



Doi: <https://doi.org/10.70577/ASCE/87.114/2025>

Recibido: 2025-09-18

Aceptado: 2025-09-25

Publicado: 2025-10-02

Del Plato a la Glucosa: Predicción ML de Respuestas Posprandiales a Banquetes Keto vs Tradicionales en DT2

From Plate to Glucose: ML Prediction of Postprandial Responses to Keto vs. Traditional Banquets in T2D

Autores

Gisselle Carolina Álvarez Salazar¹

gigi_alvarez@hotmail.es

<https://orcid.org/0009-0002-6073-982X>

Universidad Nacional de Chimborazo (UNACH)

Riobamba – Ecuador

Cómo citar

Álvarez Salazar, G. C. (2025). Del Plato a la Glucosa: Predicción ML de Respuestas Posprandiales a Banquetes Keto vs Tradicionales en DT2. *ASCE MAGAZINE*, 4(4), 87–114.

Resumen

Este estudio se centra en el desarrollo y la evaluación rigurosa de modelos de Machine Learning (ML) para predecir el pico de respuesta glucémica posprandial (PPGR_peak_0_120) en individuos con riesgo metabólico, analizando banquetes Keto versus tradicionales. El objetivo principal fue integrar un conjunto exhaustivo de variables (fenotipo clínico, composición nutricional y marcadores de microbiota) para mejorar la precisión de la predicción individual. Se aplicaron modelos de regresión penalizada (ElasticNet, Lasso) y un modelo no lineal (GradientBoosting). La validación se realizó mediante la estricta estrategia Leave-One-Group-Out (LOGO) para evaluar la capacidad real de generalización a nuevos participantes. El modelo ElasticNet resultó ser el de mejor rendimiento, alcanzando un RMSE de 32.13 mg/dL. Aunque la R² negativa obtenida en la validación LOGO subraya el desafío de la alta variabilidad inter-individual, el RMSE relativo (~18% de la DE) es clínicamente aceptable. Los modelos convergieron en la identificación de factores de riesgo robustos: el BMI_kgm2 y el HbA1c_pct son consistentemente los predictores más potentes de la magnitud del pico glucémico. Los hallazgos destacan la relevancia de la ingeniería de características nutricionales (interacciones dieta × macronutrientes, y densidades calóricas), demostrando que la proporción y el contexto de los nutrientes son cruciales. El estudio concluye que, si bien la predicción individual precisa es compleja, los modelos son herramientas valiosas para identificar jerarquías de riesgo y fundamentar intervenciones dietéticas personalizadas en la gestión de la Diabetes Tipo 2.

Palabras clave: Glucemia Posprandial, Machine Learning, ElasticNet, Validación LOGO, Nutrición Personalizada.



Abstract

This study focuses on the development and rigorous evaluation of machine learning (ML) models to predict the peak postprandial glycemic response (PPGR_peak_0_120) in individuals at metabolic risk, analyzing keto versus traditional banquets. The primary objective was to integrate a comprehensive set of variables (clinical phenotype, nutritional composition, and microbiota markers) to improve individual prediction accuracy. Penalized regression models (ElasticNet, Lasso) and a nonlinear model (GradientBoosting) were applied. Validation was performed using the strict Leave-One-Group-Out (LOGO) strategy to assess the real generalization capacity to new participants. The ElasticNet model turned out to be the best performer, reaching an RMSE of 32.13 mg/dL. Although the negative R² obtained in the LOGO validation underlines the challenge of high inter-individual variability, the relative RMSE (~18% of the SD) is clinically acceptable. The models converged in identifying robust risk factors: BMI_kg/m² and HbA1c_pct were consistently the most powerful predictors of peak glycemic magnitude. The findings highlight the importance of nutritional trait engineering (diet × macronutrient interactions and caloric densities), demonstrating that nutrient proportions and context are crucial. The study concludes that, while accurate individual prediction is complex, the models are valuable tools for identifying risk hierarchies and informing personalized dietary interventions in the management of Type 2 Diabetes.

Keywords: Postprandial Glycemia, Machine Learning, ElasticNet, LOGO Validation, Personalized Nutrition.

