

**Predicción temprana del riesgo de desarrollar Diabetes Mellitus tipo II en población adulta
mediante modelos de aprendizaje automático**

Jorge Armando Granobles Carvajal

Asesor

Dr. Luis Ángel Anillo Arrieta

Universidad Nacional Abierta y a Distancia UNAD
Escuela de Ciencias Básicas, Tecnología e Ingeniería ECBTI
Especialización en Ciencia de Datos y Analítica

2026

Resumen

El presente proyecto de investigación aborda la detección tardía de la Diabetes Mellitus Tipo II (DM2), un desafío crítico de salud pública debido a su progresión asintomática y las limitaciones de los métodos convencionales. La problemática radica en que los enfoques estadísticos tradicionales a menudo omiten patrones complejos y no lineales en los datos clínicos; por ello, se propone el aprendizaje automático como solución estratégica, dada su capacidad superior para procesar variables multidimensionales y detectar riesgos en etapas preclínicas. Bajo el marco metodológico CRISP-DM, el objetivo general fue evaluar y validar la capacidad predictiva de modelos de clasificación para el diagnóstico temprano en población adulta. La metodología integró el análisis de datos biométricos, hábitos de vida y marcadores paraclínicos, destacando la integración de la Glucemia Pre-prandial y la Hemoglobina Glicosilada como ejes del criterio diagnóstico, junto con el peso predictivo de variables metabólicas como los Triglicéridos. Se aplicaron técnicas de ingeniería de características, imputación multivariada y balanceo de datos mediante SMOTE para corregir el sesgo de la muestra original. Los resultados obtenidos demuestran la alta eficacia de los algoritmos supervisados, donde el modelo XGBoost alcanzó una exactitud del 99.09% y una sensibilidad (Recall) del 100% en la identificación de pacientes con riesgo alto. Este enfoque permite transformar la práctica sanitaria de un modelo reactivo a uno predictivo, logrando una estratificación del riesgo precisa que sustenta intervenciones tempranas y personalizadas para mitigar el impacto de la enfermedad.

Palabras clave: Diabetes Mellitus Tipo II, Aprendizaje Automático, CRISP-DM, XGBoost, Riesgo Metabólico.

Abstract

This research project addresses the late detection of Type II Diabetes Mellitus (T2DM), a critical public health challenge due to its asymptomatic progression and the limitations of conventional screening methods. The core problem lies in the fact that traditional statistical approaches often overlook complex and non-linear patterns within clinical data; therefore, machine learning is proposed as a strategic solution, given its superior capacity to process multidimensional variables and detect risks in preclinical stages. Following the CRISP-DM methodological framework, the general objective was to evaluate and validate the predictive performance of classification models for early diagnosis in the adult population. The methodology integrated the analysis of biometric data, lifestyle habits, and paraclinical markers, highlighting the integration of Pre-prandial Glycemia and Glycated Hemoglobin (HbA1c) as the pillars of the diagnostic criteria, alongside the predictive weight of metabolic variables such as Triglycerides. Feature engineering, multivariate imputation, and data balancing techniques using SMOTE were applied to correct the bias of the original sample. The results demonstrate the high efficacy of supervised algorithms, where the XGBoost model achieved an accuracy of 99.09% and a sensitivity (Recall) of 100% in identifying high-risk patients. This approach enables a transformation of healthcare practice from a reactive to a predictive model, achieving precise risk stratification that supports early and personalized interventions to mitigate the disease's impact.

Keywords: Type II Diabetes Mellitus, Machine Learning, CRISP-DM, XGBoost, Metabolic Risk

Tabla de Contenido

Introducción	9
Planteamiento del Problema	10
Pregunta Problema.....	11
Justificación	12
Objetivos	13
Objetivo General.....	13
Objetivos Específicos	13
Marco Referencial.....	14
Conceptos Fundamentales	14
Fundamentación Teórica del Problema	15
Metodología	16
Enfoque Metodológico	16
Fuente de Información y Población de Estudio.....	17
Fases Metodológicas Alineadas con los Objetivos Específicos	19
Fase 1 Identificación, Preparación y Análisis de Variables Clínicas Relevantes.....	19
Fase 2 Aplicación de Modelos de Aprendizaje Automático para la Estimación Individual del Riesgo	24
Fase 3 Evaluación, Comparación, Calibración y Selección del Modelo Predictivo	28
Resultados	32
Identificación y Análisis de Variables Clínicas Relevantes Asociadas al Riesgo de Desarrollar Diabetes Mellitus Tipo II.....	32
Variables Clínicas Analizadas	33

Relación entre Variables Clínicas y Clasificación del Riesgo Metabólico	34
Variables con Mayor Peso en la Selección de Características	37
Interpretación Clínica de las Variables Identificadas	41
Aplicación de Modelos de Aprendizaje Automático para la Estimación del Riesgo Individual de Desarrollar Diabetes Mellitus Tipo II.....	42
Distribución de las Categorías de Riesgo en los Conjuntos Analizados	43
Resultados Globales de Clasificación Según Escenario y Modelo	43
Comportamiento de los Modelos en el Escenario Completo.....	45
Matrices de Confusión del Escenario Completo	47
Comportamiento de los Modelos en el Escenario Preventivo	48
Matrices de Confusión del Escenario Preventivo	50
Interpretación General de la Aplicación de Modelos	51
Evaluación, Comparación y Selección del Modelo Predictivo con Mejor Desempeño Para la Estimación del Riesgo de Desarrollar Diabetes Mellitus Tipo II.....	53
Capacidad Discriminativa de los Modelos Según Escenario	53
Desempeño Discriminativo de XGBoost por Categoría de Riesgo.....	57
Evaluación de Calibración Probabilística.....	60
Interpretación Crítica de los Escenarios Evaluados.....	63
Selección del Modelo con Mejor Desempeño	64
Conclusiones	66
Recomendaciones	71
Fortalecimiento de la Calidad y Estandarización de los Datos Clínicos	71
Diferenciación entre Modelos Diagnósticos y Modelos Preventivos.....	71

Prevención de Fuga de Información y Validación del Diseño Analítico	72
Fortalecimiento de la Captura de Variables Conductuales y Antecedentes Clínicos.....	73
Validación Externa y Fortalecimiento de la Generalización del Modelo.....	74
Integración Institucional de Herramientas Predictivas y Tableros de Analítica Clínica	74
Desarrollo Futuro de Modelos Predictivos Preventivos	75
Implementación Progresiva y Supervisada de Inteligencia Artificial en Salud	75
Referencias Bibliográficas	77

Lista de Tablas

Tabla 1 <i>Caracterización Inicial de la Población</i>	19
Tabla 2 <i>Matriz de Variables Clínicas Utilizadas</i>	22
Tabla 3 <i>Clasificación del Riesgo Metabólico</i>	24
Tabla 4 <i>Escenarios de Modelado Aplicados</i>	25
Tabla 5 <i>Modelos de Clasificación Supervisada</i>	26
Tabla 6 <i>Variables Clínicas Evaluadas Según Dimensión Analítica</i>	33
Tabla 7 <i>Ranking de Variables Clínicas Relevantes Asociadas al Riesgo Metabólico Mediante Información Mutua</i>	38
Tabla 8 <i>Desempeño Global de los Modelos Según Escenario de Análisis</i>	44
Tabla 9 <i>Resultados AUC-ROC por Modelo y Escenario</i>	54
Tabla 10 <i>AUC-ROC por Clase para XGBoost Según Escenario</i>	57
Tabla 11 <i>Métricas de Calibración por Modelo y Escenario</i>	60

Lista de Figuras

Figura 1 <i>Estándar CRISP-DM</i>	17
Figura 2 <i>Identificación de Valores Atípicos (Outliers)</i>	21
Figura 3 <i>Arquitectura de Modelos de Aprendizaje Automático Aplicados</i>	28
Figura 4 <i>Métricas de Clasificación Utilizadas</i>	29
Figura 5 <i>Matriz de Correlación Entre Variables Clínicas Relevantes y Clasificación del Riesgo Metabólico</i>	36
Figura 6 <i>Importancia de Variables Clínicas Asociadas al Riesgo Metabólico Técnica Información Mutua</i>	40
Figura 7 <i>Matrices de Confusión de los Modelos Aplicados en el Escenario Completo</i>	47
Figura 8 <i>Matrices de Confusión de los Modelos Aplicados en el Escenario Preventivo</i>	50
Figura 9 <i>Curvas ROC Comparativas del Modelo Completo con Glucosa y HbA1c</i>	55
Figura 10 <i>Curvas ROC Comparativas del Modelo Preventivo sin Glucosa y HbA1c</i>	56
Figura 11 <i>Curvas ROC por Clase para XGBoost en el Modelo Completo</i>	58
Figura 12 <i>Curvas ROC por Clase para XGBoost en el Modelo Preventivo</i>	59
Figura 13 <i>Curva de Calibración para Riesgo Alto en el Modelo Completo</i>	61
Figura 14 <i>Curva de Calibración para Riesgo Alto en el Modelo Preventivo</i>	62

Introducción

La Diabetes Mellitus tipo II (DM2) constituye uno de los desafíos más críticos para la salud pública contemporánea, con un impacto profundo en la morbilidad y mortalidad global (Sun et al., 2022). En regiones como Colombia, el aumento sostenido de esta patología se traduce en complicaciones graves que comprometen la calidad de vida del paciente y la sostenibilidad financiera de los sistemas sanitarios (Granados Carrillo et al., 2024). Históricamente, la identificación de esta enfermedad ha sido reactiva, detectándose en etapas avanzadas debido a su carácter asintomático y a que los métodos de tamizaje tradicionales analizan variables de forma aislada, perdiendo la visión de conjunto (WHO, 2016; Boutilier et al., 2021).

Ante este escenario, la Ciencia de Datos se presenta como una disciplina capaz de transformar la atención en salud. A través de modelos de clasificación supervisada, es posible procesar volúmenes masivos de información y capturar interacciones no lineales que los métodos convencionales omiten (Larose & Larose, 2019). Este proyecto de investigación se propone desarrollar un modelo de aprendizaje automático bajo el estándar metodológico CRISP-DM. El objetivo es integrar datos clínicos, biométricos y de estilo de vida para predecir tempranamente el riesgo de DM2, proporcionando una herramienta que fortalezca la medicina preventiva y mejore la toma de decisiones clínicas basadas en evidencia.

Planteamiento del Problema

La detección tardía de la Diabetes Mellitus Tipo II (DM2) representa uno de los desafíos más críticos para el sistema de salud, dado que su fase asintomática suele pasar desapercibida hasta que se manifiestan complicaciones irreversibles (WHO, 2025; Sun et al., 2022). A pesar de la existencia de guías clínicas, en la práctica diaria los profesionales de la salud disponen de información fragmentada y tiempos de consulta limitados, lo que dificulta la integración manual de múltiples variables para identificar el riesgo individual con precisión. Esta limitación es especialmente grave en entornos de recursos limitados, donde la falta de herramientas de estratificación efectivas retrasa las intervenciones preventivas (Granados Carrillo et al., 2024; Boutilier et al., 2021).

La evidencia científica indica que el riesgo de desarrollar DM2 no depende de factores aislados, sino de interacciones complejas y no lineales entre variables clínicas, antecedentes de comorbilidades como la hipertensión arterial (HTA), y determinantes conductuales (Young-Hyman et al., 2016). Los métodos de identificación tradicionales a menudo evalúan estos factores de manera segregada por ejemplo, considerando la presión arterial y la glucosa de forma independiente fallando en capturar la sinergia patogénica real entre la hipertensión y el riesgo metabólico (Hasan, 2025). Asimismo, los modelos estadísticos clásicos suelen presentar limitaciones al procesar estas variables correlacionadas, no logrando interpretar la complejidad del riesgo con la misma eficacia que los nuevos enfoques computacionales (Sierra Santos, 2025).

En contraste, la Ciencia de Datos ofrece alternativas superiores. Estudios recientes demuestran que los modelos de aprendizaje automático (Machine Learning) pueden procesar grandes volúmenes de datos heterogéneos para descubrir patrones ocultos que los enfoques estadísticos convencionales no detectan, mejorando significativamente la precisión en la

detección temprana (Alkhaldi et al., 2025; Larose & Larose, 2019). Por tanto, existe la necesidad de diseñar una solución bajo el estándar CRISP-DM que consolide datos clínicos y de estilo de vida, utilizando la historia de hipertensión como un predictor clave, para anticiparse al desarrollo de la DM2 y fortalecer la toma de decisiones clínicas con una herramienta de alta sensibilidad (Delgado et al., 2025; Boutilier et al., 2021).

Pregunta Problema

¿En qué medida la aplicación de modelos de aprendizaje automático permite optimizar la detección temprana del riesgo de desarrollar Diabetes Mellitus Tipo II, frente a los métodos de identificación convencionales?

Justificación

La ejecución de este proyecto se justifica por la necesidad imperativa de optimizar los procesos de estratificación de riesgo en la población adulta. La detección temprana de la DM2 no solo representa un beneficio directo para el paciente al evitar daños irreversibles en órganos blancos, sino que también es una estrategia de eficiencia económica para las instituciones de salud, reduciendo los altos costos asociados al tratamiento de complicaciones crónicas (Sun et al., 2022). El uso de algoritmos avanzados permite mitigar las barreras de tiempo e información fragmentada que enfrentan los profesionales de la salud en la consulta diaria (Boutilier et al., 2021).

Asimismo, el valor de esta investigación reside en la integración de técnicas de aprendizaje automático con información clínica real proveniente de una institución de salud de primer nivel, aportando una herramienta de soporte analítico que complementa las guías de práctica clínica convencionales y favorece procesos de estratificación de riesgo más granulares y eficientes en entornos de atención primaria. Al validar la capacidad predictiva de modelos multidimensionales, este trabajo no solo cumple con los requerimientos académicos de la especialización, sino que aporta un prototipo funcional que sustenta intervenciones personalizadas. En última instancia, la implementación de modelos de aprendizaje automático en este ámbito se traduce en una práctica sanitaria más proactiva, precisa y centrada en la prevención efectiva (Larose & Larose, 2019).

Objetivos

Objetivo General

Evaluar la capacidad predictiva de modelos de aprendizaje automático para la estratificación temprana del riesgo de desarrollar diabetes mellitus tipo II en población adulta.

Objetivos Específicos

Identificar y analizar las variables clínicas relevantes asociadas al riesgo de desarrollar diabetes mellitus tipo II, mediante análisis exploratorio y técnicas estadísticas de selección de características, con el fin de definir los determinantes de riesgo más influyentes del modelo.

Aplicar modelos de aprendizaje automático para estimar el riesgo individual de desarrollar diabetes mellitus tipo II en población adulta, con el fin de apoyar las acciones de prevención y diagnóstico temprano.

Evaluar y comparar el desempeño de los modelos mediante métricas de discriminación y calibración, con el fin de seleccionar el modelo con mejor rendimiento para apoyar acciones de prevención y diagnóstico temprano.

Marco Referencial

El marco referencial integra la epidemiología de la Diabetes Mellitus Tipo II (DM2), la Ciencia de Datos aplicada a la medicina predictiva y la Teoría del Cambio Comportamental, pilares necesarios para abordar la detección temprana desde una perspectiva analítica y psicosocial (Granados Carrillo et al., 2024; Sun et al., 2022).

Conceptos Fundamentales

Aprendizaje Automático (Machine Learning) y Predicción: Rama de la inteligencia artificial que permite identificar patrones complejos en datos sin programación explícita. En el contexto clínico, algoritmos de clasificación supervisada como Random Forest y XGBoost han demostrado superar a la estadística tradicional al capturar interacciones no lineales entre variables, mejorando la estratificación del riesgo individual (Boutilier et al., 2021; Larose & Larose, 2019).

Sinergia con Hipertensión Arterial (HTA): La HTA se integra en este estudio como un factor de riesgo predictor crítico. Su presencia actúa sinérgicamente con la disglucemia, multiplicando el riesgo cardiovascular y acelerando el daño microvascular, lo que la convierte en una variable de entrada indispensable para robustecer la capacidad predictiva del modelo (Sun et al., 2022).

Factores Psicosociales y Adherencia Terapéutica: La adherencia se entiende como el grado en que la conducta del paciente coincide con las recomendaciones del profesional de salud. Factores como el estrés, el malestar emocional (diabetes distress) o las dificultades de afrontamiento se asocian directamente con un menor control metabólico (HbA1c) y baja adherencia. Desde la teoría del empowerment, empoderar al paciente influye positivamente en la

adherencia y en los resultados en salud, siendo variables que modulan el riesgo real de complicaciones (Young-Hyman et al., 2016; Anderson & Funnell, 2000).

Fundamentación Teórica del Problema

La detección tardía de la DM2 puede interpretarse como una falla de los sistemas de salud para integrar la vasta cantidad de datos que generan los pacientes. Tradicionalmente, el diagnóstico se ha basado en una "perspectiva de silos", donde se evalúan factores como el Índice de Masa Corporal (IMC), la presión arterial o la glucosa de manera aislada, sin capturar sus interdependencias complejas (Hasan, 2025). Esta limitación reduce la sensibilidad de los métodos de identificación convencionales, dejando sin diagnosticar a una gran proporción de pacientes en fase asintomática (Boutilier et al., 2021).

La Medicina Predictiva plantea que el riesgo de DM2 no es lineal, sino multidimensional. Estudios recientes indican que la integración de marcadores bioquímicos (como enzimas hepáticas y perfil lipídico) con datos antropométricos y antecedentes de hipertensión mediante técnicas de Machine Learning permite modelar el riesgo con una precisión significativamente mayor ($AUC > 0.90$) que los puntajes de riesgo tradicionales (Alkhaldi et al., 2025; Larose & Larose, 2019).

