de precisión

No. de Modelo	Texto escrito en el prompt de la herramienta de IA
3	¿Cuál es la mejor arquitectura de aprendizaje profundo que se puede utilizar para clasificar pima-indians-diabetes? ¿será posible lograr una precisión mayor del 0.90?
4	Propón un modelo avanzado de una arquitectura CNN para clasificar pima indians diabetes y lograr una precisión mayor al 80%

El código generado por las herramientas de IA determina el número de modelo (columna izquierda de la Tabla 2), de donde se obtiene la arquitectura de red neuronal, un ejemplo gráfico para el modelo 1 se ilustra en la Tabla 1. El lenguaje de programación que determinó la inteligencia artificial fue python y las librerías para los modelos estuvieron basadas en keras y tensorflow. En el caso de la Copilot, las 4 arquitecturas propuestas corresponden a modelos secuenciales, mientras que DeepSeek el 50% de sus propuestas fueron modelos secuenciales y el resto propuestas de arquitecturas más complejas.

Cabe destacar que las herramientas de IA solamente proporcionan el código fuente al interactuar con el prompt, por lo que en el presente trabajo se describe el modelo 1 en forma gráfica y con las características de cada arquitectura generada por la IA, esto se ilustra en la Tabla 3, por ejemplo la arquitectura secuencial de Copilot que se muestra en forma gráfica, consiste de 3 capas (columna 1 de la Tabla 3), todas ellas del mismo tipo (densas), en ella se señalan las dimensiones de entrada, el tipo de capa y la dimensión de salida de cada capa; por otro lado, la arquitectura propuesta por DeepSeek para este mismo modelo, consiste en 5 capas (columna 2 de la Tabla 3), donde 3 son del tipo denso y dos capas Dropout; finalmente algunas características de la cantidad de parámetros de la red neuronal se describe en la columna 3 de la Tabla 3.

Tabla 3. Modelos generados por las herramientas de IA al proporcionar el texto: crear una red neuronal para clasificar pima indians diabetes.

Copilot	DeepSeek	Características
<pre> graph TD dense1[dense (Dense) Input shape: (None, 8) Output shape: (None, 12)] --> dense2[dense_1 (Dense) Input shape: (None, 12) Output shape: (None, 8)] dense2 --> dense3[dense_2 (Dense) Input shape: (None, 8) Output shape: (None, 1)] </pre>	<pre> graph TD dropout1[dropout (Dropout) Input shape: (None, 8) Output shape: (None, 16)] --> dense1_1[dense_1 (Dense) Input shape: (None, 16) Output shape: (None, 8)] dense1_1 --> dropout1_1[dropout_1 (Dropout) Input shape: (None, 8) Output shape: (None, 8)] dropout1_1 --> dense2_1[dense_2 (Dense) Input shape: (None, 8) Output shape: (None, 1)] </pre>	<p>Copilot: 3 capas, no incluye dropout Total parámetros: 665 Parámetros entrenables: 221 Parámetros No entrenables: 0 Parámetros del Optimizador: 444</p> <p>DeepSeek: 5 capas, Total parámetros: 869 Parámetros entrenables: 289 Parámetros No entrenables: 0 Parámetros del optimizador: 580</p> <p>Capas Densas Capas dropout</p>

La cantidad de parámetros en una red neuronal impacta en el rendimiento, una mayor cantidad de parámetros indica mayor capacidad de aprendizaje, desde luego que también mayor riesgo de sobreentrenamiento de la red neuronal, además de requerir una cantidad mayor de datos y recursos computacionales para el entrenamiento de la red neuronal. En la Tabla 4 se presenta una descripción de los modelos 2 a 4 de la Tabla 1, obtenidos por las herramientas de IA al responder con el texto escrito en el prompt.

Tabla 4. Descripción de las arquitecturas generadas por las herramientas de IA, ante la respuesta del texto escrito en el prompt de cada una de ellas.