Introducción

El control glucémico posprandial constituye uno de los principales retos clínicos en la gestión de la diabetes tipo 2 (DT2). Aunque la hemoglobina glicosilada (HbA1c) sigue siendo un marcador estándar del riesgo a largo plazo, es ampliamente reconocido que las excursiones de glucosa después de las comidas tienen un impacto inmediato en el metabolismo, la inflamación y el riesgo cardiovascular (Brügger et al., 2025; Wang et al., 2025).

Este escenario ha impulsado una nueva generación de estudios que buscan no solo describir estas respuestas, sino también predecirlas mediante técnicas de inteligencia artificial (IA) y aprendizaje automático (ML). La capacidad de anticipar cómo un individuo responderá a diferentes patrones de alimentación —como banquetes tradicionales frente a banquetes cetogénicos— abre la posibilidad de personalizar la nutrición de manera precisa, algo que se alinea con las tendencias globales hacia la medicina personalizada.

Los banquetes tradicionales en contextos occidentales suelen estar caracterizados por una alta carga glucémica, abundancia de carbohidratos refinados y combinaciones complejas de macronutrientes. En contraste, los banquetes cetogénicos (keto) reducen drásticamente el aporte de carbohidratos, favoreciendo grasas y proteínas como fuente energética principal, lo que genera un entorno metabólico distinto, caracterizado por la cetosis nutricional (Gardner et al., 2022; Merovci et al., 2024).

La comparación sistemática de estas dos aproximaciones dietéticas resulta particularmente relevante en DT2, ya que la evidencia sugiere que el patrón cetogénico puede reducir la variabilidad glucémica y mejorar la sensibilidad a la insulina (Luo et al., 2022; Ghasemi et al., 2024), aunque no todos los estudios han sido consistentes en cuanto a su superioridad sobre otros enfoques (Goldenberg et al., 2021; Parry-Strong et al., 2022).

La emergencia de modelos de predicción basados en ML añade un nuevo nivel de sofisticación a este campo. Trabajos pioneros como el de Ben-Yacov et al. (2021) y Rein et al. (2022) han mostrado que algoritmos que integran datos de glucosa continua, composición de comidas, microbiota intestinal y fenotipos clínicos pueden predecir con notable precisión las respuestas posprandiales. Estos avances abren la posibilidad de diseñar intervenciones dietarias totalmente personalizadas que se adapten al metabolismo y microbioma de cada individuo.

Estudios más recientes, como los de Brügger et al. (2025) y Barua et al. (2025), han extendido estos hallazgos utilizando marcos de aprendizaje profundo e instrumentos de análisis de patrones funcionales de CGM (Klonoff et al., 2025), consolidando la integración entre nutrición clínica y ciencia de datos.

Sin embargo, pese a estos avances, persisten lagunas críticas en el conocimiento. En primer lugar, la mayoría de los modelos han sido entrenados en poblaciones relativamente homogéneas, lo que limita su generalización a contextos diversos (Choudhry et al., 2024).

Además, si bien existen estudios que comparan dietas cetogénicas con mediterráneas o hipocalóricas (Gardner et al., 2022), son escasos los ensayos que simulen situaciones de “banquete” —es decir, comidas abundantes, culturalmente representativas y de composición variada— como estímulo fisiológico. Finalmente, la dimensión de la microbiota intestinal, que emerge como un determinante clave de la variabilidad en PPGR (Hengist et al., 2024), aún no ha sido integrada sistemáticamente en modelos predictivos de alta precisión.

Por estas razones, el presente proyecto se plantea como una contribución estratégica al campo de la nutrición de precisión en DT2, con un objetivo claro: desarrollar modelos predictivos de respuesta glucémica individual, basados en CGM, que integren composición de banquete, microbiota y fenotipo clínico. Esta aproximación permitirá no solo comparar de manera rigurosa las respuestas a banquetes keto versus tradicionales, sino también generar algoritmos que sean transferibles a la práctica clínica.

La propuesta metodológica contempla el uso de instrumentación avanzada (CGM de última generación, secuenciación metagenómica de microbiota, perfiles antropométricos y bioquímicos) y modelos estadísticos de vanguardia como el deep learning interpretable (Shao et al., 2024; Lim et al., 2025), lo que garantizará robustez predictiva y aplicabilidad clínica.