En este contexto, la aplicación de la metodología CRISP-DM facilita la estructuración de estos datos heterogéneos (clínicos, biométricos y de hábitos), permitiendo desarrollar una herramienta de soporte a la decisión clínica que no solo estime el riesgo de desarrollar la enfermedad, sino que identifique los determinantes más influyentes para cada individuo, favoreciendo así una estrategia de prevención personalizada y costo-efectiva (Sierra Santos, 2025; Boutilier et al., 2021).

Metodología

Enfoque Metodológico

El presente proyecto adoptó un enfoque cuantitativo, retrospectivo y analítico, orientado al estudio de patrones clínicos, paraclínicos, biométricos, cardiovasculares y de estilo de vida registrados en historias clínicas de población adulta. El propósito metodológico fue estructurar una base de datos clínica que permitiera identificar variables asociadas al riesgo de desarrollar Diabetes Mellitus Tipo II y, posteriormente, aplicar modelos de aprendizaje automático para estimar dicho riesgo de manera individual.

Para el desarrollo del proyecto se empleó el estándar CRISP-DM (Cross-Industry Standard Process for Data Mining), debido a su utilidad en proyectos de analítica predictiva en salud. Este marco permitió organizar el proceso desde la comprensión del contexto clínico e institucional, pasando por la preparación y transformación de los datos, hasta el modelado, evaluación, calibración e interpretación de los resultados.

Figura 1

Estándar CRISP-DM

Nota. La figura representa las fases del estándar CRISP-DM utilizadas como referencia metodológica del proyecto. CRISP-DM significa *Cross-Industry Standard Process for Data Mining*.

Fuente de Información y Población de Estudio

El presente estudio corresponde a una investigación cuantitativa, retrospectiva y analítica, basada en fuente secundaria de información clínica. La investigación no implicó intervención

directa sobre pacientes, sino el análisis de información clínica anonimizada con fines académicos, analíticos y predictivos.

La fuente primaria de información utilizada en el presente estudio fue la base de datos clínica del programa institucional “Adultez”, perteneciente a los servicios de Promoción y Mantenimiento de la Salud de la E.S.E. Hospital San Roque de Guacarí, institución pública de primer nivel de atención ubicada en el municipio de Guacarí, departamento del Valle del Cauca.

La E.S.E. Hospital San Roque de Guacarí presta servicios ambulatorios, apoyos diagnósticos, urgencias, hospitalización y atención de partos a población afiliada a diferentes Entidades Promotoras de Salud (EPS) que operan en la región. Actualmente, la institución cuenta con una población contratada de 16.703 usuarios, de los cuales 15.086 pertenecen al régimen subsidiado y 1.127 al régimen contributivo.

De acuerdo con información demográfica reportada por el Departamento Administrativo Nacional de Estadística (DANE), el municipio de Guacarí cuenta con una población aproximada de 35.691 habitantes, constituyéndose como un contexto representativo para el análisis de enfermedades metabólicas y factores de riesgo asociados a la población adulta.

El programa “Adultez”, seleccionado como fuente de extracción de la muestra para el desarrollo del proyecto, tiene actualmente 10.221 pacientes inscritos, pertenecientes al grupo etario comprendido entre los 29 y 59 años. Del total de usuarios registrados en el programa, 6.866 corresponden al sexo femenino y 3.355 al sexo masculino.

Previo al acceso a la información, se realizaron reuniones técnicas con las directivas de la entidad de salud, en las cuales se presentó el alcance académico del proyecto, los objetivos de investigación y el uso previsto de los datos clínicos. Posteriormente, se obtuvo el aval

institucional para el uso de la información bajo criterios de confidencialidad, anonimización y protección de datos sensibles.

La base consolidada utilizada para el estudio estuvo conformada por 2.750 historias clínicas anonimizadas, integrando variables demográficas, antropométricas, paraclínicas, cardiovasculares y de hábitos de vida relevantes para la caracterización del riesgo de desarrollar Diabetes Mellitus Tipo II en población adulta.

Tabla 1

Caracterización Inicial de la Población

Variable	Descripción
Fuente de datos	Programa institucional “Adultez”
Tipo de información	Historias clínicas anonimizadas
Población	Adultos atendidos en programa de promoción y mantenimiento de la salud
Total de registros	2.750
Tipo de estudio	Cuantitativo, retrospectivo y analítico

Nota. Características generales de la fuente de datos y de la población analizada en el estudio, correspondiente a registros clínicos del programa institucional “Adultez”.

Fases Metodológicas Alineadas con los Objetivos Específicos

La metodología se organizó en tres fases principales, cada una vinculada directamente con los objetivos específicos del proyecto.

Fase 1 Identificación, Preparación y Análisis de Variables Clínicas Relevantes

Relación con el Objetivo Específico 1. Esta fase respondió al primer objetivo específico, orientado a identificar y analizar las variables clínicas relevantes asociadas al riesgo de

desarrollar Diabetes Mellitus Tipo II. Integró las etapas de comprensión del negocio, comprensión del dato y preparación de datos del estándar CRISP-DM.

Inicialmente, se realizó la revisión estructural de la base de datos, identificando tipos de variables, formatos, registros incompletos, posibles duplicados, inconsistencias y valores extremos. La base incluyó variables demográficas, antropométricas, paraclínicas, cardiovasculares, metabólicas, conductuales y escalas clínicas de riesgo.

Las variables categóricas, como sexo, actividad física, consumo de alcohol, consumo de cigarrillo, dislipidemia e hipertensión arterial, fueron transformadas a formato numérico binario para garantizar su compatibilidad con los análisis estadísticos y los modelos de clasificación supervisada.

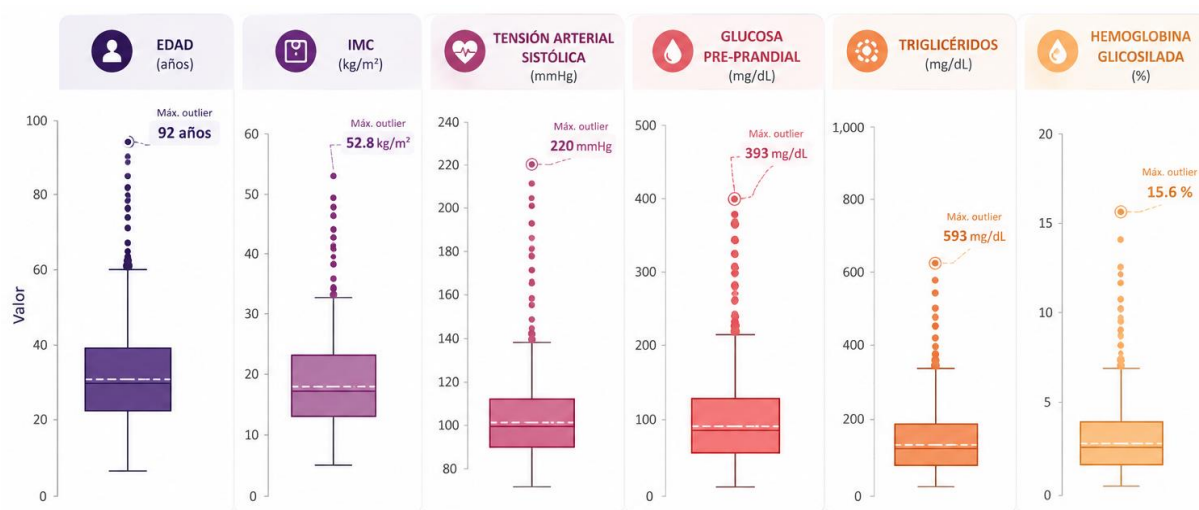
El análisis exploratorio de datos incluyó revisión de distribuciones, medidas descriptivas, diagramas de caja, mapas de calor de valores faltantes y análisis de asociación entre variables. Este proceso permitió caracterizar el comportamiento general de la información clínica y establecer la estructura analítica de las variables incluidas en el estudio.

Durante la auditoría de calidad de datos se identificaron vacíos principalmente en variables paraclínicas de laboratorio. Para su tratamiento, la imputación multivariada mediante IterativeImputer de Scikit-learn se realizó siguiendo el principio de no contaminación del conjunto de prueba. El modelo de imputación fue ajustado exclusivamente sobre el conjunto de entrenamiento, correspondiente al 80 % de los datos, y posteriormente aplicado para transformar tanto el conjunto de entrenamiento como el conjunto de prueba, garantizando así que la información del conjunto de prueba no influyera en el proceso de estimación de valores ausentes. Esta estrategia fue seleccionada por conservar mejor la estructura relacional de las variables clínicas frente a métodos simples de imputación, como media o mediana.

Los valores atípicos fueron evaluados mediante diagramas de caja y análisis de plausibilidad clínica. Aquellos valores extremos compatibles con escenarios reales de descompensación metabólica fueron conservados, debido a que representan pacientes con alteraciones relevantes para la caracterización del riesgo metabólico y para las fases posteriores de modelado predictivo.

Figura 2

Identificación de Valores Atípicos (Outliers)



Nota. Diagramas de caja utilizados para identificar valores atípicos en variables clínicas críticas asociadas al riesgo metabólico. Los valores extremos fueron evaluados según su compatibilidad clínica y conservados cuando correspondían a escenarios reales de descompensación metabólica o riesgo agudo. IMC = Índice de Masa Corporal.

Posteriormente, se consolidó una matriz analítica final compuesta por variables agrupadas en dimensiones clínicas: glucémicas, lipídicas, antropométricas, cardiovasculares, metabólicas complementarias, conductuales, demográficas y escalas clínicas de riesgo.

Tabla 2*Matriz de Variables Clínicas Utilizadas*

Grupo de Variables	Variables Incluidas	Utilidad dentro del análisis
VARIABLES glucémicas	Glucosa Pre-prandial, Hemoglobina Glicosilada (HbA1c)	Identificación y clasificación del riesgo metabólico
VARIABLES lipídicas	Triglicéridos, Colesterol HDL, Colesterol LDL, Colesterol VLDL	Evaluación del perfil metabólico y cardiovascular
VARIABLES antropométricas	Índice de Masa Corporal (IMC)	Caracterización del estado nutricional y obesidad
VARIABLES demográficas	Edad, Sexo	Caracterización poblacional y análisis de susceptibilidad
VARIABLES cardiovasculares	Tensión Arterial Sistólica, Tensión Arterial Diastólica, Hipertensión Arterial	Evaluación del riesgo cardiovascular asociado
VARIABLES metabólicas complementarias	Creatinina, Dislipidemia	Evaluación de alteraciones metabólicas y renales asociadas
VARIABLES conductuales y hábitos de vida	Actividad Física, Consumo de Alcohol, Consumo de Cigarrillo	Identificación de factores modificables de riesgo
Escalas clínicas de riesgo	Resultado FINDRISC, Resultado Framingham	Estimación complementaria del riesgo cardiometabólico

Nota. Variables clínicas incluidas en el análisis, organizadas por grupos según su naturaleza y utilidad para la caracterización multidimensional del riesgo metabólico en la población estudiada. HbA1c = hemoglobina glicosilada; HDL = lipoproteína de alta densidad; LDL = lipoproteína de baja densidad; VLDL = lipoproteína de muy baja densidad; IMC = índice de

masa corporal; FINDRISC = *Finnish Diabetes Risk Score*, escala finlandesa de riesgo de diabetes. Framingham corresponde a una escala de estimación del riesgo cardiovascular.

Para el análisis de asociación entre variables se aplicó el *coeficiente de correlación de Pearson*, con el propósito de identificar la fuerza y dirección de las relaciones lineales entre las variables clínicas seleccionadas y la clasificación del riesgo metabólico. Este análisis permitió observar qué variables presentaban mayor asociación lineal con la variable objetivo y facilitó la construcción de la matriz de correlación utilizada en la interpretación de resultados.

De manera complementaria, se aplicó la técnica de *Información Mutua* como método de selección de características. Esta técnica permitió estimar la dependencia estadística entre cada variable predictora y la clasificación del riesgo metabólico, identificando variables relevantes incluso cuando su relación con la variable objetivo no fuera estrictamente lineal. A diferencia de la correlación de Pearson, la Información Mutua permite capturar asociaciones más amplias entre variables clínicas y categorías de riesgo.

Para la implementación de esta técnica se utilizó la función `mutual_info_classif` de Scikit-learn, considerando la naturaleza de las variables continuas y discretas incluidas en la matriz analítica. Los puntajes obtenidos fueron normalizados para expresar la importancia relativa de cada variable y facilitar su representación gráfica en el ranking de características.

En conjunto, la combinación de *correlación de Pearson e Información Mutua* permitió analizar la relevancia de las variables desde dos enfoques complementarios: el primero orientado a asociaciones lineales y el segundo a dependencia estadística general entre variables predictoras y clasificación del riesgo metabólico. Este procedimiento permitió establecer los determinantes clínicos más relevantes para las fases posteriores de modelado y evaluación.

Fase 2 Aplicación de Modelos de Aprendizaje Automático para la Estimación Individual del Riesgo

Relación con el Objetivo Específico 2. Esta fase respondió al segundo objetivo específico, relacionado con la aplicación de modelos de aprendizaje automático para estimar el riesgo individual de desarrollar Diabetes Mellitus Tipo II en población adulta.

Antes del entrenamiento de los modelos, se construyó la variable objetivo, denominada *Clasificación del Riesgo Metabólico*, definida a partir de los criterios diagnósticos de la Asociación Americana de Diabetes. Esta variable permitió transformar el problema clínico en una tarea de clasificación supervisada multicategoría.

Tabla 3

Clasificación del Riesgo Metabólico

Clasificación	Criterio clínico	Estado
Bajo (0)	Glucosa < 100 mg/dL y HbA1c < 5.7 %	Paciente metabólicamente sano
Moderado (1)	Glucosa entre 100–125 mg/dL o HbA1c entre 5.7–6.4 %	Prediabetes
Alto (2)	Glucosa \geq 126 mg/dL o HbA1c \geq 6.5 %	Diabetes Mellitus

Nota. Criterios clínicos empleados para clasificar el riesgo metabólico en bajo, moderado y alto, según los valores de glucosa y hemoglobina glicosilada. HbA1c = hemoglobina glicosilada; mg/dL = miligramos por decilitro.

Con el fin de fortalecer la validez analítica del modelado y controlar el posible riesgo de fuga de información (*target leakage*), se definieron dos escenarios de análisis. Esta decisión se fundamentó en que la variable objetivo, denominada Clasificación del Riesgo Metabólico, fue construida a partir de los valores de Glucosa Pre-prandial y Hemoglobina Glicosilada. Por tanto,

incluir simultáneamente estas variables como predictoras podía generar un desempeño artificialmente elevado, debido a que los modelos tendrían acceso a variables que contienen parte de la regla clínica utilizada para definir la etiqueta objetivo. Por esta razón, además del modelo completo, se incluyó un escenario preventivo sin variables diagnósticas directas, orientado a evaluar la capacidad predictiva de variables clínicas indirectas.

Tabla 4

Escenarios de Modelado Aplicados

Escenario	Variables incluidas	Propósito analítico
Modelo completo o diagnóstico	Incluye Glucosa Pre-prandial, HbA1c y demás variables clínicas	Evaluar el desempeño del modelo cuando dispone de marcadores diagnósticos directos
Modelo preventivo sin variables diagnósticas directas	Excluye Glucosa Pre-prandial y HbA1c; conserva variables clínicas, lipídicas, cardiovasculares, antropométricas y conductuales	Evaluar la capacidad predictiva del modelo sin depender directamente de las variables usadas para construir la etiqueta objetivo

Nota. Escenarios de modelado aplicados en el estudio, diferenciando entre un modelo completo y un modelo preventivo. HbA1c = hemoglobina glicosilada.

El modelo completo permitió valorar la capacidad de clasificación cuando el algoritmo dispone de marcadores diagnósticos directos. Por su parte, el modelo preventivo permitió analizar si variables indirectas como edad, IMC, triglicéridos, colesterol, presión arterial, hipertensión, dislipidemia, actividad física, consumo de alcohol, consumo de cigarrillo,

FINDRISC y Framingham aportan información suficiente para estimar el riesgo metabólico sin depender directamente de Glucosa Pre-prandial ni HbA1c.

El conjunto de datos fue dividido en subconjuntos de entrenamiento y prueba antes del proceso de imputación multivariada, con el fin de evitar contaminación del conjunto de prueba durante la estimación de valores ausentes. La partición se realizó bajo una proporción de *80 % para entrenamiento y 20 % para prueba*, quedando conformado el conjunto de entrenamiento por *2.200 registros* y el conjunto de prueba por *550 registros*. Esta misma partición fue conservada durante la construcción de los escenarios completo y preventivo, garantizando comparabilidad entre los modelos evaluados.

Debido al desbalance natural entre las clases de riesgo, se aplicó la técnica *SMOTE* exclusivamente sobre el conjunto de entrenamiento. Posterior a su aplicación, el conjunto de entrenamiento quedó balanceado con *1.637 registros por cada categoría de riesgo metabólico*: bajo, moderado y alto. Este balanceo se aplicó únicamente sobre los datos de entrenamiento, manteniendo intacto el conjunto de prueba para evitar sesgos en la evaluación final.

Asimismo, se aplicó *RobustScaler* para estandarizar las variables numéricas, preservando la influencia de valores extremos clínicamente plausibles y reduciendo el efecto de diferencias de escala entre variables.