Modelo	Copilot	DeepSeek
2	<p>Modelo Secuencial Capas Totales: 5 I(8) Densa O(64), I(64)Dropout O(64), I(64)Densa O(32), I(32) Dropout O(32), I(32)Densa O(1) Total de parámetros: 8069</p>	<p>Modelo Secuencial Capas Totales: 10 I(8) Densa O(32), I(32) Batch Normalization O(32), I(32)Dropout O(32), I(32)Densa O(64), I(64) Batch Normalization O(64), I(64) Dropout O(64), I(64)Densa O(32), I(32) Batch Normalization O(32), I(32)Dropout O(32), I(32) Densa O(1) Total de parámetros: 14566</p>
3	<p>Modelo Secuencial Capas Totales: 8 I(8) Densa O(64), I(64) Batch Normalization O(64), I(64) Dropout O(64), I(64)Densa O(32), I(32) Batch Normalization O(32), I(32) Dropout O(32), I(32)Densa O(16), I(16)Densa O(1) Total de parámetros: 10373</p>	<p>Modelo Híbrido: Red Residual con Concatenación (ResNet/DenseNet) Capas Totales: 15 Capa de Entrada O(11), I(11) Densa O(64), I(64) Batch Normalization O(64), I(64)Dropout O(64) I(64)Densa O(128), I(128) Batch Normalization O(128), I(128)Dropout O(128), I(126)Densa O(64), I(64) Batch Normalization O(64) I(64,64) Concatena (128), I(128)Densa O(32), I(32) Batch Normalization O(32), I(32)Dropout O(32), I(32)Densa O(1) Total de Parámetros: 66821</p>
4	<p>Modelo Secuencial: Red Neuronal Convolucionar 2D Capas Totales: 10 I(8,1,1) Conv2D O(7,1,32), I(7,1,32) MaxPooling 2D O(3,1,32), I(3,1,32) Dropout O(3,1,32), I(3,1,32) Conv2D O(2,1,64), I(2,1,64) MaxPooling 2D O(1,1,64), I(1,1,64) Dropout O(1,1,64), I(1,1,64)Flatten O(64), I(64)Densa O(128), I(128) Dropout O(128), I(128)Densa O(1) Total de parámetros: 38,117</p>	<p>Modelo Red Neuronal Convolucionar 1D Capas Totales: 17 Capa de Entrada O(10,1), I(10,1) Conv1D O(10,64), I(10,64) Batch Normalization O(10,64), I(10,64)Dropout O(10,64), I(10,64) MaxPooling 1D (5,64), I(5,64) Conv1D O(5,128), I(5,128) Batch Normalization O(5,128), I(5,128)Dropout O(5,128), I(5,128) MaxPooling 1D (2,128), I(2,128) Conv1D O(2,256), I(2,256) Batch Normalization O(2,256), I(2,256)Dropout O(2,256), I(2,256) Flatten (512), I(512)Densa O(128), I(128) Batch Normalization O(128), I(128)Dropout O(128), I(128)Densa O(1) Total de Parámetros: 474,245</p>

En la Tabla 4 se observa que los modelos de Copilot son todos ellos secuenciales, proponiendo en el modelo 4 una red neuronal convolucional en 2 dimensiones, mientras que DeepSeek propone arquitecturas más complejas a partir del segundo modelo. Para la propuesta de red neuronal híbrida hace una mezcla de red residual con concatenación, no es ningún modelo establecido como Resnet o DenseNet, la propuesta es algo diferente. En el modelo 3 de DeepSeek se propone el uso de redes neuronales convolucionales en una dimensión, utilizadas principalmente en problemas de series de tiempo. En todos los casos el prompt responde con una propuesta más compleja que la anterior, esto sugiere que la especificación del problema planteado se ha contextualizado de una mejor manera. Al proponer modelos más complejos la cantidad de parámetros totales es mayor en cada uno de los casos propuestos, desde luego que ello no garantiza un mejor desempeño y confiabilidad de la arquitectura para los fines de diagnóstico temprano o mejor clasificación de la información.

3. PRUEBAS Y RESULTADOS

Para validar los modelos producidos por las herramientas de IA y con ello obtener información que permita soportar la eficacia de cada uno, se determinó realizar el entrenamiento de los modelos por 10 veces y se reportaron las métricas de clasificación: precisión, sensibilidad y AUC-ROC (este último por sus siglas en Inglés de Area Under Curve: AUC y ROC: Receiver Operating Characteristic), en cada caso se promediaron los valores obtenidos en las 10 ejecuciones. La Tabla 6 nos muestra la información de las métricas en cada uno de los modelos generados por las herramientas de IA.