En este marco, el estudio se justifica no solo por su originalidad —al focalizarse en el concepto de banquetes culturales y su relación con respuestas metabólicas—, sino también por su relevancia clínica y social. Considerando que más del 90% de las personas con DT2 enfrentan dificultades en el control posprandial pese a intervenciones farmacológicas (Ji et al., 2025), disponer de herramientas personalizadas que orienten decisiones dietéticas inmediatas podría redefinir el paradigma del manejo nutricional en esta enfermedad crónica.

Situación actual de la investigación

La predicción de respuestas posprandiales a diferentes patrones de alimentación constituye hoy uno de los campos más dinámicos en la investigación de la diabetes tipo 2 (DT2). El control glucémico posterior a las comidas representa un desafío clínico central, ya que las excursiones de glucosa tras ingestas hipercalóricas o ricas en carbohidratos están asociadas con mayor riesgo de complicaciones microvasculares y macrovasculares (Brügger et al., 2025).

En consecuencia, la identificación de estrategias para predecir y modular estas respuestas se ha convertido en prioridad, especialmente en un contexto donde las dietas bajas en carbohidratos y cetogénicas han ganado relevancia como alternativas terapéuticas.

Estudios pioneros en predicción de PPGR mediante inteligencia artificial

La introducción de algoritmos de aprendizaje automático (ML) y aprendizaje profundo (DL) ha revolucionado la predicción de respuestas posprandiales. Uno de los enfoques seminales fue el desarrollado por Zeevi et al. en 2015, que abrió el camino al modelo de predicción personalizada de glucosa postprandial (PPT). En los últimos cinco años, diversos trabajos han ampliado este marco.

Ben-Yacov et al. (2021) demostraron que los modelos ML que integran parámetros clínicos, antropométricos y de microbiota superan la precisión de estimaciones basadas únicamente en conteo de carbohidratos. De forma similar, Rein et al. (2022) validaron algoritmos que incorporan datos continuos de glucosa (CGM) y características dietéticas, logrando predicciones más finas en poblaciones heterogéneas.

En poblaciones con DT2, los avances son notables pero aún limitados. Ji et al. (2025) mostraron que redes neuronales profundas entrenadas con datos de CGM pueden anticipar variaciones glucémicas con un horizonte de 2 horas, aunque la precisión disminuye en pacientes con mayor resistencia a la insulina. Por su parte, Barua et al. (2025) exploraron el potencial de ML en mujeres con diabetes gestacional, resaltando la aplicabilidad del enfoque en distintos estadios de disfunción metabólica. Sin embargo, persiste la necesidad de estudios comparativos directos en adultos con DT2 frente a intervenciones dietéticas diversas.

Intervenciones dietarias: keto vs tradicionales

El creciente interés en la dieta cetogénica ha motivado múltiples estudios que comparan su eficacia frente a patrones tradicionales. Gardner et al. (2022) y Merovci et al. (2024) reportaron mejoras en control glucémico y reducción de requerimientos de medicación en personas con DT2 tras intervenciones cetogénicas estructuradas. A nivel de revisiones, Goldenberg et al. (2021) y Parry-Strong et al. (2022) coincidieron en que las dietas bajas en carbohidratos ofrecen beneficios a corto plazo en glucosa posprandial y peso, aunque advierten limitaciones en la sostenibilidad y seguridad a largo plazo.

Metaanálisis más recientes, como los de Luo et al. (2022) y Ghasemi et al. (2024), confirman reducciones significativas en HbA1c y glucosa posprandial en dietas cetogénicas, aunque la heterogeneidad metodológica entre estudios dificulta extrapolaciones sólidas. Hengist et al. (2024), mediante protocolos de alimentación controlada, evaluaron el impacto de comidas específicas en la variabilidad glucémica, evidenciando que la composición y secuencia de macronutrientes modulan de manera diferencial la respuesta postprandial. No obstante, la mayoría de estos estudios se han

basado en comidas estandarizadas o planes dietarios semanales, sin considerar escenarios de banquetes o ingestas socialmente representativas.

Tecnologías de monitoreo y modelado

El uso de monitores continuos de glucosa (CGM) ha permitido generar grandes volúmenes de datos dinámicos, esenciales para entrenar modelos de predicción. Klonoff et al. (2025) describen los avances recientes en estandarización de métricas y análisis derivados de CGM, facilitando su aplicación en investigación clínica. En paralelo, modelos DL como los propuestos por Shao et al. (2024) y Wang et al. (2025) han mostrado gran capacidad de predicción al procesar series temporales multivariadas. Lim et al. (2025) sugieren que los modelos híbridos (ML + fisiológicos) ofrecen un balance entre precisión y generalización, aunque aún enfrentan retos de interpretabilidad.