En cada escenario se entrenaron tres modelos de clasificación supervisada:

Tabla 5

Modelos de Clasificación Supervisada

Modelo	Rol metodológico	Configuración general utilizada
Regresión	Modelo lineal de referencia	Modelo multiclase con aumento del número máximo
Logística		de iteraciones para asegurar convergencia

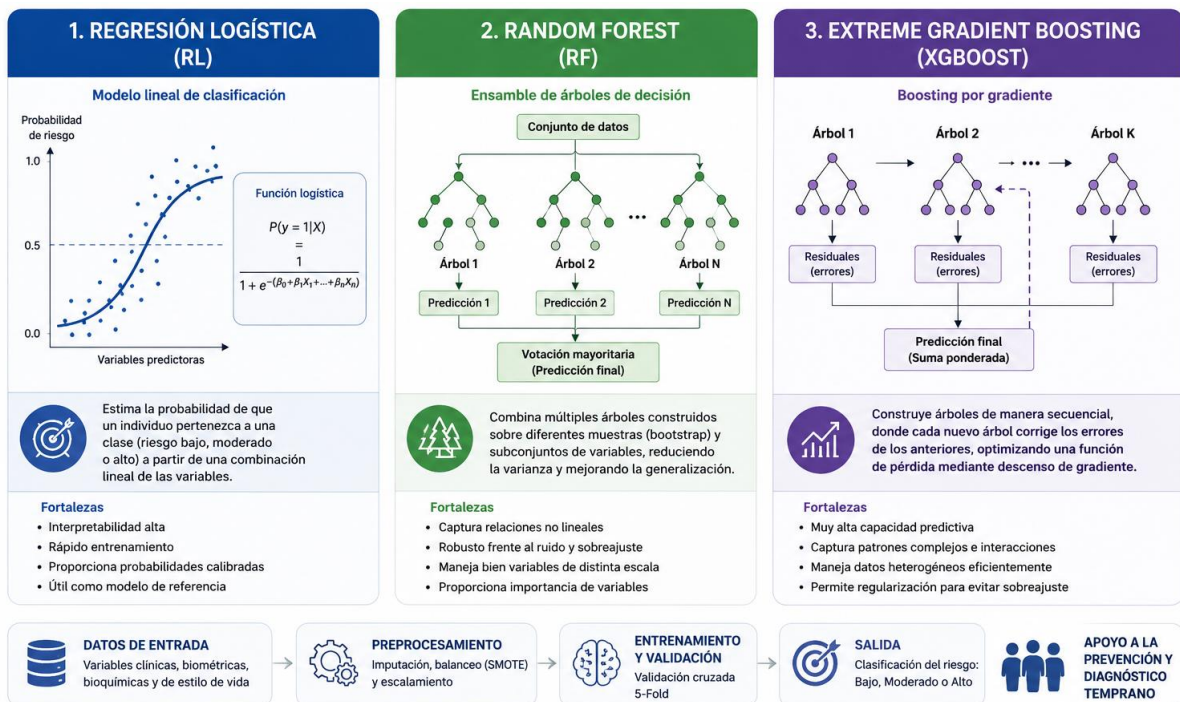
Modelo	Rol metodológico	Configuración general utilizada
Random	Modelo de ensamble paralelo	Ensamble de árboles de decisión con 300
Forest	basado en árboles	estimadores
XGBoost	Modelo de ensamble secuencial basado en boosting	Modelo de boosting con 300 estimadores, tasa de aprendizaje de 0,05 y profundidad máxima de 5

Nota. modelos de clasificación supervisada aplicados en el estudio y su configuración general.

La Regresión Logística se utilizó como línea base interpretable para contrastar el desempeño de modelos lineales frente a modelos de ensamble. Random Forest y XGBoost fueron incorporados por su capacidad para representar relaciones no lineales e interacciones complejas entre variables clínicas, metabólicas, cardiovasculares y antropométricas. Los modelos fueron entrenados bajo una configuración controlada y reproducible, empleando una semilla aleatoria fija para mantener la consistencia de los resultados.

Figura 3

Arquitectura de Modelos de Aprendizaje Automático Aplicados



Nota. Representación esquemática de la arquitectura general de los modelos de aprendizaje automático aplicados en el estudio: Regresión Logística, Random Forest y XGBoost

Fase 3 Evaluación, Comparación, Calibración y Selección del Modelo Predictivo

Relación con el Objetivo Específico 3. Esta fase respondió al tercer objetivo específico, orientado a evaluar y comparar el desempeño de los modelos mediante métricas de discriminación y calibración, con el fin de seleccionar el modelo con mejor rendimiento para apoyar acciones de prevención y diagnóstico temprano.

La evaluación se realizó en ambos escenarios definidos: modelo completo y modelo preventivo sin variables diagnósticas directas. Para cada modelo se calcularon métricas de clasificación por clase de riesgo y métricas globales de desempeño.

Las métricas de clasificación incluyeron:

- Exactitud (*Accuracy*).
- Precisión (*Precision*).
- Sensibilidad (*Recall*).
- F1-score.
- Matriz de confusión.

Se dio especial importancia al *Recall* en las categorías de riesgo moderado y alto, debido a que en salud preventiva resulta crítico minimizar falsos negativos, es decir, evitar que pacientes con alteraciones metabólicas sean clasificados erróneamente como sanos.

Figura 4

Métricas de Clasificación Utilizadas



Nota. métricas utilizadas para evaluar el desempeño de los modelos de clasificación: exactitud, precisión, sensibilidad, F1-score, AUC-ROC y matriz de confusión. Estas métricas permitieron analizar la capacidad de los modelos para clasificar correctamente a los pacientes en riesgo bajo, moderado y alto.

Para evaluar la capacidad discriminativa de los modelos se construyeron curvas ROC bajo un enfoque multiclase One-vs-Rest y se calcularon valores de AUC-ROC en ambos escenarios. Este análisis permitió observar la capacidad de los modelos para separar correctamente pacientes en riesgo bajo, moderado y alto.

Adicionalmente, se incorporó una evaluación de calibración probabilística, debido a que no basta con que un modelo clasifique correctamente, sino que también resulta importante que las probabilidades estimadas sean confiables. Para ello se calcularon:

- *Log Loss.*
- *Brier Score multiclase.*
- *Curvas de calibración*, especialmente para la categoría de riesgo alto.

Estas métricas permitieron valorar no solo la capacidad de clasificación de los modelos, sino también la confiabilidad de las probabilidades estimadas. En este sentido, la calibración complementó la evaluación discriminativa, permitiendo analizar si las probabilidades asignadas por los modelos fueron consistentes con la proporción real de pacientes observados en cada categoría de riesgo.

Finalmente, se analizó la importancia de variables del modelo con mejor desempeño, con el propósito de fortalecer la interpretabilidad clínica del sistema predictivo. Este análisis permitió identificar qué variables tuvieron mayor influencia en la estimación del riesgo y valorar si las decisiones del modelo fueron coherentes con criterios clínicos reconocidos.

La selección del modelo óptimo se realizó considerando de manera conjunta:

- desempeño global.
- sensibilidad en riesgo moderado y alto.
- capacidad discriminativa mediante AUC-ROC.

- comportamiento de calibración.
- interpretabilidad clínica.
- pertinencia del escenario completo frente al escenario preventivo.

Consideraciones Éticas y de Confidencialidad. El estudio se desarrolló a partir de información secundaria anonimizada, sin intervención directa sobre pacientes. El uso de los datos fue autorizado por la entidad de salud, garantizando la confidencialidad de la información clínica y evitando el uso de identificadores personales.

El tratamiento de los datos se realizó exclusivamente con fines académicos e investigativos, respetando los principios de privacidad, confidencialidad, anonimización y uso responsable de información sensible en salud.

Resultados

Identificación y Análisis de Variables Clínicas Relevantes Asociadas al Riesgo de Desarrollar Diabetes Mellitus Tipo II

Los resultados del análisis de variables clínicas, paraclínicas, antropométricas, cardiovasculares y conductuales permitieron identificar los factores con mayor relación frente a la clasificación del riesgo metabólico de desarrollar Diabetes Mellitus Tipo II en la población estudiada. La base analizada estuvo conformada por 2.750 registros clínicos anonimizados, correspondientes a pacientes adultos del programa institucional “Adultez”.

El conjunto analítico incluyó 19 variables predictoras y una variable objetivo denominada *Clasificación del Riesgo Metabólico*. Las variables evaluadas integraron dimensiones glucémicas, lipídicas, antropométricas, cardiovasculares, metabólicas complementarias, conductuales, demográficas y escalas clínicas de riesgo. Esta estructura permitió analizar el riesgo metabólico desde una perspectiva multidimensional, considerando tanto marcadores bioquímicos directos como factores clínicos, cardiovasculares y de estilo de vida.

Los hallazgos evidenciaron que las variables de mayor relevancia se concentraron principalmente en los marcadores glucémicos, especialmente *Hemoglobina Glicosilada (HbA1c)* y *Glucosa Pre-prandial*. Estas variables presentaron la mayor relación con la diferenciación entre pacientes clasificados en riesgo bajo, moderado y alto. Asimismo, se identificó un grupo de variables complementarias conformado por *Edad, Hipertensión Arterial, Resultado FINDRISC, Creatinina, Resultado Framingham, Tensión Arterial Sistólica y Triglicéridos*, las cuales aportaron información adicional para caracterizar el perfil cardiometabólico integral del paciente.

Variables Clínicas Analizadas

La matriz final de análisis estuvo compuesta por variables agrupadas en dimensiones clínicas relevantes para el riesgo metabólico. Esta organización permitió interpretar el fenómeno desde una perspectiva multidimensional, integrando marcadores bioquímicos, antecedentes cardiovasculares, características antropométricas y hábitos de vida.

Tabla 6

Variables Clínicas Evaluadas Según Dimensión Analítica

Dimensión clínica	Variables incluidas	Aporte clínico dentro del análisis
Glucémica	Glucosa Pre-prandial, Hemoglobina Glicosilada (HbA1c)	Marcadores principales para identificar alteraciones compatibles con prediabetes o diabetes
Lipídica	Triglicéridos, Colesterol HDL, Colesterol LDL, Colesterol VLDL	Indicadores asociados a dislipidemia y alteración metabólica
Antropométrica	Índice de Masa Corporal (IMC)	Factor asociado a exceso de peso, obesidad y resistencia a la insulina
Cardiovascular	Tensión Arterial Sistólica, Tensión Arterial Diastólica, Hipertensión Arterial, Resultado Framingham	Variables asociadas a riesgo cardiometabólico y comorbilidad vascular
Conductual	Actividad Física, Consumo de Alcohol, Consumo de Cigarrillo	Factores modificables relacionados con estilos de vida
Demográfica	Edad, Sexo	Variables de caracterización poblacional y susceptibilidad clínica

Dimensión clínica	VARIABLES INCLUIDAS	Aporte clínico dentro del análisis
Metabólica complementaria	Creatinina, Dislipidemia	VARIABLES ASOCIADAS AL ESTADO METABÓLICO Y FUNCIÓN ORGÁNICA COMPLEMENTARIA
Escala clínica	Resultado FINDRISC	Indicador de riesgo futuro de diabetes

Nota. Variables clínicas incluidas en el análisis, clasificadas por dimensión analítica y aporte clínico. HbA1c = hemoglobina glicosilada; HDL = lipoproteína de alta densidad; LDL = lipoproteína de baja densidad; VLDL = lipoproteína de muy baja densidad; IMC = índice de masa corporal; FINDRISC = *Finnish Diabetes Risk Score*, escala finlandesa de riesgo de diabetes. Framingham corresponde a una escala clínica utilizada para la estimación del riesgo cardiovascular.

La agrupación de variables permitió observar que el riesgo metabólico no estuvo representado por una única dimensión clínica, sino por la interacción entre factores glucémicos, lipídicos, cardiovasculares, antropométricos, conductuales y demográficos. No obstante, dentro del conjunto evaluado, los marcadores glucémicos fueron los que presentaron la mayor relación con la clasificación del riesgo, seguidos por variables asociadas al componente cardiovascular, la edad, las escalas clínicas de riesgo y el perfil metabólico complementario.

Relación entre Variables Clínicas y Clasificación del Riesgo Metabólico

La matriz de correlación correspondió al *coeficiente de correlación de Pearson*, lo que permitió identificar las asociaciones lineales más relevantes entre las variables clínicas seleccionadas y la *Clasificación del Riesgo Metabólico*. Los resultados evidenciaron que las variables con mayor correlación frente al riesgo fueron *Glucosa Pre-prandial* y *Hemoglobina*

Glicosilada (HbA1c), con coeficientes de *0,68* y *0,63*, respectivamente. Este comportamiento es clínicamente coherente, dado que ambos marcadores reflejan de manera directa el estado glucémico del paciente y permiten diferenciar con mayor claridad los niveles de riesgo bajo, moderado y alto.

Asimismo, se observó una correlación alta entre *Glucosa Pre-prandial* y *HbA1c* (*0,76*), lo cual confirma la consistencia entre el estado glucémico inmediato y el comportamiento glucémico sostenido en el tiempo. Esta relación respalda la relevancia de ambos marcadores dentro de la caracterización del riesgo metabólico.

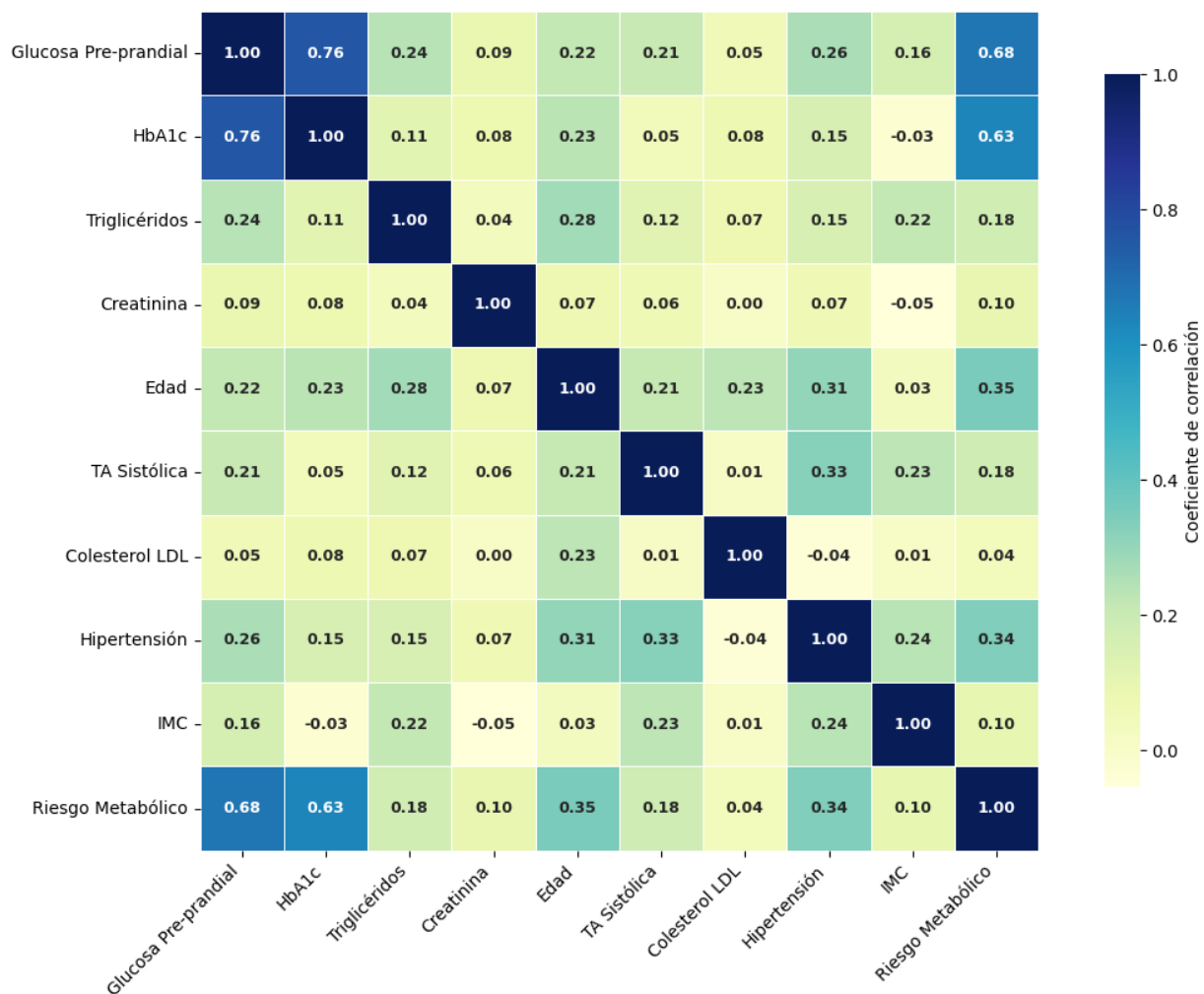
En un segundo nivel de asociación frente al riesgo metabólico se ubicó la *Edad*, con una correlación de *0,35*, seguida por *Hipertensión Arterial*, con *0,34*. Este hallazgo evidencia que el riesgo metabólico en la población estudiada no estuvo asociado únicamente a variables glucémicas, sino también a factores demográficos y condiciones cardiovasculares que aumentan la susceptibilidad clínica del paciente.

De igual forma, la *Tensión Arterial Sistólica* presentó una correlación de *0,18* con el riesgo metabólico y una asociación de *0,33* con Hipertensión Arterial, lo cual refuerza la relación entre el componente cardiovascular y el perfil metabólico del paciente. Por su parte, los *Triglicéridos* presentaron una correlación de *0,18* con el riesgo, indicando una asociación menor, pero clínicamente relevante dentro del componente lipídico.

VARIABLES COMO *IMC*, *Creatinina* Y *COLESTEROL LDL* presentaron correlaciones más débiles frente al riesgo metabólico, con valores de *0,10*, *0,10* y *0,04*, respectivamente. Esto indica que, de manera individual y bajo una relación lineal simple, estas variables no explican fuertemente la clasificación del riesgo; sin embargo, pueden aportar información clínica útil cuando son interpretadas de forma conjunta con otros factores cardiometabólicos.