Tabla 6. Valores promedio obtenidos (10 iteraciones) de las métricas de evaluación en los modelos generados por las herramientas de IA.

Modelo	Copilot			DeepSeek		
	Precisión	Sensibilidad	AUC-ROC	Precisión	Sensibilidad	AUC-ROC
1	75.00%	65.45%	0.7958	69.80%	49.81%	0.7961
2	74.22%	61.27%	0.7656	80.93%	91.60%	0.8725
3	70.65%	60.36%	0.7339	89.81%	87.34%	0.9531
4	74.28%	62.36%	0.7709	51.27%	99.77%	0.8147

El análisis de los valores promedio indicados en la Tabla 6, permite identificar el modelo que puede ser el mejor candidato a utilizarse en una aplicación médica, siempre y cuando se logren realizar mejoras al modelo y entonces lograr aplicar un criterio de cero errores en la detección de enfermos positivos dentro de un contexto médico en que algún modelo pueda utilizarse.

El AUC es el área bajo la curva. Mientras que AUC-ROC se refiere al área bajo la curva característica operativa del receptor y mide el rendimiento de un modelo en todos sus posibles umbrales de decisión. Los valores de AUC que se obtienen para los modelos de Copilot indica que los modelos son aceptables pero que se pueden mejorar, mientras que los valores de AUC para los modelos generados por la herramienta DeepSeek, en específico el modelo 3 tiene un AUC que señala una discriminación excelente entre clases; los valores de AUC para los modelos 2 y 4 indican que los modelos son aceptables cuantitativamente, con cierto margen de mejora. El modelo 1 de DeepSeek con un valor AUC de 0.7961 es un modelo aceptable pero se puede mejorar.

En relación a la sensibilidad para un modelo en el área de la salud, primero es significativo señalar que es más importante maximizar la métrica de sensibilidad, para el estudio realizado en este trabajo significa que se debe evitar el mayor número de Falsos Negativos; esto quiere decir que un modelo de clasificación de datos, en este contexto está clasificando erróneamente los casos de diabéticos positivos como negativos, denominando estos últimos Falsos Negativos. En este estudio se refiere a la capacidad de los modelos para detectar los casos de diabéticos que son positivos; es decir, detectar si una persona es diabética. Con este criterio como elemento principal, al observar los valores de la Tabla 6 para la columna de los modelos obtenidos por Copilot y en lo que se refiere a la subcolumna de Sensibilidad, estos se encuentran en un rango del 60% al 65.5%, con esta información se establece que los modelos se encuentran fuera de lograr ser implementados en una aplicación para diagnóstico médico, debido a que entre el 40% y el 34.5% de los casos positivos (diabéticos) son erróneamente clasificados como negativos (falsos negativos); las implicaciones en el contexto médico es que un Falso Negativo es un paciente diabético no diagnosticado, este puede desarrollar complicaciones graves al ser diagnosticado erróneamente. Los resultados de sensibilidad obtenidos para los modelos generados por DeepSeek, el modelo 1 quedaría descartado al obtener el 49.81% de sensibilidad, los modelos 2