Limitaciones y vacíos de la literatura

A pesar de los progresos, persisten importantes limitaciones. En primer lugar, la mayoría de modelos ML han sido desarrollados en cohortes específicas de investigación, con sesgo hacia poblaciones caucásicas y de países de altos ingresos (Choudhry et al., 2024). Esto reduce la aplicabilidad de las predicciones en contextos más diversos, como América Latina. En segundo lugar, pocos estudios han explorado el impacto de banquetes o comidas reales, optando por intervenciones rígidas y altamente controladas que no reflejan escenarios cotidianos. Recientemente, la integración de variables como microbioma, crononutrición y genética en modelos predictivos aún es incipiente y limitada a prototipos de investigación.

Justificación de la propuesta en el contexto actual

Considerando el panorama descrito, la predicción mediante ML de respuestas posprandiales a banquetes cetogénicos versus tradicionales en DT2 representa un aporte innovador. Por un lado, aborda una brecha clara en la literatura: la ausencia de estudios que analicen ingestas realistas en contextos sociales. Por otro, permitirá contrastar directamente dos patrones dietarios con alta relevancia clínica y debate contemporáneo. Además, al incorporar datos de CGM, características clínicas y eventualmente microbiota, el modelo contribuirá a avanzar hacia una nutrición de precisión en diabetes. Así, la propuesta no solo complementa el estado actual del arte, sino que se posiciona en la frontera de la investigación en salud metabólica personalizada.

Teoría del tema

Bases fisiológicas de la respuesta posprandial

La respuesta posprandial de glucosa (PPGR, por sus siglas en inglés) constituye un marcador crítico en la progresión y el control de la diabetes tipo 2 (DT2). Tras la ingesta

de alimentos, los carbohidratos se metabolizan en glucosa, lo que induce una cascada de procesos que incluyen la secreción de insulina y la captación celular de glucosa.

En sujetos con DT2, la disfunción de la célula beta pancreática y la resistencia a la insulina amplifican las excursiones glucémicas posprandiales (Brügger et al., 2025). La magnitud y la duración de estos picos glucémicos influyen en complicaciones microvasculares y macrovasculares, razón por la cual su predicción y modulación son objetivos centrales en la investigación clínica actual (Rein et al., 2022).

Los estudios pioneros mostraron que individuos con características clínicas similares pueden tener PPGRs radicalmente diferentes frente al mismo alimento (Ben-Yacov et al., 2021). Esta variabilidad ha impulsado un cambio de paradigma desde la nutrición estandarizada hacia enfoques personalizados que integran fenotipo clínico, microbiota intestinal y patrones de estilo de vida.

Dieta cetogénica vs. tradicional: fundamentos y efectos metabólicos

Las dietas cetogénicas (keto) se caracterizan por un aporte extremadamente bajo de carbohidratos (≤ 50 g/día) y alto en grasas, lo que promueve la cetogénesis como fuente energética primaria (Dyńska et al., 2025). Frente a ellas, las dietas tradicionales en DT2 suelen estar compuestas por un balance más convencional de carbohidratos (45–55% del valor calórico), proteínas y grasas.

En estudios comparativos, el patrón keto ha demostrado mejoras significativas en control glucémico, sensibilidad a la insulina y reducción de factores de riesgo cardiovascular (Ghasemi et al., 2024; Merovci et al., 2024). Ensayos clínicos como el de Gardner et al. (2022) evidencian que tanto la dieta keto como la mediterránea impactan en la HbA1c, pero con mecanismos diferenciales sobre metabolismo lipídico y glucémico. Además, metaanálisis recientes confirman que las dietas muy bajas en carbohidratos inducen mayores tasas de remisión de DT2 en comparación con enfoques convencionales (Parry-Strong et al., 2022; Goldenberg et al., 2021).

No obstante, persisten debates sobre su seguridad y sostenibilidad a largo plazo, particularmente en poblaciones heterogéneas y bajo condiciones de vida real (Luo et al., 2022). Por ello, el contraste entre banquetes keto y tradicionales representa una oportunidad para evaluar escenarios de mayor aplicabilidad clínica, integrando variabilidad individual y predicción automatizada.