Figura 5

Matriz de Correlación Entre Variables Clínicas Relevantes y Clasificación del Riesgo Metabólico



Nota. Matriz de correlación de Pearson entre variables clínicas relevantes y la clasificación del riesgo metabólico, permitiendo identificar asociaciones lineales entre los factores evaluados.

HbA1c = hemoglobina glicosilada; TA = tensión arterial; LDL = lipoproteína de baja densidad; IMC = índice de masa corporal.

En conjunto, la matriz de correlación evidencia que la clasificación del riesgo metabólico estuvo principalmente asociada con marcadores glucémicos, pero también mostró una relación relevante con *Edad e Hipertensión Arterial*. Estos hallazgos refuerzan la necesidad de interpretar

el riesgo de desarrollar Diabetes Mellitus Tipo II desde una perspectiva integrada, combinando marcadores glucémicos directos con factores cardiovasculares, demográficos, metabólicos y antropométricos.

Variables con Mayor Peso en la Selección de Características

El ranking de características basado en *Información Mutua* permitió identificar el grado de dependencia estadística entre cada variable clínica y la clasificación del riesgo metabólico. A diferencia de la correlación de Pearson, que refleja asociaciones lineales simples, la Información Mutua permitió valorar relaciones más amplias entre las variables predictoras y la variable objetivo, incluyendo posibles asociaciones no lineales.

Los resultados evidenciaron que las variables con mayor importancia relativa fueron *Hemoglobina Glicosilada (HbA1c)* y *Glucosa Pre-prandial*, con valores de 0,354 y 0,327, respectivamente. Estas dos variables concentraron la mayor proporción de información respecto a la clasificación del riesgo metabólico, confirmando su papel central en la diferenciación entre pacientes con riesgo bajo, moderado y alto.

En un segundo nivel de relevancia se ubicaron variables complementarias como *Resultado FINDRISC (0,059)*, *Edad (0,039)*, *Hipertensión Arterial (0,036)*, *Creatinina (0,034)*, *Resultado Framingham (0,032)*, *Colesterol VLDL (0,024)*, *Triglicéridos (0,021)* y *Colesterol HDL (0,021)*. Aunque estas variables presentaron una importancia relativa menor frente a los marcadores glucémicos, su presencia dentro del ranking confirma que el riesgo metabólico también está influenciado por factores cardiovasculares, metabólicos, renales, lipídicos, demográficos y de susceptibilidad clínica.

Por su parte, variables como *Colesterol LDL*, *Tensión Arterial Sistólica*, *Tensión Arterial Diastólica*, *Actividad Física*, *IMC*, *Consumo de Alcohol*, *Dislipidemia*, *Sexo* y *Consumo de*

Cigarrillo mostraron valores de importancia relativa más bajos. Este comportamiento no implica ausencia de relevancia clínica, sino menor aporte individual dentro de la diferenciación estadística de las categorías de riesgo metabólico en la muestra analizada.

Tabla 7

Ranking de Variables Clínicas Relevantes Asociadas al Riesgo Metabólico Mediante

Información Mutua

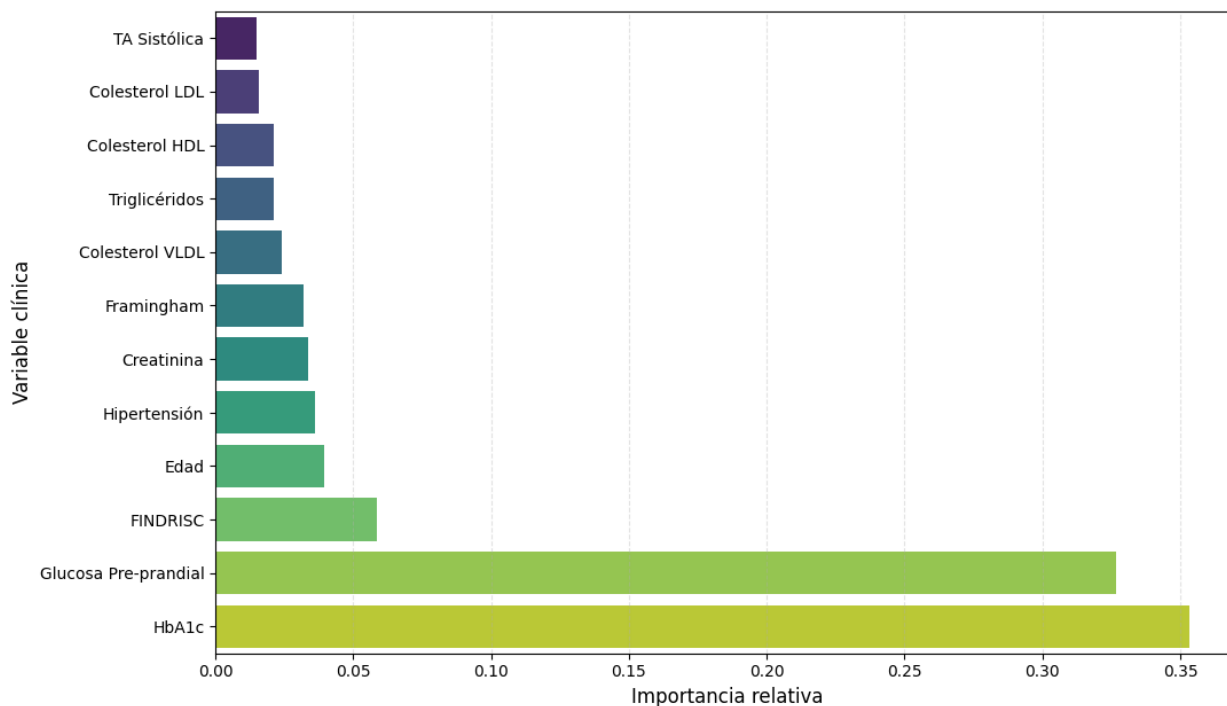
Orden	Variable clínica	Importancia relativa	Nivel de relevancia	Interpretación
1	Hemoglobina Glicosilada (HbA1c)	0,354	Muy alto	Refleja el comportamiento glucémico sostenido y permite diferenciar estados de prediabetes y diabetes
2	Glucosa Pre-prandial	0,327	Muy alto	Principal marcador asociado al estado glucémico inmediato y a la clasificación del riesgo metabólico
3	Resultado FINDRISC	0,059	Alto	Escala clínica asociada al riesgo futuro de diabetes
4	Edad	0,039	Moderado	Factor demográfico asociado al incremento progresivo del riesgo
5	Hipertensión Arterial	0,036	Moderado	Comorbilidad cardiovascular asociada a mayor riesgo cardiometabólico
6	Creatinina	0,034	Moderado	Variable complementaria relacionada con función renal y estado metabólico
7	Resultado Framingham	0,032	Moderado	Escala complementaria de riesgo cardiovascular
8	Colesterol VLDL	0,024	Bajo	Variable lipídica complementaria

Orden	Variable clínica	Importancia relativa	Nivel de relevancia	Interpretación
9	Triglicéridos	0,021	Bajo	Indicador lipídico asociado a alteración metabólica
10	Colesterol HDL	0,021	Bajo	Componente del perfil lipídico
11	Colesterol LDL	0,016	Bajo	Variable lipídica complementaria
12	Tensión Arterial Sistólica	0,015	Bajo	Indicador cardiovascular complementario

Nota. Resumen de las variables con mayor relevancia estadística frente a la clasificación del riesgo metabólico, según la técnica de Información Mutua. HbA1c = hemoglobina glicosilada; FINDRISC = *Finnish Diabetes Risk Score*, escala finlandesa de riesgo de diabetes; VLDL = lipoproteína de muy baja densidad; HDL = lipoproteína de alta densidad; LDL = lipoproteína de baja densidad. Framingham corresponde a una escala clínica utilizada para la estimación del riesgo cardiovascular.

Figura 6

Importancia de Variables Clínicas Asociadas al Riesgo Metabólico Técnica: Información Mutua



Nota. Importancia relativa de las variables clínicas asociadas a la clasificación del riesgo metabólico, calculada mediante la técnica de Información Mutua. HbA1c = hemoglobina glicosilada; TA = tensión arterial; LDL = lipoproteína de baja densidad; HDL = lipoproteína de alta densidad; VLDL = lipoproteína de muy baja densidad.

La gráfica evidencia una concentración marcada de la importancia relativa en los marcadores glucémicos. La *Hemoglobina Glicosilada (HbA1c)* obtuvo el mayor valor de importancia relativa (0,354), seguida por la *Glucosa Pre-prandial* (0,327). Este resultado confirma que las variables directamente relacionadas con el metabolismo de la glucosa fueron los principales determinantes estadísticos asociados a la clasificación del riesgo metabólico.

El tercer lugar fue ocupado por el *Resultado FINDRISC*, con una importancia relativa de 0,059, lo cual resulta clínicamente relevante, dado que esta escala integra factores asociados al

riesgo futuro de diabetes. Posteriormente se ubicaron *Edad, Hipertensión Arterial, Creatinina y Framingham*, variables que aportan información complementaria sobre el perfil demográfico, cardiovascular, renal y clínico del paciente.

En conjunto, los resultados de Información Mutua complementan los hallazgos obtenidos mediante la correlación de Pearson. Mientras Pearson mostró que *Glucosa Pre-prandial, HbA1c, Edad e Hipertensión Arterial* presentaron las mayores asociaciones lineales con el riesgo metabólico, la Información Mutua confirmó que las variables glucémicas concentran la mayor dependencia estadística con la clasificación del riesgo, incorporando además variables clínicas complementarias como *FINDRISC, Creatinina, Framingham* y variables del perfil lipídico.

Interpretación Clínica de las Variables Identificadas

Los resultados obtenidos permiten establecer que el riesgo metabólico en la población estudiada estuvo explicado principalmente por alteraciones glucémicas, representadas por la *HbA1c* y la *Glucosa Pre-prandial*. Estas variables fueron las más relevantes porque reflejan tanto el comportamiento sostenido de la glucosa en el tiempo como el estado glucémico inmediato del paciente.

La presencia de *FINDRISC* dentro de las variables con mayor importancia relativa evidencia la utilidad de las escalas clínicas para contextualizar el riesgo futuro de Diabetes Mellitus Tipo II. Asimismo, la participación de *Edad, Hipertensión Arterial, Framingham y Tensión Arterial Sistólica* refuerza la relación entre riesgo metabólico, riesgo cardiovascular y susceptibilidad clínica asociada al envejecimiento.

La inclusión de *Creatinina* entre las variables relevantes aporta una dimensión metabólica y funcional complementaria, debido a su relación con el estado renal y la condición clínica general del paciente. Por su parte, los *Triglicéridos, el Colesterol VLDL, el Colesterol*

HDL y el *Colesterol LDL* confirman la importancia del perfil lipídico dentro de la caracterización del riesgo cardiometabólico.

En conjunto, los hallazgos muestran que las variables más influyentes no pertenecen a una única dimensión clínica, sino que integran factores glucémicos, cardiovasculares, metabólicos, renales, lipídicos y de susceptibilidad clínica. Esto confirma que el riesgo de Diabetes Mellitus Tipo II debe interpretarse desde una perspectiva multidimensional y no únicamente desde valores aislados de glucosa.

Aplicación de Modelos de Aprendizaje Automático para la Estimación del Riesgo Individual de Desarrollar Diabetes Mellitus Tipo II

Los resultados obtenidos evidenciaron diferencias importantes en la capacidad de clasificación de los modelos para estimar el riesgo individual de desarrollar Diabetes Mellitus Tipo II. En el escenario completo, donde estuvieron presentes la Glucosa Pre-prandial y la Hemoglobina Glicosilada como marcadores glucémicos directos, los tres modelos presentaron un desempeño elevado, con exactitudes superiores al 90%, destacándose XGBoost y Random Forest por su mayor capacidad para clasificar pacientes en riesgo bajo, moderado y alto.

En contraste, el escenario preventivo, en el cual no se incluyeron estos marcadores glucémicos directos, presentó una disminución considerable en las métricas de desempeño, especialmente en las categorías de riesgo moderado y alto. Este hallazgo confirma el alto peso que tienen la Glucosa Pre-prandial y la Hemoglobina Glicosilada dentro de la clasificación del riesgo metabólico. Sin embargo, también evidencia que las variables clínicas indirectas conservan valor analítico para una aproximación inicial al riesgo, principalmente como apoyo en procesos de tamizaje y priorización preventiva.

La comparación entre ambos escenarios permitió diferenciar dos niveles de utilidad del sistema predictivo. El escenario completo mostró mayor precisión para la clasificación del riesgo metabólico cuando se dispone de información paraclínica completa, mientras que el escenario preventivo permitió valorar la capacidad de las variables no glucémicas para anticipar posibles perfiles de riesgo cuando no se dispone inicialmente de marcadores diagnósticos directos.

Distribución de las Categorías de Riesgo en los Conjuntos Analizados

La distribución de la variable objetivo evidenció predominio de pacientes clasificados en riesgo bajo, seguido por riesgo moderado y, en menor proporción, riesgo alto. En el conjunto de entrenamiento se observaron *1.637 registros de riesgo bajo, 437 de riesgo moderado y 126 de riesgo alto*. En el conjunto de prueba se registraron *410 pacientes de riesgo bajo, 119 de riesgo moderado y 21 de riesgo alto*.

Esta distribución confirma el comportamiento esperado en un contexto de tamizaje poblacional, donde los casos de riesgo alto suelen representar una proporción menor frente a los pacientes con bajo riesgo. Esta característica resulta relevante para interpretar el desempeño de los modelos, especialmente en las clases moderada y alta, que son las de mayor interés clínico para prevención y diagnóstico temprano.

Resultados Globales de Clasificación Según Escenario y Modelo

Los resultados globales muestran una diferencia clara entre los dos escenarios evaluados. En el modelo completo, la exactitud fue superior al *90%* en los tres algoritmos, alcanzando el valor más alto con *XGBoost*, con *99,45%*. En el escenario preventivo, la exactitud se ubicó entre *68,00%* y *78,18%*, lo que evidencia una reducción esperada del desempeño al excluir las variables glucémicas utilizadas como marcadores diagnósticos directos.

Tabla 8*Desempeño Global de los Modelos Según Escenario de Análisis*

Escenario	Modelo	Accuracy	Precision macro	Recall macro	F1score macro	Precision ponderada	Recall ponderado	F1-score ponderado
Modelo completo con Glucosa y HbA1c	Regresión Logística	90,73%	0,81	0,88	0,84	0,92	0,91	0,91
Modelo completo con Glucosa y HbA1c	Random Forest	98,36%	0,92	0,97	0,94	0,99	0,98	0,98
Modelo completo con Glucosa y HbA1c	XGBoost	99,45%	0,97	0,99	0,98	0,99	0,99	0,99
Modelo preventivo sin Glucosa/HbA1c	Regresión Logística	68,00%	0,50	0,62	0,50	0,77	0,68	0,71
Modelo preventivo sin Glucosa/HbA1c	Random Forest	76,91%	0,58	0,65	0,60	0,78	0,77	0,77
Modelo preventivo sin Glucosa/HbA1c	XGBoost	78,18%	0,59	0,62	0,59	0,78	0,78	0,78

Nota. Desempeño global de los modelos evaluados en los escenarios completo y preventivo. Las métricas macro otorgan igual peso a cada clase de riesgo, mientras que las métricas ponderadas consideran la distribución real de registros por clase. HbA1c = hemoglobina glicosilada.

Los resultados del modelo completo muestran que la inclusión de Glucosa Pre-prandial y HbA1c permitió una clasificación altamente precisa de los pacientes. En este escenario, la Regresión Logística alcanzó una exactitud de 90,73%, mientras que Random Forest y XGBoost alcanzaron desempeños superiores, con 98,36% y 99,45%, respectivamente.

Por su parte, el modelo preventivo evidenció un comportamiento más moderado. Al excluir las variables diagnósticas directas, la exactitud disminuyó a $68,00\%$ en Regresión Logística, $76,91\%$ en Random Forest y $78,18\%$ en XGBoost. Esta reducción no debe interpretarse como una debilidad del análisis, sino como un hallazgo relevante: las variables indirectas aportan información útil para la estimación inicial del riesgo, pero no reemplazan completamente la capacidad discriminativa de los marcadores glucémicos directos.

Comportamiento de los Modelos en el Escenario Completo

En el escenario completo, los modelos presentaron una alta capacidad para clasificar correctamente a los pacientes en los tres niveles de riesgo metabólico. Este comportamiento se explica por la presencia de variables directamente relacionadas con la definición clínica del riesgo, principalmente Glucosa Pre-prandial y HbA1c.

Regresión Logística en el Modelo Completo. La Regresión Logística alcanzó una exactitud global de $90,73\%$. En la clase de riesgo bajo, obtuvo una precisión de $0,98$, un recall de $0,92$ y un F1-score de $0,95$, lo que evidencia una buena capacidad para identificar pacientes sin alteraciones metabólicas relevantes.

En la clase de riesgo moderado, el modelo alcanzó una precisión de $0,75$, un recall de $0,86$ y un F1-score de $0,80$. Esto indica que logró identificar una proporción importante de pacientes con características compatibles con prediabetes, aunque con menor precisión que los modelos de ensamble.

En la clase de riesgo alto, obtuvo una precisión de $0,69$, un recall de $0,86$ y un F1-score de $0,77$. Este resultado muestra una sensibilidad adecuada para detectar pacientes de mayor compromiso metabólico, aunque con mayor margen de error frente a Random Forest y XGBoost.

Random Forest en el Modelo Completo. Random Forest mostró un desempeño superior a la Regresión Logística, alcanzando una exactitud global de 98,36%. En riesgo bajo obtuvo una precisión de 1,00, un recall de 0,99 y un F1-score de 1,00. En riesgo moderado alcanzó una precisión de 0,97, un recall de 0,96 y un F1-score de 0,96. En riesgo alto logró una precisión de 0,80, un recall de 0,95 y un F1-score de 0,87.