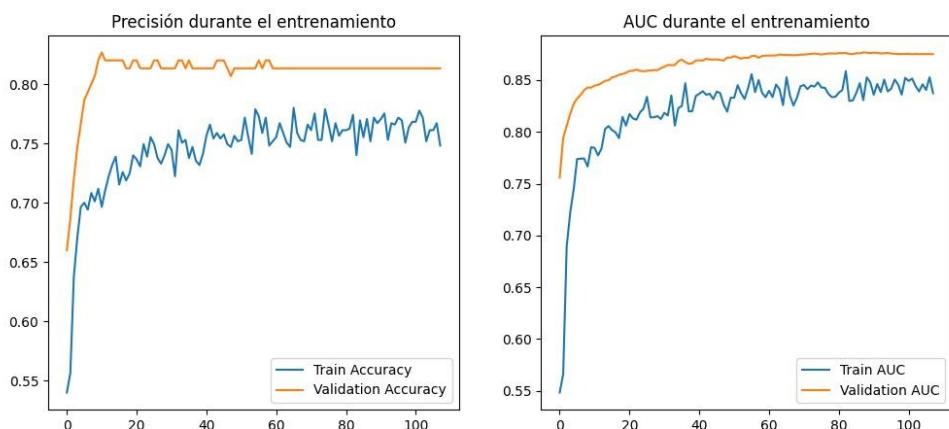
al 4 se encuentran dentro de un valor considerado deseable de sensibilidad (mayor al 85%) y factibles para ser implementados dentro de una aplicación médica, cabe señalar que lo ideal es obtener una sensibilidad del 100%, en aplicaciones médicas no deben aceptarse Falsos Negativos, para lograr acercarse, un primer paso que se sugiere es realizar un balanceo de clases en el conjunto de datos, el segundo paso es sintonizar parámetros del modelo y una tercera opción es proponer modelos con arquitecturas de redes neuronales más profundas.

Respecto a la precisión de los modelos obtenidos en Copilot, se tiene que en promedio los modelos clasifican correctamente 7 de cada 10 casos totales (tanto diabéticos como los no diabéticos), con estos valores surge el problema de que existe una alta tasa de errores, del 25% al 30% de predicciones incorrectas. Lograr una precisión mayor al 80% en los modelos sugeridos, son candidatos a realizarles mejoras para que puedan ser utilizados en aplicaciones médicas. Para el caso de los modelos obtenidos con DeepSeek, en el valor de precisión promedio obtenido para el modelo 4 (~51%), se puede establecer que se comportaría igual que un clasificador que proporciona el resultado por casualidad, para el caso de un clasificador binario que implementa al azar la obtención del resultado de clasificación es del 50%, por lo que debe desecharse en cualquier circunstancia. Para el modelo 1, el valor de precisión promedio que se obtuvo está cercano al 70%, inferior a los resultados de los modelos generados por Copilot, por lo que es un modelo que no es candidato para desempeñarse en alguna aplicación médica. El modelo 2 logró una precisión promedio de ~80%, con este valor este modelo es candidato a posibles mejoras y lograr un mejor desempeño general, al mismo tiempo se puede observar que el valor promedio de sensibilidad que se obtiene por este modelo es superior al 91%, lo que refuerza el argumento de mejorar el modelo y con ello ser un candidato aceptable en una aplicación médica; lo ideal es tener el mejor valor de sensibilidad posible aún cuando se sacrifique la precisión. Finalmente el valor promedio de precisión del modelo 3 es el mejor valor obtenido de todos los modelos, sin embargo el valor de sensibilidad es inferior al obtenido por el modelo 2, esto sugiere elegir como segunda opción al modelo 3 para mejorarlo y entonces ser un candidato aceptable para una aplicación médica.

En el desarrollo de un clasificador de redes neuronales para diagnóstico médico temprano (como detección de diabetes, cáncer o enfermedades

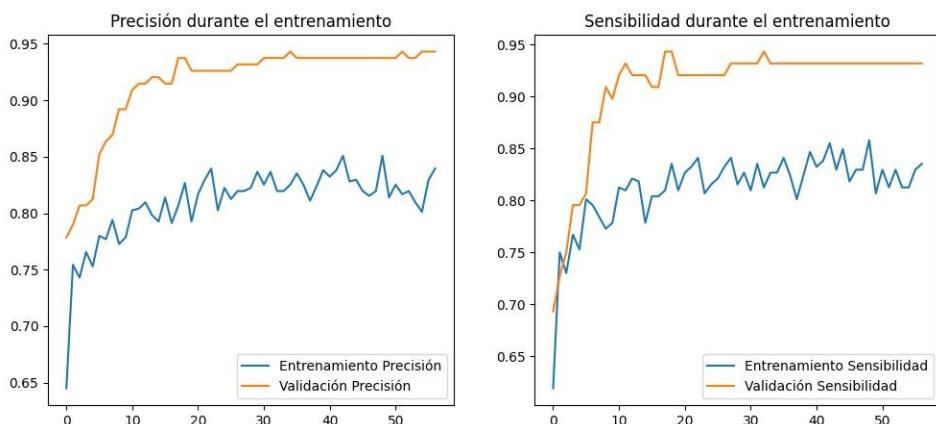
cardiovasculares), los elementos visuales que permite analizar la eficiencias y comportamientos de los modelos, son las curvas de entrenamiento y las métricas de evaluación (precisión, sensibilidad y AUC-ROC), estas herramientas son críticas para garantizar que el modelo sea confiable, seguro y clínicamente útil. Las Figuras 1 y 2 son ejemplos del comportamiento de entrenamiento de los modelos obtenidos.