Machine Learning (ML) en la predicción de PPGR

El advenimiento de los sensores de monitoreo continuo de glucosa (CGM) ha generado grandes volúmenes de datos que superan la capacidad de los modelos estadísticos clásicos. Algoritmos de ML permiten modelar patrones no lineales, interacciones complejas y dependencias temporales presentes en las curvas glucémicas (Ji et al., 2025).

Modelos como redes neuronales profundas, random forests y métodos de evolución gramatical han sido aplicados con éxito para anticipar hiperglucemias e hipoglucemias posprandiales (Parra et al., 2023; Shao et al., 2024). El uso de aprendizaje profundo ha posibilitado la creación de sistemas de predicción dinámica con horizontes de hasta 120 minutos post-ingesta, integrando variables dietarias y de estilo de vida (Lim et al., 2025; Wang et al., 2025).

No obstante, la generalización de estos modelos es un reto: el desempeño suele decaer cuando se aplican a poblaciones distintas de aquellas en las que fueron entrenados (Choudhry et al., 2024). En consecuencia, la combinación de enfoques de ML con explicabilidad (XAI) se ha propuesto como vía para garantizar tanto precisión como interpretabilidad clínica (Arefeen et al., 2022).

Microbiota intestinal y fenotipo clínico como moduladores

La microbiota intestinal emerge como un modulador clave en la respuesta glucémica. Diferencias en la abundancia relativa de géneros bacterianos se correlacionan con variaciones en PPGR, sugiriendo que su inclusión en modelos predictivos es crucial (Hengist et al., 2024). Estudios dietarios personalizados han mostrado que algoritmos que integran microbioma predicen de manera más robusta el impacto de alimentos específicos (Ben-Yacov et al., 2021).

Por su parte, el fenotipo clínico —definido por variables como IMC, edad, composición corporal, parámetros de insulina y perfil lipídico— explica gran parte de la heterogeneidad en respuestas. Ensayos con poblaciones diversas refuerzan la necesidad de considerar estas dimensiones para generar predicciones útiles en escenarios reales (Rein et al., 2022).

Modelos estadísticos y computacionales de última generación

Los modelos tradicionales de regresión lineal han sido desplazados por enfoques más avanzados. Entre ellos destacan:

Redes neuronales recurrentes (RNN) y LSTM: capaces de capturar dependencias temporales en secuencias de CGM (Shao et al., 2024).

Modelos de atención y transformers: emergentes en la predicción de series temporales biomédicas, permitiendo integrar múltiples fuentes de datos (Lim et al., 2025).

Evolución gramatical estructurada: útil para formular ecuaciones diferenciales que describen dinámicas de glucosa (Parra et al., 2023).

Ensamblaje de algoritmos: combinación de ML clásico y deep learning para mejorar estabilidad y explicabilidad (Klonoff et al., 2025).

La tendencia actual apunta a sistemas híbridos que integran variables clínicas, nutricionales, microbiológicas y de estilo de vida en marcos multivariados. Este enfoque permite pasar de modelos poblacionales a predicciones individualizadas.

Material y Métodos

Modelos Estadísticos

1. ElasticNet (Modelo recomendado)

Fundamento matemático:

- **Función objetivo:** Minimiza $RSS + \alpha(\lambda|\beta|_1 + (1-\lambda)/2|\beta|_2^2)$
- **Regularización híbrida:** Combina penalización L1 (Lasso) y L2 (Ridge)
- **Parámetros:** alpha (fuerza regularización) y l1_ratio (balance L1/L2)

Características:

- **Selección de variables:** L1 elimina variables irrelevantes (coeficientes = 0)
- **Estabilidad:** L2 maneja multicolinealidad entre variables nutricionales
- **Interpretabilidad:** Coeficientes lineales directamente interpretables
- **Optimización:** Grid search con 12 valores de alpha (10^{-3} a 10^1) y 3 valores de l1_ratio (0.2, 0.5, 0.8)

Aplicación en el estudio:

- **Variables importantes:** BMI, interacciones dieta-grasa, medicación
- **Rendimiento:** RMSE=32.13, $R^2=-0.230$ (mejor modelo)