Estos resultados indican que Random Forest clasificó de forma muy consistente a los pacientes en los tres niveles de riesgo. La alta sensibilidad observada en la clase de riesgo alto resulta especialmente relevante desde una perspectiva preventiva, ya que esta categoría representa el grupo con mayor prioridad clínica.

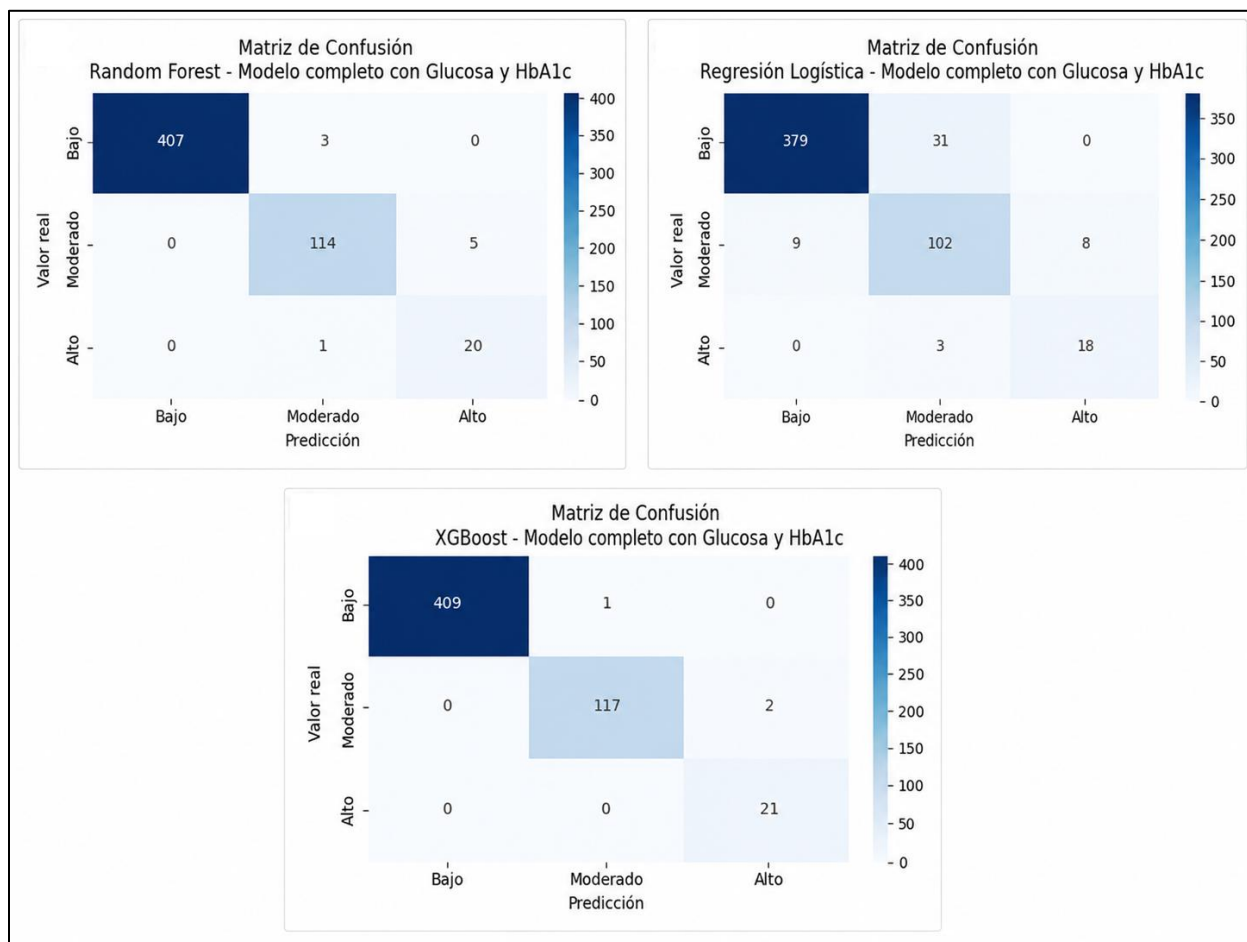
XGBoost en el Modelo Completo. XGBoost presentó el mejor comportamiento en el escenario completo, con una exactitud global de 99,45%. En riesgo bajo alcanzó una precisión de 1,00, un recall de 1,00 y un F1-score de 1,00. En riesgo moderado obtuvo una precisión de 0,99, un recall de 0,98 y un F1-score de 0,99. En riesgo alto logró una precisión de 0,91, un recall de 1,00 y un F1-score de 0,95.

Este comportamiento evidencia una capacidad sobresaliente para clasificar pacientes en las tres categorías de riesgo. En particular, la identificación completa de los pacientes clasificados en riesgo alto muestra que el modelo logró reconocer patrones clínicos asociados a alteraciones metabólicas avanzadas.

Matrices de Confusión del Escenario Completo

Figura 7

Matrices de Confusión de los Modelos Aplicados en el Escenario Completo



Nota. Matrices de confusión de Regresión Logística, Random Forest y XGBoost en el escenario completo, HbA1c = hemoglobina glicosilada.

Las matrices de confusión del escenario completo muestran que los errores de clasificación fueron bajos en los tres modelos, especialmente en Random Forest y XGBoost.

En Regresión Logística, se clasificaron correctamente *379 de 410* pacientes de riesgo bajo, *102 de 119* pacientes de riesgo moderado y *18 de 21* pacientes de riesgo alto. Los errores se

concentraron principalmente en la clasificación de pacientes de riesgo bajo como moderados y en algunos pacientes moderados clasificados como alto riesgo.

En Random Forest, el desempeño fue más estable: clasificó correctamente *407 de 410* pacientes de riesgo bajo, *114 de 119* pacientes de riesgo moderado y *20 de 21* pacientes de riesgo alto. Este resultado muestra una reducción significativa de errores frente a Regresión Logística.

XGBoost presentó el menor número de errores en el escenario completo, clasificando correctamente *409 de 410* pacientes de riesgo bajo, *117 de 119* pacientes de riesgo moderado y *21 de 21* pacientes de riesgo alto. Este comportamiento confirma su capacidad para diferenciar con alta precisión los tres niveles de riesgo metabólico.

Comportamiento de los Modelos en el Escenario Preventivo

El escenario preventivo permitió evaluar la capacidad de los modelos para estimar el riesgo individual sin utilizar Glucosa Pre-prandial ni HbA1c. Los resultados muestran una disminución importante del desempeño, especialmente en las clases de riesgo moderado y alto.

Este hallazgo es relevante porque evidencia que, aunque las variables clínicas indirectas contienen información útil, la estimación precisa del riesgo metabólico depende en gran medida de los marcadores glucémicos directos. Sin embargo, el escenario preventivo permite valorar el potencial de variables como edad, hipertensión arterial, tensión arterial, triglicéridos, IMC, dislipidemia, FINDRISC y Framingham para apoyar procesos de tamizaje inicial.

Regresión Logística en el Modelo Preventivo. En el escenario preventivo, la Regresión Logística alcanzó una exactitud global de *68,00%*. El modelo conservó un desempeño moderado en riesgo bajo, con precisión de *0,90*, recall de *0,75* y F1-score de *0,82*. Sin embargo, su capacidad para clasificar riesgo moderado y alto fue limitada.

En riesgo moderado obtuvo una precisión de *0,44*, recall de *0,43* y F1-score de *0,43*. En riesgo alto, aunque alcanzó un recall de *0,67*, la precisión fue de apenas *0,16*, lo cual indica una tendencia a clasificar como alto riesgo a pacientes que realmente pertenecían a otras categorías.

Este comportamiento refleja las limitaciones de un modelo lineal cuando no dispone de los marcadores glucémicos directos y debe estimar el riesgo a partir de variables clínicas indirectas.

Random Forest en el Modelo Preventivo. Random Forest alcanzó una exactitud global de *76,91%* en el escenario preventivo. El modelo presentó un buen desempeño en la clasificación de riesgo bajo, con precisión de *0,87*, recall de *0,86* y F1-score de *0,87*.

En riesgo moderado obtuvo una precisión de *0,54*, recall de *0,47* y F1-score de *0,50*. En riesgo alto logró una precisión de *0,33*, recall de *0,62* y F1-score de *0,43*. Estos resultados muestran que Random Forest logró reconocer parcialmente los patrones de riesgo moderado y alto, pero presentó dificultades para separarlos con precisión cuando no contó con Glucosa ni HbA1c.

XGBoost en el Modelo Preventivo. XGBoost obtuvo el mejor desempeño global dentro del escenario preventivo, con una exactitud de *78,18%*. En riesgo bajo alcanzó una precisión de *0,86*, recall de *0,89* y F1-score de *0,88*, mostrando una buena capacidad para reconocer pacientes con bajo riesgo metabólico.

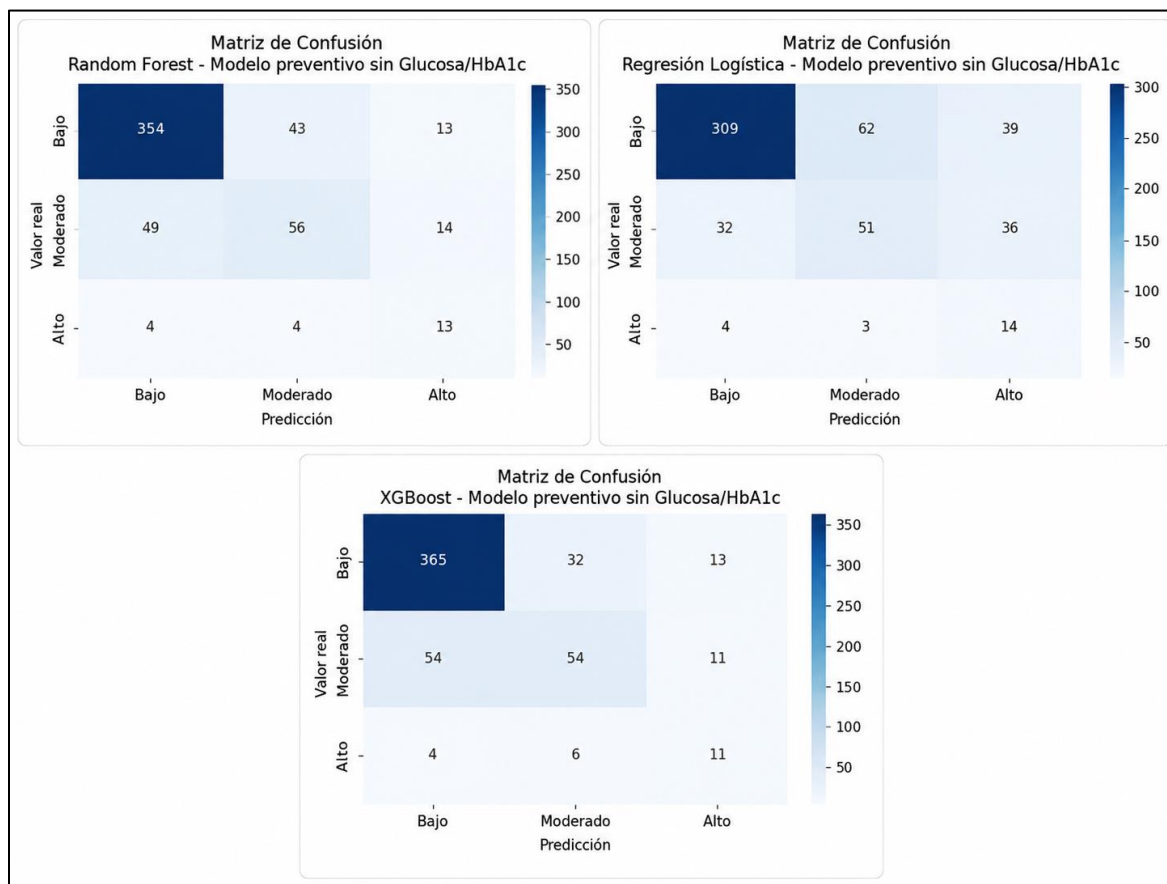
En riesgo moderado alcanzó una precisión de *0,59*, recall de *0,45* y F1-score de *0,51*. En riesgo alto obtuvo una precisión de *0,31*, recall de *0,52* y F1-score de *0,39*. Esto indica que, aunque XGBoost fue el modelo con mejor exactitud global en el escenario preventivo, su capacidad para identificar correctamente pacientes de riesgo alto fue menor que la observada en el escenario completo.

Este resultado confirma que las variables indirectas permiten una aproximación inicial al riesgo, pero no sustituyen el valor clínico de los marcadores glucémicos directos para una clasificación precisa.

Matrices de Confusión del Escenario Preventivo

Figura 8

Matrices de Confusión de los Modelos Aplicados en el Escenario Preventivo



Nota. matrices de confusión de Regresión Logística, Random Forest y XGBoost en el escenario preventivo. HbA1c = hemoglobina glicosilada.

Las matrices de confusión del escenario preventivo muestran un incremento considerable de errores en comparación con el escenario completo.

En Regresión Logística, se clasificaron correctamente *309 de 410* pacientes de riesgo bajo, *51 de 119* pacientes de riesgo moderado y *14 de 21* pacientes de riesgo alto. Se observaron errores relevantes en la clasificación de pacientes de riesgo bajo como moderado o alto, así como en la clasificación de pacientes moderados hacia las demás categorías.

En Random Forest, se clasificaron correctamente *354 de 410* pacientes de riesgo bajo, *56 de 119* pacientes de riesgo moderado y *13 de 21* pacientes de riesgo alto. Aunque mantuvo un desempeño aceptable en la identificación de pacientes de riesgo bajo, continuó mostrando dificultad para diferenciar adecuadamente los grupos de riesgo moderado y alto.

En XGBoost, se clasificaron correctamente *365 de 410* pacientes de riesgo bajo, *54 de 119* pacientes de riesgo moderado y *11 de 21* pacientes de riesgo alto. Aunque este modelo obtuvo la mayor exactitud global del escenario preventivo, las matrices muestran que su desempeño se concentró principalmente en la identificación de pacientes de riesgo bajo, con menor capacidad para reconocer las categorías de mayor riesgo.

Interpretación General de la Aplicación de Modelos

Los resultados evidencian que los modelos de aprendizaje automático permitieron estimar el riesgo individual de desarrollar Diabetes Mellitus Tipo II mediante la clasificación de pacientes en niveles de riesgo bajo, moderado y alto. En el escenario completo, los modelos alcanzaron desempeños elevados, destacándose Random Forest y XGBoost por su alta capacidad para identificar pacientes con riesgo moderado y alto.

No obstante, el desempeño sobresaliente observado en el escenario completo debe interpretarse con cautela, debido a que la Glucosa Pre-prandial y la Hemoglobina Glicosilada participaron en la construcción de la clasificación del riesgo metabólico y, al mismo tiempo, estuvieron presentes como variables predictoras. Esta condición puede favorecer un desempeño

artificialmente alto, dado que los modelos aprenden a partir de variables que contienen una parte importante de la regla clínica utilizada para definir la etiqueta de riesgo.

En contraste, el escenario preventivo mostró una disminución considerable del desempeño, especialmente en las categorías de riesgo moderado y alto. Este resultado confirma que la Glucosa Pre-prandial y la HbA1c son variables altamente determinantes para la clasificación del riesgo metabólico. Sin embargo, también evidencia que variables clínicas indirectas como edad, hipertensión arterial, tensión arterial sistólica, triglicéridos, IMC, dislipidemia, FINDRISC y Framingham conservan valor analítico para una aproximación inicial al riesgo.

Este hallazgo permite diferenciar dos niveles de utilidad del sistema predictivo. El escenario completo resulta más adecuado como modelo de clasificación asistida cuando se dispone de información paraclínica completa, mientras que el escenario preventivo aporta valor como herramienta de tamizaje inicial para priorizar pacientes que podrían requerir pruebas confirmatorias, seguimiento clínico oportuno o intervención preventiva.

En síntesis, la aplicación de los modelos permitió cumplir el propósito de estimar el riesgo individual de desarrollar Diabetes Mellitus Tipo II en población adulta. Los resultados muestran que el desempeño predictivo fue superior cuando se incorporaron marcadores glucémicos directos, mientras que el escenario preventivo aportó una visión más conservadora y realista sobre la capacidad de las variables clínicas indirectas para apoyar acciones de prevención y diagnóstico temprano.

Evaluación, Comparación y Selección del Modelo Predictivo con Mejor Desempeño Para la Estimación del Riesgo de Desarrollar Diabetes Mellitus Tipo II

Los resultados de evaluación evidenciaron diferencias importantes en el desempeño de los modelos según el escenario analizado. En el *modelo completo con Glucosa Pre-prandial y Hemoglobina Glicosilada*, los tres algoritmos presentaron una capacidad discriminativa elevada, con valores AUC cercanos a 1,00. Este comportamiento fue especialmente marcado en *Random Forest* y *XGBoost*, cuyos resultados mostraron una separación prácticamente perfecta entre las categorías de riesgo bajo, moderado y alto.

En contraste, el *modelo preventivo sin Glucosa Pre-prandial ni HbA1c* presentó una disminución considerable en la capacidad discriminativa. Este resultado confirma que los marcadores glucémicos directos tienen un peso determinante en la clasificación del riesgo metabólico. No obstante, el escenario preventivo permitió observar que las variables clínicas indirectas aún conservan valor predictivo para aproximar perfiles de riesgo, aunque con menor precisión que el modelo completo.

Capacidad Discriminativa de los Modelos Según Escenario

Los valores AUC-ROC mostraron un desempeño claramente superior en el escenario completo. En este escenario, *XGBoost* alcanzó el mayor valor AUC micro, con *1,000*, seguido por *Random Forest* con *1,000* y *Regresión Logística* con *0,985*. Estos resultados indican que, cuando se dispone de marcadores glucémicos directos, los modelos logran diferenciar con alta precisión entre los distintos niveles de riesgo metabólico.