Figura 1. Comportamiento del modelo 2 obtenido por la herramienta DeepSeek, respecto a la precisión y AUC en una de las ejecuciones de entrenamiento.



Las curvas de entrenamiento del modelo 2 en lo que se refiere a la precisión, en los datos de entrenamiento a partir de la época 50 en adelante no se mejora el entrenamiento y con los datos de validación obtiene una precisión por arriba del 80%, en aplicaciones médicas el modelo es aceptable pero puede mejorarse. Cuando el AUC se encuentra entre el 0.8 y 0.9 se debe mejorar el modelo hasta alcanzar un AUC superior al 0.9, en cuyo caso el modelo tendrá una capacidad satisfactoria para distinguir entre enfermos y sanos. La Figura 2 es un ejemplo de las curvas de entrenamiento del modelo 3, en ella se puede observar el comportamiento de la sensibilidad tanto para los datos de entrenamiento como de validación, cuando la sensibilidad en los datos de validación superan el 90%, es posible considerar la arquitectura para aplicaciones médicas, en donde el ideal es el 100%, que significa siempre detectar enfermos positivos o sobreentrenamiento del modelo.

Figura 2. Comportamiento del entrenamiento para el modelo 3, generado por la herramienta DeepSeek. Se muestran las curvas de sensibilidad y precisión.



4. CONCLUSIONES

La generación de modelos para la clasificación del conjunto de datos Pima Indians Diabetes mediante el uso de herramientas de IA (Copilot y DeepSeek) en este estudio determina que solamente el 25% de los modelos son candidatas a ser mejorados y con posibilidad de ser utilizados en una aplicación médica. El 100% de los modelos generados por la herramienta Copilot se desecharon al obtener valores promedio de sensibilidad inferiores al 66%. Solamente el 50% de los modelos generados por DeepSeek pueden considerarse candidatos a mejorarse para ser empleados en una aplicación médica al obtener una sensibilidad promedio superior al 87%. Con la información obtenida de las pruebas se logra señalar que DeepSeek proporciona modelos de clasificación que implementan redes neuronales más confiables y robustos respecto a aquellos generados por la herramienta Copilot, en el contexto del conjunto de datos Pima Indians Diabetes. Se concluye además que la evidencia empírica que se proporciona es un elemento que debe considerarse en el uso de modelos o arquitecturas de aprendizaje profundo generado por herramientas de IA. Se corrobora la importancia de contextualizar el texto escrito en el prompt en las herramientas de IA, la aproximación a una respuesta más correcta depende en gran medida a la forma en que se es más específico en la pregunta o comando. Es importante mencionar que la IA determinó el lenguaje

de programación y el tipo de algoritmos de aprendizaje profundo a implementar, en este caso Redes Neuronales Convolucionales, para realizar la tarea de clasificación binaria del conjunto de datos especificado. La evidencia obtenida nos permite indicar que como trabajo futuro se deben mejorar los modelos de redes neuronales obtenidos y con ello superar los resultados del estado del arte en la clasificación del conjunto de datos Pima Indians Diabetes.

REFERENCIAS

World Health Organization. (2023, April 13). Diabetes fact sheet. World Health Organization. URL: <https://www.who.int/news-room/fact-sheets/detail/diabetes>

Zheng, Y., Ley, S. H., & Hu, F. B. (2022). Global aetiology and epidemiology of type 2 diabetes mellitus and its complications. *Nature Reviews Endocrinology*, *18*(2), 88-98. <https://doi.org/10.1038/s41574-021-00608-9>

Brinker, T. J., Hekler, A., Enk, A. H., Klode, J., Hauschild, A., Berking, C., ... & Schrüfer, P. (2019). Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *European Journal of Cancer*, 113, 47-54.