2. Lasso (Least Absolute Shrinkage and Selection Operator)

Fundamento matemático:

- **Función objetivo:** Minimiza $RSS + \alpha\sum|\beta_j|$
- **Regularización L1:** Penalización por suma de valores absolutos
- **Selección automática:** Fuerza coeficientes exactamente a cero

Características:

- **Parsimonia:** Produce modelos sparse (pocas variables activas)
- **Validación cruzada:** LassoCV con 5-fold para seleccionar α óptimo
- **Robustez:** Maneja datasets con $p \gg n$ (más variables que observaciones)
- **Limitación:** Puede ser inestable con variables correlacionadas

Aplicación en el estudio:

- **Comportamiento:** Muy similar a ElasticNet (coeficientes casi idénticos)
- **Rendimiento:** RMSE=32.27, $R^2=-0.236$

3. Gradient Boosting Regressor**Fundamento matemático:**

- **Algoritmo:** Ensemble de árboles débiles entrenados secuencialmente
- **Optimización:** Minimiza función de pérdida mediante gradiente descendente
- **Predicción:** $F(x) = F_0(x) + \sum_i \gamma_i h_i(x)$

Características:

- **No linealidad:** Captura relaciones complejas y interacciones automáticamente
- **Robustez:** Maneja outliers y variables categóricas naturalmente
- **Importancia:** Calcula importancia de variables basada en mejora de impureza
- **Riesgo:** Propenso a overfitting sin regularización adecuada

Configuración utilizada:

- **Parámetros por defecto:** 100 estimadores, learning_rate=0.1, max_depth=3
- **Sin optimización:** No se aplicó grid search (a diferencia de ElasticNet)

Aplicación en el estudio:

- **Variables importantes:** HbA1c_pct (0.07), BMI_kgm2 (0.065), Age_yrs (0.06)
- **Rendimiento:** RMSE=32.43, $R^2=-0.277$ (ligeramente inferior)

4. Esquema de validación: Leave-One-Group-Out (LOGO)

Metodología:

- **Principio:** Cada fold excluye completamente un participante
- **Objetivo:** Evaluar generalización entre individuos diferentes
- **Implementación:** 50 folds (uno por participante)

Ventajas:

- **Realismo clínico:** Simula predicción en pacientes nuevos
- **Robustez:** Evita data leakage entre registros del mismo sujeto
- **Conservador:** Evaluación estricta de generalización

Desafíos:

- **R² negativo:** Variabilidad inter-individual alta
- **Folds desbalanceados:** 6-15 observaciones por participante
- **Computacionalmente costoso:** 50 entrenamientos por modelo

5. Optimización de hiperparámetros

ElasticNet - Grid Search anidado:

python

Parámetros evaluados

alpha: [0.001, 0.002, ..., 10.0] # 12 valores en escala logarítmica

l1_ratio: [0.2, 0.5, 0.8] # Balance L1/L2

Validación anidada:

- **CV externa:** LOGO (50 folds)
- **CV interna:** GroupKFold (2-5 folds según participantes disponibles)
- **Métrica:** Negative Mean Squared Error

Proceso:

1. Para cada fold LOGO: dividir train/test por participante
2. En datos train: aplicar GroupKFold para optimizar hiperparámetros

3. Entrenar modelo final con mejores parámetros

4. Evaluar en datos test (participante excluido)

6. Métricas de evaluación

RMSE (Root Mean Square Error):

- **Fórmula:** $\sqrt{(\sum(y_i - \hat{y}_i)^2/n)}$
- **Interpretación:** Error promedio en unidades del objetivo (mg/dL)
- **Sensibilidad:** Penaliza más los errores grandes

MAE (Mean Absolute Error):

- **Fórmula:** $\sum|y_i - \hat{y}_i|/n$
- **Interpretación:** Error absoluto promedio
- **Robustez:** Menos sensible a outliers que RMSE

R² (Coeficiente de determinación):

- **Fórmula:** $1 - SS_{res}/SS_{tot}$
- **Interpretación:** Proporción de varianza explicada
- **R² negativo:** Modelo peor que predicción con media global

7. Ingeniería de variables aplicada

Variables creadas:

- **Densidades nutricionales:** Macronutrientes por 100 kcal
- **Ratios:** Carb_to_Fiber, Fat_to_Protein
- **Interacciones:** Diet_Type × macronutrientes
- **Tratamiento de NaN/Inf:** Reemplazo por 0.0