En el escenario preventivo, los valores AUC disminuyeron de forma relevante. *Random Forest* obtuvo el mejor desempeño discriminativo global, con un AUC micro de *0,923*, seguido por *XGBoost* con *0,919* y *Regresión Logística* con *0,827*. Este comportamiento indica que, al

excluir Glucosa Pre-prandial y HbA1c, los modelos aún logran discriminar mejor que una clasificación aleatoria, pero con una separación menos precisa entre las categorías de riesgo.

Tabla 9

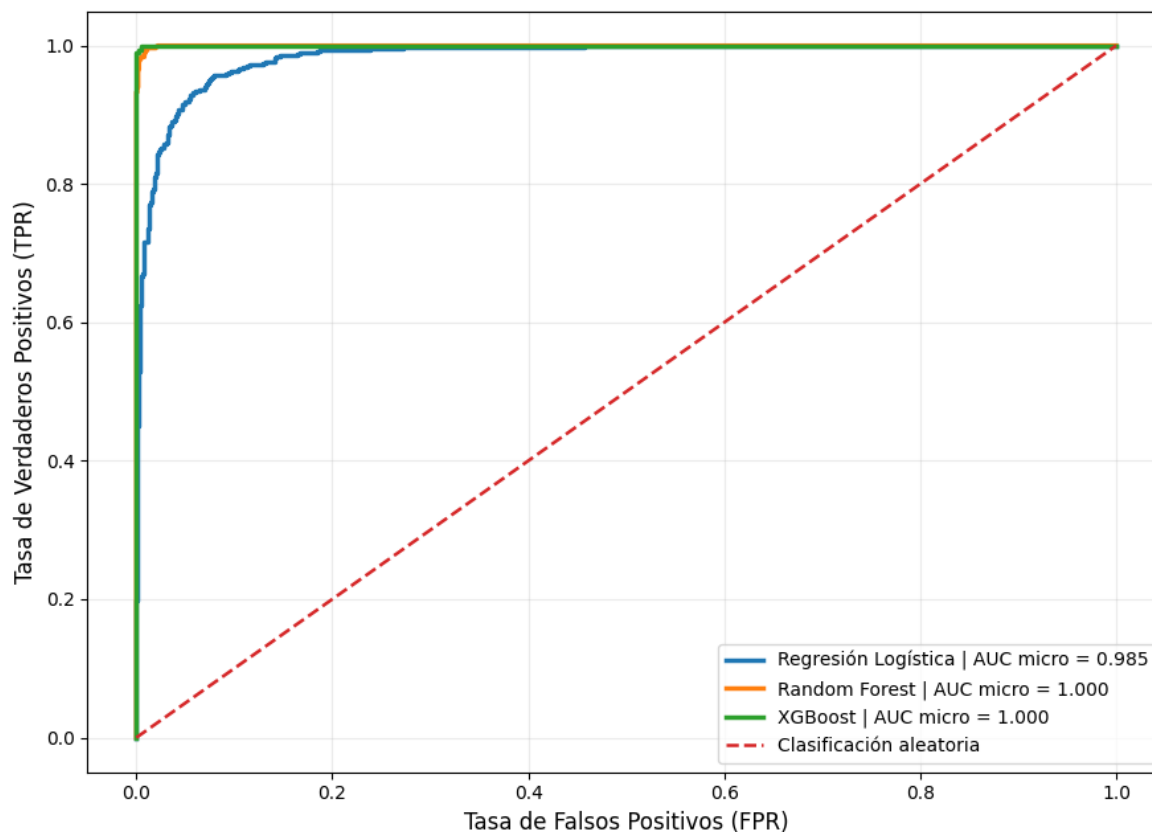
Resultados AUC-ROC por Modelo y Escenario

Escenario	Modelo	AUC	AUC	AUC
		micro	macro	ponderado
Modelo completo con Glucosa y HbA1c	Regresión	0,985	0,976	0,976
	Logística			
Modelo completo con Glucosa y HbA1c	Random Forest	1,000	0,999	1,000
Modelo completo con Glucosa y HbA1c	XGBoost	1,000	1,000	1,000
Modelo preventivo sin Glucosa/HbA1c	Regresión	0,827	0,778	0,785
	Logística			
Modelo preventivo sin Glucosa/HbA1c	Random Forest	0,923	0,851	0,839
Modelo preventivo sin Glucosa/HbA1c	XGBoost	0,919	0,839	0,822

Nota. capacidad discriminativa de los modelos mediante AUC-ROC en los escenarios completo y preventivo. El AUC micro evalúa el desempeño global agregado, el AUC macro asigna igual peso a cada clase y el AUC ponderado considera la distribución real de registros por clase. AUC = área bajo la curva; ROC = *Receiver Operating Characteristic*; HbA1c = hemoglobina glicosilada; XGBoost = *Extreme Gradient Boosting*.

Figura 9

Curvas ROC Comparativas del Modelo Completo con Glucosa y HbA1c

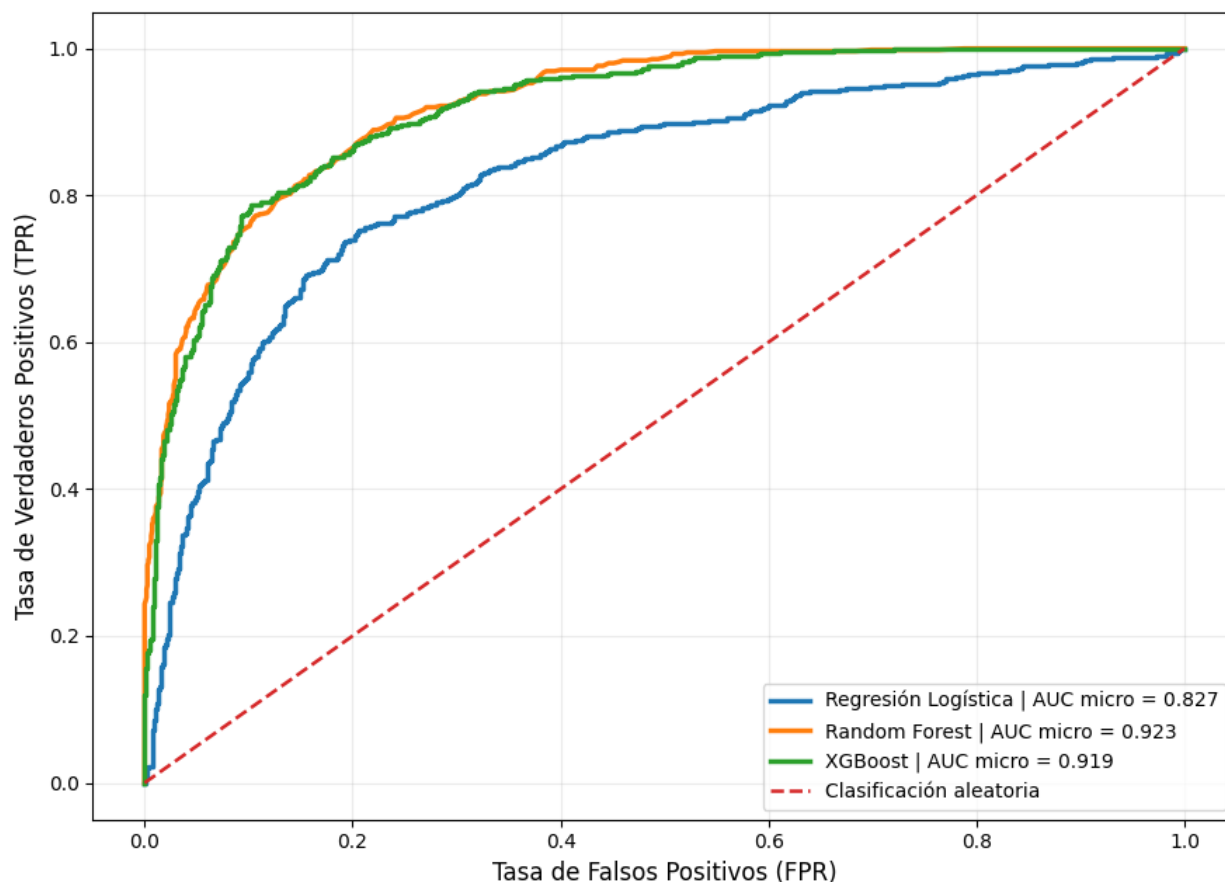


Nota. Curvas ROC comparativas de los modelos evaluados en el escenario completo, incluyendo Glucosa Pre-prandial y HbA1c. ROC = Receiver Operating Characteristic; AUC = Área Bajo la Curva; TPR = tasa de verdaderos positivos; FPR = tasa de falsos positivos; HbA1c = hemoglobina glicosilada.

La curva ROC del escenario completo evidencia que *Random Forest* y *XGBoost* alcanzaron una capacidad discriminativa prácticamente óptima. Las curvas se ubicaron muy cerca del extremo superior izquierdo, lo que indica alta sensibilidad y baja tasa de falsos positivos. La Regresión Logística también presentó un desempeño elevado, aunque ligeramente inferior frente a los modelos de ensemble.

Figura 10

Curvas ROC Comparativas del Modelo Preventivo sin Glucosa y HbA1c



Nota. Curvas ROC comparativas de los modelos evaluados en el escenario preventivo, excluyendo Glucosa Pre-prandial y HbA1c. ROC = Receiver Operating Characteristic; AUC = área bajo la curva; TPR = tasa de verdaderos positivos; FPR = tasa de falsos positivos; HbA1c = hemoglobina glicosilada.

En el escenario preventivo, las curvas ROC se alejaron del comportamiento casi perfecto observado en el modelo completo. *Random Forest* presentó la mejor capacidad discriminativa global, seguido muy de cerca por *XGBoost*. La Regresión Logística obtuvo el desempeño más bajo dentro de este escenario. Este resultado muestra que las variables clínicas indirectas permiten una aproximación al riesgo metabólico, pero no alcanzan el mismo nivel de separación obtenido cuando están presentes Glucosa Pre-prandial y HbA1c.

Desempeño Discriminativo de XGBoost por Categoría de Riesgo

El análisis por clase del modelo *XGBoost* mostró un comportamiento sobresaliente en el escenario completo. Los valores AUC fueron de 0,999 para riesgo bajo, 0,999 para riesgo moderado y 1,000 para riesgo alto. Esto indica que el modelo logró separar de forma prácticamente perfecta las tres categorías de riesgo cuando dispuso de los marcadores glucémicos directos.

En el escenario preventivo, el desempeño por clase disminuyó. *XGBoost* alcanzó un AUC de 0,830 para riesgo bajo, 0,779 para riesgo moderado y 0,908 para riesgo alto. Aunque estos valores siguen siendo superiores a una clasificación aleatoria, muestran una reducción importante en la capacidad de diferenciación, especialmente para la categoría de riesgo moderado.

Tabla 10

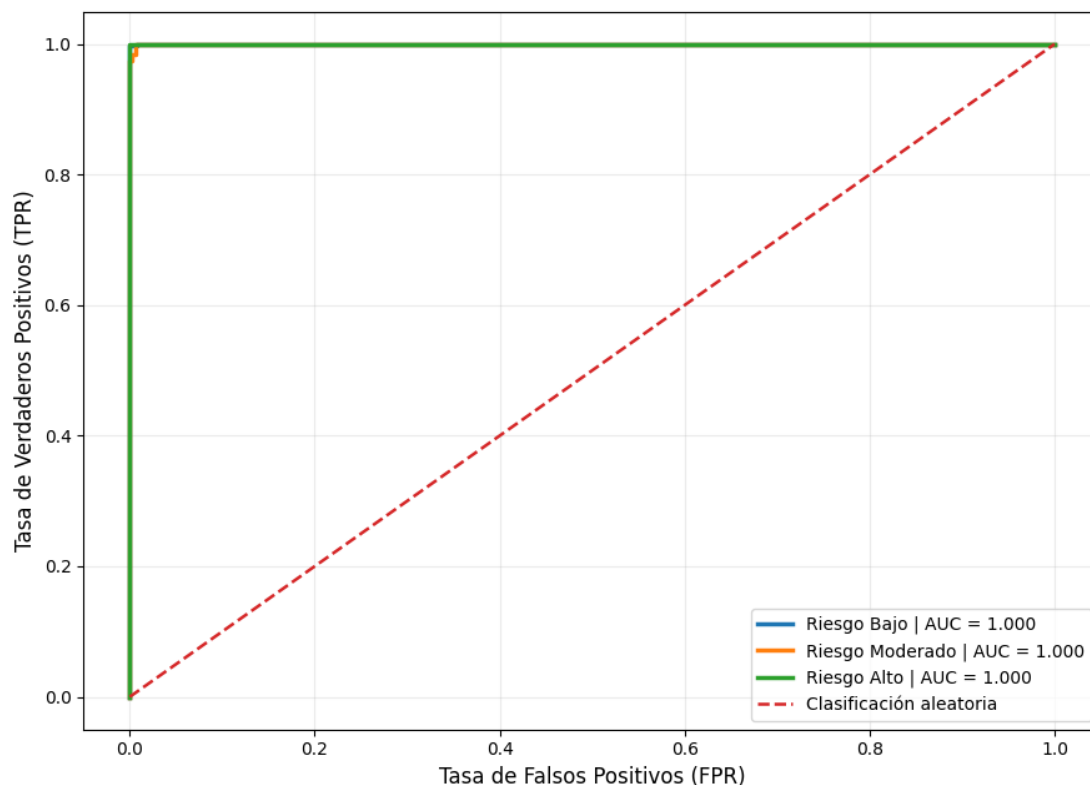
AUC-ROC por Clase para XGBoost Según Escenario

Escenario	Clase	AUC
Modelo completo con Glucosa y HbA1c	Riesgo Bajo	0,999
Modelo completo con Glucosa y HbA1c	Riesgo Moderado	0,999
Modelo completo con Glucosa y HbA1c	Riesgo Alto	1,000
Modelo preventivo sin Glucosa/HbA1c	Riesgo Bajo	0,830
Modelo preventivo sin Glucosa/HbA1c	Riesgo Moderado	0,779
Modelo preventivo sin Glucosa/HbA1c	Riesgo Alto	0,908

Nota. Valores de AUC-ROC por clase obtenidos por el modelo *XGBoost* en los escenarios completo y preventivo. AUC = área bajo la curva; ROC = *Receiver Operating Characteristic*; HbA1c = hemoglobina glicosilada; *XGBoost* = *Extreme Gradient Boosting*.

Figura 11

Curvas ROC por Clase para XGBoost en el Modelo Completo

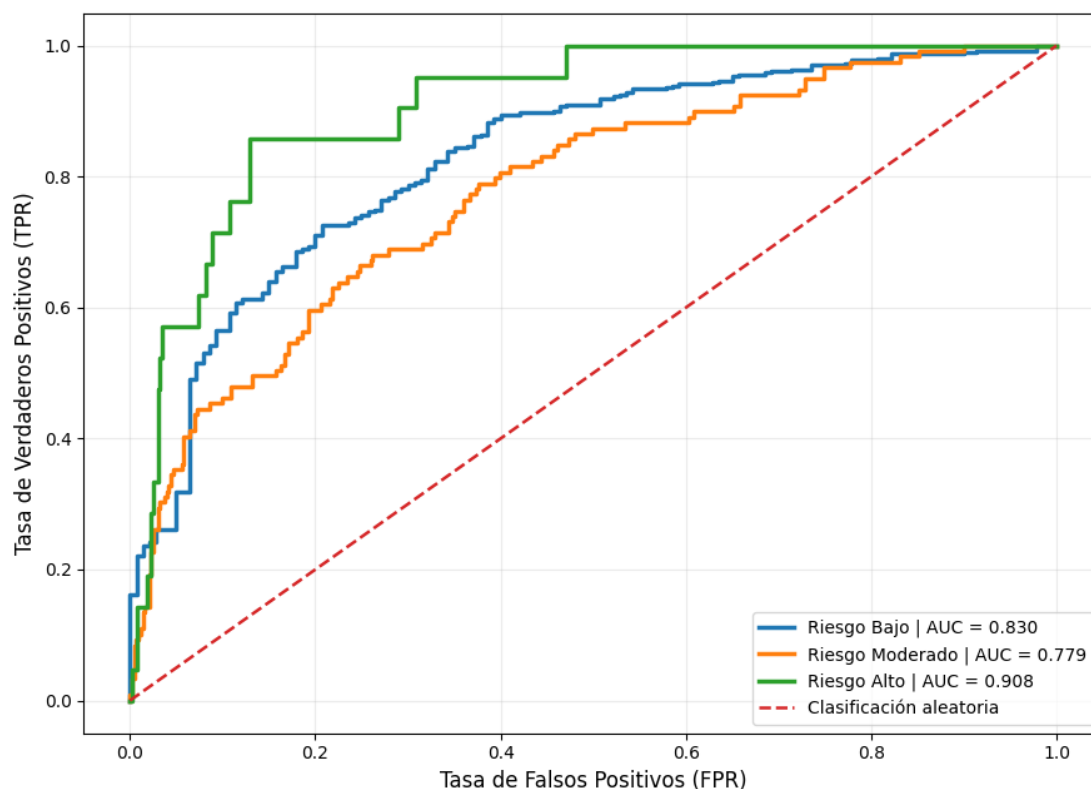


Nota. Curvas ROC por clase del modelo XGBoost en el escenario completo, permitiendo evaluar su capacidad para diferenciar las categorías de riesgo bajo, moderado y alto cuando se incluyen Glucosa Preprandial y HbA1c. ROC = Receiver Operating Characteristic; AUC = área bajo la curva; TPR = tasa de verdaderos positivos; FPR = tasa de falsos positivos; HbA1c = hemoglobina glicosilada.