Sharma, H., Tatiya, M., Aswal, U. S., Laxminarayananamma, K., Tripathi, N., & Singh, D. (2023, September). Real-Time Patient Monitoring Using Deep Learning For Medical Diagnosis. In *2023 6th International Conference on Contemporary Computing and Informatics (IC3I)* (Vol. 6, pp. 1629-1634). IEEE.

Agarwal, S., Rengarajan, V., Chinnakotla, B., Padmanaban, G., & Mani, S. (2022). AI-based screening for prediabetes in low-resource settings. *Journal of Diabetes Science and Technology*, *16*(6), 1386-1394. <https://doi.org/10.1177/19322968211053381>

Patel, R., Krishnan, A., Sharma, P., Mehta, N., & Jain, S. (2023). Cost-effectiveness of artificial intelligence triage for diabetes screening in Indian primary care: A cluster randomized trial. *BMJ Innovations*, *9*(2), 78-89. <https://doi.org/10.1136/bmjjinnov-2022-001234>

UNICEF. (2023). Leveraging AI and mobile health technologies for non-communicable disease screening in low-resource settings: Case studies from Sub-Saharan Africa and South Asia. Fondo de las Naciones Unidas para la Infancia. URL: <https://www.unicef.org/innovation/ai-health-ncd-screening> DOI: 10.5316/unicef.2023.healthtech01

Shakya, R., Vadiee, F., & Khalil, M. (2025, April). A Showdown of ChatGPT vs DeepSeek in Solving Programming Tasks. In *2025 International Conference on New Trends in Computing Sciences (ICTCS)* (pp. 413-418). IEEE.

Yadav, S., Qureshi, A. M., Kaushik, A., Sharma, S., Loughran, R., Kazhuparambil, S., ... & Lillis, D. (2025). From Idea to Implementation: Evaluating the Influence of Large Language Models in Software Development--An Opinion Paper. *arXiv preprint arXiv:2503.07450*.

Tian, Z., Chen, J., & Zhang, X. (2023). Test-case-driven programming understanding in large language models for better code generation. arXiv preprint arXiv:2309.16120.

Wagner, G., Müller, T., Schmidt, A., Zhang, L., & Kim, H. (2024). Large language models in biomedicine: A comparative study of ChatGPT and DeepSeek for clinical text analysis and diagnostic support. *Nature Computational Science*, *4*(2), 112-125. <https://doi.org/10.1038/s43588-024-00605-8>

Rajeswari, S., & Suthendran, K. (2022). Diabetes prediction using machine learning: A systematic review. *Computers in Biology and Medicine*, *141*, 105133. DOI: 10.1016/j.combiomed.2021.105133

GitHub. (2023). GitHub Copilot documentation: Your AI pair programmer. GitHub, Inc. URL: <https://docs.github.com/en/copilot> Fecha de consulta: 12 de abril de 2025.

DeepSeek. (2023). DeepSeek Chat: Advanced AI assistant for technical and scientific applications [Computer software]. DeepSeek Inc. URL: <https://www.deepseek.com>

Puspitasari, F. D., Zhang, C., Dam, S. K., Zhang, M., Kim, T. H., Hong, C. S., ... & Yang, Y. (2025). Deepseek models: A comprehensive survey of methods and applications. *Authorea Preprints*.

Wagner, G., Müller, T., Schmidt, A., Zhang, L., & Kim, H. (2024). Large language models in biomedicine: A comparative study of ChatGPT and DeepSeek for clinical text analysis and diagnostic support. *Nature Computational Science*, 4(2), 112-125. <https://doi.org/10.1038/s43588-024-00605-8>