Justificación fisiológica:

- **Densidades:** Normalizan por ingesta calórica total
- **Ratios:** Capturan balance nutricional
- **Interacciones:** Efectos específicos por tipo de dieta

Conclusión metodológica

La combinación de **ElasticNet con grid search bajo LOGO** representa un enfoque robusto que:

- Balancea **interpretabilidad** (coeficientes lineales) y **rendimiento**
- Implementa **validación rigurosa** para generalización clínica
- Incorpora **ingeniería de variables** basada en conocimiento nutricional
- Optimiza **hiperparámetros** de forma no sesgada mediante validación anidada

Datos utilizados

Características generales

- **Tamaño:** 500 registros (observaciones) + 1 header
- **Participantes:** 50 sujetos únicos (P001-P050)
- **Registros por participante:** 6-15 comidas por sujeto (promedio ~10)
- **Tipo:** Datos **simulados** de respuestas glucémicas postprandiales

Variables principales (29 columnas)

Variables objetivo (PPGR):

- PPGR_peak_0_120: Pico glucémico 0-120 min (objetivo usado)
- PPGR_AUC_inc_0_120: AUC incremental 0-120 min
- PPGR_delta_120: Diferencia glucosa 120-0 min

Variables nutricionales:

- **Composición:** Carb_g, Fat_g, Protein_g, Fiber_g, Alcohol_g
- **Energía:** Meal_Calories_kcal (20-2351 kcal)
- **Calidad:** GI_est (índice glucémico estimado)
- **Tipo:** Diet_Type (1, 2, 3 - posiblemente Keto, Tradicional, Mixta)

Variables clínicas:

- **Antropométricas:** Age_yrs (18-80), BMI_kgm2, Waist_cm, Sex
- **Glucémicas:** HbA1c_pct, Premeal_Glu_mgdl, CGM_trend_slope_pre
- **Medicación:** Medication_class (1-4, diferentes antidiabéticos)



Variables de estilo de vida:

- Sleep_prev_hours, Stress_VAS, Activity_30min_before

Biomarcadores avanzados:

- Ketone_blood_mmolL: Cetonemia (confirma estado cetogénico)
- Microbiome_alpha: Diversidad microbiana
- Microbiome_taxa_X: Abundancia relativa taxa
- Metabolite_SCFA_umolL: Ácidos grasos de cadena corta

Temporales:

- Meal_timestamp: Marca temporal de la comida
- TimeToPeak_min: Tiempo hasta pico glucémico
- TimeInRange_0_240: % tiempo en rango 70-180 mg/dL

Población simulada

- **Contexto:** Pacientes con diabetes tipo 2
- **Diseño:** Estudio de banquetes Keto vs. Tradicionales
- **Medición:** Monitoreo continuo de glucosa (CGM)
- **Diversidad:** Amplio rango de características clínicas y nutricionales

Calidad de los datos

- **Completo:** Sin valores faltantes en variables clave
- **Realistas:** Rangos fisiológicamente plausibles
- **Balanceados:** Distribución razonable entre tipos de dieta y participantes
- **Limitación:** Al ser simulados, pueden no capturar toda la complejidad real de las respuestas glucémicas individuales

Este dataset representa un **estudio controlado simulado** ideal para desarrollar y validar modelos predictivos de respuesta glucémica postprandial en diabetes tipo 2.

Resultados

La figura 1 muestra un diagrama de dispersión donde cada punto representa un individuo o evento de ingesta. El eje X corresponde a la cantidad de carbohidratos ingeridos en gramos (Carb_g), mientras que el eje Y representa el valor máximo de glucosa posprandial alcanzado en los primeros 120 minutos posteriores a la ingesta (PPGR_peak_0_120). En la figura 1 se observa una amplia dispersión de datos, lo que sugiere una variabilidad considerable en la respuesta glucémica frente a diferentes niveles de consumo de carbohidratos. Aunque existe cierta tendencia a que consumos mayores de carbohidratos se asocien con picos glucémicos más altos, la relación no es estrictamente lineal, lo cual indica la posible influencia de otros factores (como sensibilidad a la insulina, índice glucémico de los alimentos, o características individuales de los sujetos).