La curva por clase del escenario completo muestra que XGBoost presentó una capacidad discriminativa muy alta en las tres categorías de riesgo. La clase de riesgo alto alcanzó un AUC prácticamente perfecto, lo cual evidencia una fuerte capacidad para identificar pacientes con alteraciones metabólicas avanzadas cuando el modelo dispone de información glucémica directa.

Figura 12

Curvas ROC por Clase para XGBoost en el Modelo Preventivo



Nota. Curvas ROC por clase del modelo XGBoost en el escenario preventivo, permitiendo evaluar su capacidad de discriminación para las categorías de riesgo bajo, moderado y alto sin incluir Glucosa Pre-prandial ni HbA1c como predictores. ROC = *Receiver Operating Characteristic*; AUC = área bajo la curva; TPR = tasa de verdaderos positivos; FPR = tasa de falsos positivos; HbA1c = hemoglobina glicosilada.

En el escenario preventivo, la curva por clase muestra una reducción del desempeño, especialmente en riesgo moderado. El mejor resultado se observó en riesgo alto, con un AUC de 0,908, lo cual indica que, aun sin marcadores glucémicos directos, el modelo logró conservar cierta capacidad para reconocer pacientes con perfiles clínicos de mayor riesgo. Sin embargo, el

menor AUC en riesgo moderado evidencia mayor dificultad para diferenciar pacientes en fases intermedias de alteración metabólica.

Evaluación de Calibración Probabilística

Los resultados de calibración mostraron diferencias relevantes entre escenarios. En el modelo completo, *XGBoost* presentó el mejor comportamiento probabilístico, con un *Log Loss* de 0,014 y un *Brier Score multiclase* de 0,009, seguido por Random Forest y Regresión Logística. Estos valores bajos indican que, además de clasificar correctamente, el modelo completo generó probabilidades más consistentes con las categorías reales observadas.

En el escenario preventivo, las métricas de calibración fueron menos favorables. *Random Forest* obtuvo el mejor comportamiento entre los modelos preventivos, con un *Log Loss* de 0,533 y un *Brier Score* de 0,314, seguido por *XGBoost* con 0,554 y 0,321, respectivamente. Esto indica que, aunque *XGBoost* alcanzó una exactitud ligeramente superior en el escenario preventivo, *Random Forest* mostró un comportamiento probabilístico más estable.

Tabla 11

Métricas de Calibración por Modelo y Escenario

Escenario	Modelo	Log Loss	Brier Score multiclase
Modelo completo con Glucosa y HbA1c	Regresión Logística	0,229	0,135
Modelo completo con Glucosa y HbA1c	Random Forest	0,099	0,038
Modelo completo con Glucosa y HbA1c	XGBoost	0,014	0,009
Modelo preventivo sin Glucosa/HbA1c	Regresión Logística	0,841	0,458
Modelo preventivo sin Glucosa/HbA1c	Random Forest	0,533	0,314
Modelo preventivo sin Glucosa/HbA1c	XGBoost	0,554	0,321

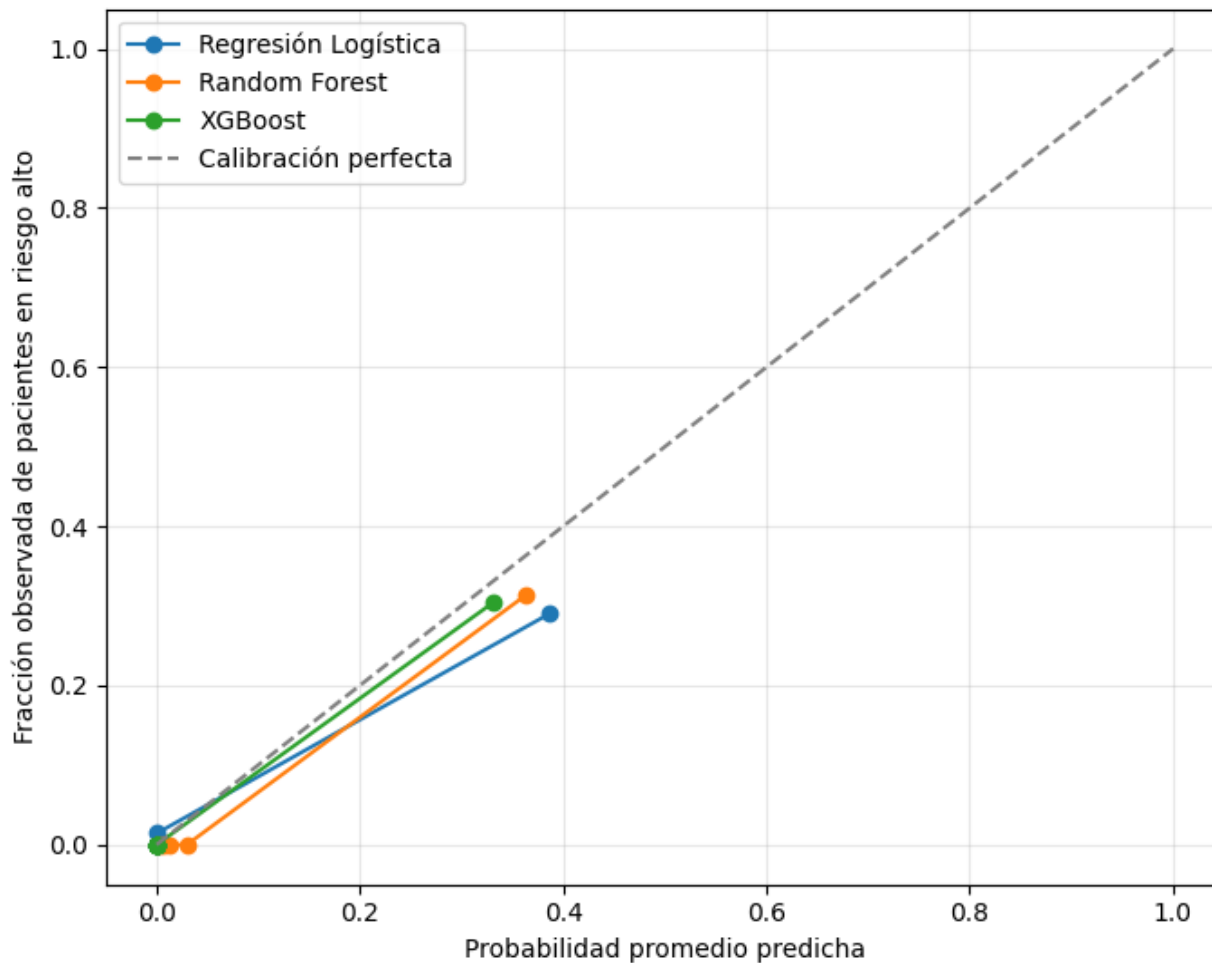
Nota. Calibración probabilística de los modelos mediante Log Loss y Brier Score multiclase.

Ambas métricas evalúan qué tan confiables son las probabilidades estimadas por el modelo;

valores más cercanos a cero indican mejor calibración. HbA1c = hemoglobina glicosilada;
XGBoost = *Extreme Gradient Boosting*.

Figura 13

Curva de Calibración para Riesgo Alto en el Modelo Completo

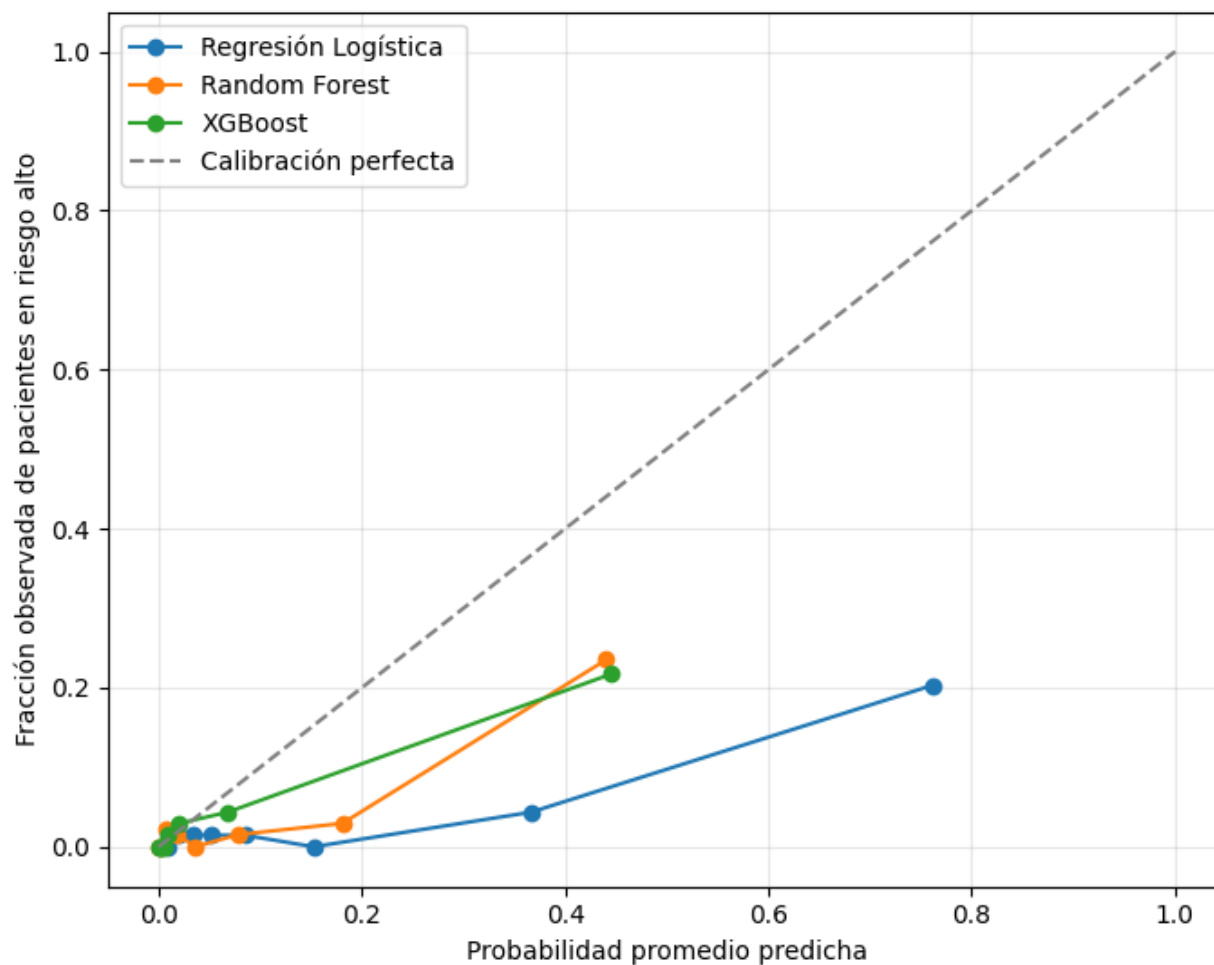


Nota. Curva de calibración para la categoría de riesgo alto en el escenario completo, permitiendo comparar la probabilidad promedio predicha por cada modelo con la fracción observada de pacientes en dicha categoría.

La curva de calibración del modelo completo mostró una mayor proximidad respecto a la línea de calibración perfecta, especialmente en *Random Forest* y *XGBoost*. Este comportamiento indica que las probabilidades estimadas para riesgo alto fueron más consistentes con la proporción observada de pacientes en dicha categoría.

Figura 14

Curva de Calibración para Riesgo Alto en el Modelo Preventivo



Nota. curva de calibración para la categoría de riesgo alto en el escenario preventivo, comparando la probabilidad promedio predicha por cada modelo con la fracción observada de pacientes en dicha categoría.

En el escenario preventivo, las curvas de calibración mostraron una mayor separación respecto a la línea ideal. Este resultado evidencia que, al excluir Glucosa Pre-prandial y HbA1c, las probabilidades estimadas para riesgo alto fueron menos estables. En términos analíticos, el modelo preventivo conserva utilidad para aproximar perfiles de riesgo, pero requiere interpretación cautelosa cuando se utiliza para estimar probabilidades individuales.

Interpretación Crítica de los Escenarios Evaluados

La comparación entre escenarios permitió identificar que el desempeño sobresaliente del modelo completo estuvo fuertemente influenciado por la presencia de *Glucosa Pre-prandial* y *HbA1c*, variables que también participaron en la definición de la clasificación del riesgo metabólico. Este hallazgo no invalida el modelo completo, pero sí exige interpretarlo como un modelo de *clasificación asistida con marcadores diagnósticos directos*, más que como un modelo puramente predictivo independiente.

En contraste, el escenario preventivo permitió evaluar la capacidad de los modelos para anticipar perfiles de riesgo a partir de variables clínicas indirectas. Aunque su desempeño fue menor, este escenario aporta una lectura más conservadora sobre la utilidad del modelo en contextos de tamizaje inicial, donde todavía no se cuenta con resultados de Glucosa Pre-prandial o HbA1c.

Este contraste es uno de los hallazgos más importantes del estudio. El modelo completo demuestra que los algoritmos pueden clasificar con alta precisión cuando cuentan con información paraclínica diagnóstica, mientras que el modelo preventivo muestra el límite real de predicción cuando solo se dispone de variables indirectas como edad, hipertensión, tensión arterial, triglicéridos, IMC, FINDRISC, Framingham y antecedentes metabólicos.

Selección del Modelo con Mejor Desempeño

A partir de los resultados obtenidos, *XGBoost en el escenario completo* presentó el mejor desempeño global del estudio. Este modelo alcanzó una exactitud de 99,45%, AUC micro de 1,000, AUC macro de 1,000, y los mejores indicadores de calibración, con el menor *Log Loss* (0,014) y el menor *Brier Score multiclase* (0,009). Sin embargo, este resultado debe interpretarse como un desempeño asociado a un modelo de *clasificación asistida*, dado que el escenario completo incluye marcadores diagnósticos directos como Glucosa Pre-prandial y Hemoglobina Glicosilada, variables que también participaron en la construcción de la clasificación del riesgo metabólico.

Por tanto, aunque *XGBoost* representa el modelo con mayor rendimiento cuando se dispone de información clínica y paraclínica completa, sus resultados no deben interpretarse como evidencia de predicción preventiva independiente, sino como una clasificación altamente precisa apoyada en variables diagnósticas directas. Esta precisión elevada confirma la capacidad del algoritmo para reproducir patrones clínicos asociados al riesgo metabólico, pero también exige cautela frente a posibles interpretaciones sobreestimadas del desempeño.

En el escenario preventivo, donde se excluyeron Glucosa Pre-prandial y HbA1c, los modelos presentaron desempeños más conservadores. En este contexto, *Random Forest* se destacó por ofrecer el mejor equilibrio entre capacidad discriminativa y calibración probabilística. Aunque *XGBoost* obtuvo una exactitud ligeramente superior en el escenario preventivo, *Random Forest* alcanzó mejores valores globales de AUC y mejores métricas de calibración, lo cual lo convierte en una alternativa más estable cuando el propósito es aproximar el riesgo a partir de variables clínicas indirectas.

En consecuencia, la selección del modelo debe interpretarse según el contexto de uso. Para una clasificación asistida con información paraclínica completa, *XGBoost en el escenario completo* representa el modelo con mejor rendimiento. Para procesos de tamizaje inicial sin marcadores glucémicos directos, *Random Forest en el escenario preventivo* ofrece una lectura más conservadora y equilibrada del riesgo.

En síntesis, el tercer objetivo específico se cumple al comparar el desempeño de los modelos mediante métricas de discriminación y calibración, identificando el mayor rendimiento global en XGBoost dentro del escenario completo, pero reconociendo que dicho desempeño corresponde a un modelo asistido por marcadores diagnósticos directos. De manera complementaria, el escenario preventivo aporta una visión más realista para apoyar acciones de priorización, tamizaje y diagnóstico temprano cuando aún no se dispone de resultados paraclínicos completos.

Conclusiones

El presente estudio permitió demostrar que la aplicación de modelos de aprendizaje automático sobre información clínica, paraclínica, antropométrica, cardiovascular, metabólica y conductual constituye una estrategia relevante para la estimación del riesgo de desarrollar Diabetes Mellitus Tipo II en población adulta. Los resultados evidenciaron que la analítica predictiva puede transformar registros clínicos rutinarios en herramientas de apoyo para la clasificación del riesgo metabólico, facilitando la identificación de pacientes en categorías de riesgo bajo, moderado y alto.

En relación con el primer objetivo específico, el análisis de variables permitió identificar que los marcadores glucémicos fueron los factores con mayor asociación frente a la clasificación del riesgo metabólico. La Glucosa Pre-prandial y la Hemoglobina Glicosilada presentaron la mayor relevancia dentro del análisis, lo cual es coherente con su papel clínico en la identificación de alteraciones compatibles con prediabetes y Diabetes Mellitus Tipo II. Asimismo, variables como Edad, Hipertensión Arterial, Resultado FINDRISC, Creatinina, Framingham, Tensión Arterial Sistólica y Triglicéridos aportaron información complementaria para caracterizar el perfil cardiometabólico integral del paciente.

Los hallazgos del análisis de variables confirman que el riesgo metabólico no debe interpretarse únicamente desde un marcador aislado, sino desde una perspectiva multidimensional. Aunque los marcadores glucémicos concentraron la mayor capacidad explicativa, la participación de variables cardiovasculares, lipídicas, antropométricas y de antecedentes clínicos evidencia que la Diabetes Mellitus Tipo II responde a una interacción compleja entre factores metabólicos y condiciones clínicas asociadas.