SOBRE O ORGANIZADOR

Xosé Somoza Medina (1969, Ourense, España) Licenciado con Grado y premio extraordinario en Geografía e Historia por la Universidad de Santiago de Compostela (1994). Doctor en Geografía e Historia por la misma universidad (2001) y premio extraordinario de doctorado por su Tesis “Desarrollo urbano en Ourense 1895-2000”. Profesor Titular en la Universidad de León, donde imparte clases desde 1997. En la Universidad de León fue Director del Departamento de Geografía entre 2004 y 2008 y Director Académico de la Escuela de Turismo entre 2005 y 2008. Entre 2008 y 2009 ejerció como Director del Centro de Innovación y Servicios de la Xunta de Galicia en Ferrol. Entre 2007 y 2009 fue vocal del comité “Monitoring cities of tomorrow” de la Unión Geográfica Internacional. En 2012 fue Director General de Rehabilitación Urbana del Ayuntamiento de Ourense y ha sido vocal del Consejo Rector del Instituto Ourensano de Desarrollo Local entre 2011 y 2015. Ha participado en diversos proyectos y contratos de investigación, en algunos de ellos como investigador principal, con temática relacionada con la planificación urbana, la ordenación del territorio, las nuevas tecnologías de la información geográfica, el turismo o las cuestiones demográficas. Autor de más de 100 publicaciones relacionadas con sus líneas de investigación preferentes: urbanismo, turismo, gobernanza, desarrollo, demografía, globalización y ordenación del territorio. Sus contribuciones científicas más importantes se refieren a la geografía urbana de las ciudades medias, la crisis del medio rural y sus posibilidades de desarrollo, la evolución del turismo cultural como generador de transformaciones territoriales y más recientemente las posibilidades de reindustrialización de Europa ante una nueva etapa posglobalización. Ha participado como docente en masters y cursos de especialización universitaria en Brasil, Bolivia, Colombia, Paraguay y Venezuela y como docente invitado en la convocatoria Erasmus en universidades de Bulgaria (Sofia), Rumanía (Bucarest) y Portugal (Porto, Guimarães, Coimbra, Aveiro y Lisboa). Ha sido evaluador de proyectos de investigación en la Agencia Estatal de Investigación de España y en la Organización de Estados Iberoamericanos (OEI). Como experto europeo en Geografía ha participado en reuniones de la Comisión Europea en Italia y Bélgica. Impulsor y primer coordinador del proyecto europeo URBACT, “come Ourense”, dentro del Programa de la Unión Europea “Sostenibilidad alimentaria en comunidades urbanas” (2012-2014). Dentro de la experiencia en organización de actividades de I+D+i se pueden destacar la organización de diferentes reuniones científicas desarrolladas dentro de la Asociación de Geógrafos Españoles (en 2002, 2004, 2012 y 2018).

ÍNDICE REMISSIVO

A

Análisis de variogramas 40
Arquitectura 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14, 15, 61, 62, 65, 68

C

Ceramic applicator 82, 83, 84, 87, 88, 89, 90, 91
Chícharo (*Pisum sativum*) 72, 73, 74, 76
Comb applicator 82, 83, 84, 85, 86, 87, 89, 90, 91
Conservación genética 17, 21, 24
Copilot 54, 55, 58, 60, 62, 63, 64, 65, 66, 67, 69, 71
Curvas de secado 73, 76, 78, 79

D

DeepSeek 54, 55, 58, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71
Diabetes Mellitus tipo 2 54, 55, 56

E

Eficiencia productiva 27
Energía 73, 77, 79, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108

F

Fotovoltaica 93, 94, 95, 101, 102, 103, 104, 105, 106, 107, 108

I

IA en salud 55
Innovación agropecuaria 27

K

Kriging ordinario 40, 45, 46, 49, 50

L

LLMs 55, 57, 61

M

Mapeo de combustibles 40, 49

Microwave 82, 87

P

Periodo formativo 1, 2, 6, 8, 11, 14, 15

Producción extensiva 17, 19, 20, 23, 25

R

Redes Neuronales Convolucionales 55, 65, 70

Región pampeana 27, 28, 29, 30, 38, 39

Renovable 94, 95, 102, 105, 106, 107

S

Secador de lecho fluidizado 72, 73, 81

Sistemas productivos 23, 27

Slow wave 82, 83, 92

Soil microwave treatment 82

T

Tecnología constructiva 1, 2, 3, 4, 5, 7, 8, 11, 13, 14, 15, 16

Tecnología frutícola 27

V

Valor agregado comunitario 17

Velocidad mínima de fluidización 73, 75, 77, 78

Vivienda 16, 94, 96, 97, 102, 108