Respecto al segundo objetivo específico, la aplicación de modelos de aprendizaje automático permitió estimar el riesgo individual de los pacientes mediante una clasificación multicategoría. Los modelos evaluados lograron generar salidas predictivas diferenciadas para riesgo bajo, moderado y alto, demostrando que este tipo de técnicas puede aportar valor en procesos de tamizaje, priorización clínica y apoyo a la toma de decisiones en programas de promoción y mantenimiento de la salud.

Uno de los aportes más importantes del estudio fue la comparación entre dos escenarios analíticos. El escenario completo presentó un riesgo metodológico de fuga de información (*target leakage*), debido a que la Glucosa Pre-prandial y la Hemoglobina Glicosilada (HbA1c) fueron utilizadas para definir la variable objetivo según criterios ADA y, simultáneamente, como variables predictoras del modelo. Esta condición explica el desempeño elevado observado en dicho escenario, especialmente en Random Forest y XGBoost, ya que los modelos tuvieron acceso a variables que contienen parte de la regla clínica empleada para clasificar el riesgo metabólico. Por tanto, el escenario completo debe entenderse como un validador de coherencia interna del sistema de clasificación, no como un modelo de predicción prospectiva o preventiva independiente.

El escenario preventivo, construido sin Glucosa Pre-prandial ni Hemoglobina Glicosilada, permitió obtener una lectura más conservadora y metodológicamente más transferible sobre la capacidad real de las variables clínicas indirectas para anticipar perfiles de riesgo. Aunque su desempeño fue menor, este escenario resulta valioso porque refleja condiciones más cercanas a contextos de tamizaje inicial, donde aún no se dispone de marcadores glucémicos completos y se requiere priorizar pacientes para evaluación paraclínica, seguimiento clínico o intervención preventiva. En este sentido, el verdadero potencial predictivo

del estudio reside en el escenario preventivo, dado que opera con variables clínicas indirectas y no presenta el mismo problema estructural de fuga de información del escenario completo.

En relación con el tercer objetivo específico, la evaluación comparativa mostró que el modelo XGBoost en el escenario completo obtuvo el mejor desempeño global, con una exactitud de 99,45 %, AUC micro de 1,000, AUC macro de 1,000, Log Loss de 0,014 y Brier Score multiclase de 0,009. Estos resultados evidencian una capacidad sobresaliente para clasificar pacientes cuando se dispone de información paraclínica completa, especialmente variables glucémicas directamente relacionadas con la clasificación del riesgo.

No obstante, la evaluación del escenario preventivo mostró un comportamiento distinto. Aunque XGBoost obtuvo la mayor exactitud global en este escenario, con 78,18%, Random Forest presentó mejor equilibrio en términos de capacidad discriminativa y calibración probabilística, con AUC micro de 0,923, Log Loss de 0,533 y Brier Score multiclase de 0,314. Este hallazgo indica que, cuando se excluyen los marcadores glucémicos directos, Random Forest puede ofrecer una alternativa más estable para estimar riesgo a partir de variables clínicas indirectas.

La incorporación de métricas de calibración fortaleció la evaluación de los modelos, debido a que permitió analizar no solo la capacidad de clasificación, sino también la confiabilidad de las probabilidades estimadas. En este sentido, el modelo completo presentó mejores indicadores de calibración, mientras que el escenario preventivo mostró probabilidades menos estables, especialmente para la categoría de riesgo alto. Esto confirma que los modelos preventivos pueden ser útiles para aproximar perfiles de riesgo, pero sus resultados deben interpretarse con cautela y no reemplazan la valoración clínica ni las pruebas diagnósticas confirmatorias.

Desde una perspectiva aplicada, el estudio permite diferenciar dos niveles de utilidad. El modelo completo puede entenderse como una herramienta de clasificación asistida cuando se dispone de información clínica y paraclínica completa. En cambio, el modelo preventivo puede funcionar como una herramienta de apoyo al tamizaje inicial, orientada a priorizar pacientes que requieren confirmación diagnóstica, control metabólico o seguimiento oportuno.

La interpretabilidad de los modelos también aportó valor al estudio, dado que las variables con mayor peso estuvieron relacionadas con criterios clínicamente reconocidos en el riesgo metabólico. La importancia de HbA1c, Glucosa Pre-prandial, FINDRISC, Edad, Hipertensión Arterial, Creatinina, Framingham y variables del perfil lipídico respalda la coherencia clínica de los resultados y reduce la percepción del modelo como una herramienta completamente opaca. Sin embargo, la importancia de variables debe entenderse como una medida de contribución predictiva o dependencia estadística, no como una relación causal directa.

Desde el punto de vista metodológico, el uso del estándar CRISP-DM permitió estructurar de manera ordenada y reproducible las fases del proyecto, desde la comprensión de los datos y la preparación de la matriz analítica hasta el modelado, evaluación, calibración e interpretación de los resultados. La incorporación de dos escenarios de análisis fortaleció la validez del estudio, al permitir diferenciar el rendimiento de un modelo con marcadores diagnósticos directos frente a un modelo preventivo basado en variables clínicas indirectas.

El estudio presenta limitaciones que deben ser consideradas. La base de datos proviene de una única institución de salud y de una población específica vinculada al programa de adultez, por lo que la generalización de los modelos a otros contextos requiere validación externa. Asimismo, el alto desempeño del modelo completo debe interpretarse con cautela por la

participación de Glucosa Pre-prandial y HbA1c tanto en la construcción de la variable objetivo como en el conjunto predictor.

Finalmente, se concluye que los modelos de aprendizaje automático tienen un alto potencial para apoyar estrategias institucionales de prevención, tamizaje y diagnóstico temprano de Diabetes Mellitus Tipo II. Sin embargo, su implementación práctica debe diferenciar claramente entre modelos de clasificación asistida con información paraclínica completa y modelos preventivos basados en variables clínicas indirectas. Esta distinción es fundamental para evitar interpretaciones sobreestimadas del desempeño y para orientar el uso responsable de la inteligencia artificial en contextos clínicos reales.

Recomendaciones

Fortalecimiento de la Calidad y Estandarización de los Datos Clínicos

Se recomienda fortalecer los procesos institucionales de captura, validación y estandarización de la información clínica y paraclínica registrada en los sistemas de información en salud, especialmente en variables con alta relevancia para la clasificación del riesgo metabólico, como *Glucosa Pre-prandial*, *Hemoglobina Glicosilada (HbA1c)*, *Triglicéridos*, *Hipertensión Arterial*, *Tensión Arterial Sistólica*, *IMC* y perfil lipídico.

Los resultados del estudio evidenciaron que la calidad del dato constituye un factor determinante para el desempeño de los modelos de aprendizaje automático. En consecuencia, la reducción de vacíos de información, inconsistencias, registros incompletos y errores de digitación permitiría disminuir la dependencia de procesos de imputación y aumentaría la confiabilidad de futuros sistemas predictivos aplicados al entorno clínico.

Asimismo, se recomienda establecer protocolos institucionales de validación periódica de bases de datos clínicas, incorporando procesos automatizados de auditoría, detección de inconsistencias, validación de rangos clínicos y monitoreo continuo de calidad de datos. Esto permitiría garantizar la integridad analítica de la información utilizada en proyectos de inteligencia artificial en salud.

Diferenciación entre Modelos Diagnósticos y Modelos Preventivos

Se recomienda que futuras implementaciones institucionales diferencien claramente entre modelos de *clasificación asistida con marcadores diagnósticos directos* y modelos de *tamizaje preventivo basados en variables clínicas indirectas*.

El presente estudio demostró que el modelo completo, al incluir *Glucosa Pre-prandial* y *HbA1c*, alcanzó un desempeño sobresaliente; sin embargo, dicho comportamiento debe

interpretarse con cautela, debido a que estas variables también participaron en la construcción de la clasificación del riesgo metabólico. Por esta razón, en escenarios reales de implementación, este tipo de modelo debe entenderse como una herramienta de apoyo cuando ya se cuenta con información paraclínica completa, no como un modelo independiente de predicción temprana.

Por otra parte, el modelo preventivo sin Glucosa Pre-prandial ni HbA1c ofrece una aproximación más conservadora para contextos donde aún no se dispone de resultados paraclínicos. Aunque su desempeño fue menor, puede aportar valor como herramienta inicial para priorizar pacientes que requieren toma de laboratorios, seguimiento clínico o intervención preventiva temprana.

Prevención de Fuga de Información y Validación del Diseño Analítico

Se recomienda que en futuros desarrollos de modelos predictivos se realice una revisión explícita del riesgo de *fuga de información* o *target leakage*, especialmente cuando la variable objetivo se construye a partir de variables clínicas que también podrían incorporarse como predictoras.

En el caso de modelos para riesgo de desarrollar Diabetes Mellitus Tipo II, variables como Glucosa Pre-prandial y HbA1c tienen una relación directa con los criterios diagnósticos. Por tanto, su inclusión dentro del conjunto predictor puede incrementar artificialmente el desempeño de los modelos si también fueron utilizadas para definir la etiqueta objetivo. Para evitar interpretaciones sobreestimadas del rendimiento, se recomienda mantener análisis comparativos entre escenarios completos y escenarios preventivos sin marcadores diagnósticos directos.

Esta práctica permitiría diferenciar entre modelos útiles para clasificación clínica asistida y modelos con potencial de predicción preventiva, fortaleciendo la validez metodológica y la confianza en los resultados.

Asimismo, se recomienda que los procesos de preprocesamiento de datos, tales como imputación de valores faltantes, escalado de variables y balanceo de clases, se realicen siguiendo una separación estricta entre entrenamiento y prueba. En futuras implementaciones, las técnicas de transformación deberán ajustarse exclusivamente sobre el conjunto de entrenamiento y aplicarse posteriormente al conjunto de prueba, evitando que la información destinada a validación influya directa o indirectamente en el aprendizaje del modelo. Esta práctica fortalece la validez experimental y mejora la confiabilidad de las métricas reportadas.

Fortalecimiento de la Captura de Variables Conductuales y Antecedentes Clínicos

Se recomienda mejorar la recolección estructurada de variables relacionadas con hábitos de vida, antecedentes cardiovasculares y escalas de riesgo, incluyendo *actividad física, consumo de alcohol, consumo de cigarrillo, Dislipidemia, Hipertensión Arterial, FINDRISC y Framingham*.

Los resultados del escenario preventivo evidenciaron que, aunque las variables indirectas no sustituyen el poder discriminativo de Glucosa Pre-prandial y HbA1c, sí aportan información útil para una aproximación inicial al riesgo metabólico. Por tanto, mejorar la calidad y completitud de estas variables podría fortalecer futuros modelos preventivos orientados al tamizaje inicial.

Desde una perspectiva institucional, se recomienda capacitar al personal asistencial sobre la importancia clínica, epidemiológica y analítica de estas variables, promoviendo registros más completos y consistentes dentro de la historia clínica electrónica.

Validación Externa y Fortalecimiento de la Generalización del Modelo

Se recomienda realizar procesos de validación externa utilizando bases de datos provenientes de otras instituciones, regiones geográficas o poblaciones con características epidemiológicas diferentes, con el propósito de evaluar la capacidad de generalización de los modelos desarrollados.

Aunque el modelo completo, especialmente XGBoost, alcanzó un desempeño sobresaliente dentro de la población analizada, su rendimiento podría estar influenciado por las características específicas de la base de datos institucional y por la presencia de variables diagnósticas directas. Por esta razón, la validación sobre nuevas cohortes permitiría confirmar la estabilidad predictiva del modelo y reducir el riesgo de sobreajuste.

Adicionalmente, se recomienda implementar esquemas de actualización y reentrenamiento periódico, considerando que los perfiles epidemiológicos, hábitos poblacionales, criterios clínicos y dinámicas asistenciales pueden modificarse con el tiempo. Esta estrategia permitiría mantener la vigencia analítica del sistema frente a nuevas condiciones poblacionales.

Integración Institucional de Herramientas Predictivas y Tableros de Analítica Clínica

Se recomienda considerar el desarrollo de una plataforma institucional o tablero interactivo que integre los resultados del modelo predictivo dentro de los procesos asistenciales y administrativos del programa de promoción y mantenimiento de la salud.

Esta herramienta podría clasificar a los pacientes según niveles de riesgo metabólico bajo, moderado y alto, facilitando procesos de tamizaje poblacional, priorización de citas, solicitud de pruebas paraclínicas, seguimiento preventivo y control de pacientes con mayor probabilidad de alteraciones metabólicas.

Además, se recomienda que la plataforma incorpore elementos de interpretabilidad clínica, mostrando las variables que más influyen en la clasificación individual del paciente. Esto facilitaría la comprensión del resultado por parte del profesional de salud, fortalecería la confianza institucional en el uso de modelos predictivos y permitiría orientar intervenciones preventivas más personalizadas.

Desarrollo Futuro de Modelos Predictivos Preventivos

Como línea futura de investigación, se recomienda fortalecer el desarrollo de modelos preventivos que no dependan directamente de variables diagnósticas como Glucosa Pre-prandial o HbA1c, sino que integren variables clínicas indirectas, antecedentes, estilos de vida, escalas de riesgo y datos longitudinales de seguimiento.

La incorporación de información histórica del paciente permitiría evaluar trayectorias de riesgo y no solo estados clínicos puntuales. Esto podría mejorar la capacidad de anticipar perfiles de riesgo antes de que se manifiesten alteraciones glucémicas evidentes.

También se recomienda explorar nuevas fuentes de información, como antecedentes farmacológicos, datos nutricionales, adherencia a controles, antecedentes familiares, resultados seriados de laboratorio y datos provenientes de dispositivos de monitoreo. Estas fuentes podrían contribuir al desarrollo de modelos predictivos más robustos, explicables y aplicables en escenarios reales de atención primaria.

Implementación Progresiva y Supervisada de Inteligencia Artificial en Salud

Finalmente, se recomienda que cualquier implementación futura del modelo se realice de manera progresiva, supervisada y validada clínicamente. Antes de su despliegue operativo, el sistema debería ser evaluado en condiciones reales de atención, comparando sus recomendaciones con el criterio médico y con los resultados clínicos observados.

La inteligencia artificial puede aportar valor significativo en la priorización de riesgo, el tamizaje poblacional y la detección temprana de enfermedades crónicas no transmisibles. Sin embargo, su adopción institucional debe acompañarse de procesos de validación, capacitación, monitoreo, auditoría y mejora continua.

En este sentido, el modelo desarrollado debe considerarse una base analítica inicial con alto potencial de aplicación, pero su uso futuro debe integrarse bajo un enfoque responsable, transparente y centrado en el apoyo a la toma de decisiones clínicas.

Referencias Bibliográficas

- Alkhaldi, M., Khuoj, A., Almsfer, M., Alshulail, B., & Abudalfa, S. (2025). Machine Learning for Early Detection of Type 2 Diabetes Based on Liver Enzymes and BMI. In *Sustainable Data Management* (pp. 293–301). Springer.
- Alvarado Medina, H. S., & Barreiro Mendoza, G. S. (2025). Impacto de la telemedicina versus consultas presenciales en la gestión de diabetes e hipertensión. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 6(1), 3008–3021.
- Anderson, R. M., & Funnell, M. M. (2000). The Diabetes Empowerment Scale: a measure of psychosocial self-efficacy. *Diabetes Care*, 23(6), 739–743. DOI: .2337/diacare.23.6.739.
- Barragán-Silva, C. D., D'Espaux-Garrido, A. S., Vasconez-Garrido, D. T., & Pinto-Matos, E. J. (2025). Efectividad del uso de telemedicina en seguimiento de diabetes mellitus tipo 2 en comunidades andinas. *Sanitas Revista arbitrada de ciencias de la salud*, 4(especial), 117-127.
- Boutilier, J. J., Chan, T. C. Y., Ranjan, M., & Deo, S. (2021). Risk Stratification for Early Detection of Diabetes and Hypertension in Resource-Limited Settings: Machine Learning Analysis. *Journal of Medical Internet Research*, 23(1), e20123. DOI: 10.2196/20123.
- Delgado Ayala, F., Díaz Greene, E. J., & Rodríguez Weber, F. L. (2025). Utilidad clínica del Machine Learning con Python para la predicción de factores de riesgo cardiovascular. *Acta Médica Grupo Ángeles*, 23(4), 323-328.
- Granados Carrillo, N. Y., Gélvez Araque, J. D. C., & Paredes Ardila, M. L. (2024). Salud Móvil: Una Innovadora Solución para el Control y Prevención de Diabetes e Hipertensión en Comunidades Rurales. *Biociencias (UNAD)*, 8(1), 121–126. DOI: 10.22490/26194759.9204.

- Hasan, K. S. (2025). *Diabetes-Early-Warning* [Computer Software]. GitHub. Recuperado de <https://github.com/SakibHasanSimanto/Diabetes-Early-Warning>.
- Larose, C. D., & Larose, D. T. (2019). *Data Science Using Python and R*. Hoboken, NJ: John Wiley & Sons.
- Sierra Santos, I. (2025). *Aplicación Técnicas de Machine Learning para la Predicción de Costos de Seguros de Salud: Un Modelo Automatizado, Interactivo*. Fundación Universitaria Los Libertadores.
- Sun, H., Saeedi, P., Karuranga, S., et al. (2022). IDF Diabetes Atlas: Global, Regional and Country-Level Diabetes Prevalence Estimates for 2021 and Projections for 2045. *Diabetes Research and Clinical Practice*, 183, 109119. DOI: 10.1016/j.diabres.2021.109119.
- World Health Organization (WHO). (2003). *Adherence to Long-Term Therapies: Evidence for Action*. Geneva: WHO.
- World Health Organization (WHO). (2016). *Global Report on Diabetes*. Geneva: WHO.
- Young-Hyman, D., De Groot, M., Hill-Briggs, F., Gonzalez, J. S., Hood, K., & Peyrot, M. (2016). Psychosocial Care for People With Diabetes: A Position Statement of the American Diabetes Association. *Diabetes Care*, 39(12), 2126–2140. DOI: 10.2337/dc16-2053.