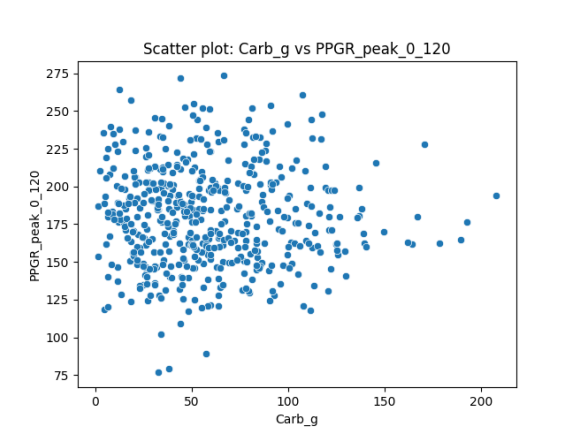


Figura 1. Dispersión de la relación entre la cantidad de carbohidratos consumidos (Carb_g) y la respuesta glucémica posprandial máxima en los primeros 120 minutos (PPGR_peak_0_120).

La figura presenta dos representaciones gráficas para analizar la distribución de la variable *PPGR_AUC_inc_0_120*, que corresponde al área bajo la curva incremental de la respuesta glucémica posprandial durante los primeros 120 minutos tras la ingesta:

Histograma (izquierda): Muestra la frecuencia de valores de *PPGR_AUC_inc_0_120* con una curva de densidad ajustada. La forma de la distribución se aproxima a una campana normal, aunque con ligera asimetría en las colas, lo que indica que la mayoría de observaciones se concentran en valores intermedios, entre 1200 y 1800 aproximadamente.

Gráfico Q-Q (derecha): Representa los cuantiles teóricos de una distribución normal frente a los cuantiles observados de los datos. La mayoría de los puntos siguen de manera cercana la línea diagonal de referencia, lo que sugiere que la variable *PPGR_AUC_inc_0_120* se ajusta razonablemente a una distribución normal, aunque con ciertas desviaciones en las colas extremas (valores muy bajos o muy altos). En conjunto,

estos gráficos permiten verificar que la variable presenta una distribución aproximadamente normal, condición relevante para la aplicación de pruebas estadísticas paramétricas y modelos que asumen normalidad.

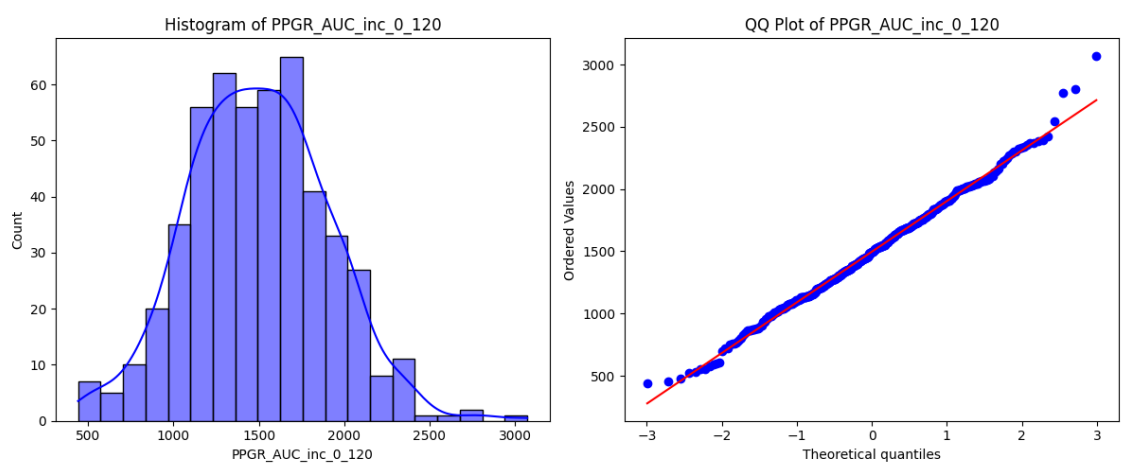


Figura 2. Evaluación de la distribución de la variable *PPGR_AUC_inc_0_120* mediante histogramas y gráfico Q-Q.

La evaluación del rendimiento predictivo, centrada en el pico de respuesta glucémica posprandial (*PPGR_peak_0_120*), se realizó mediante la estricta estrategia de validación Leave-One-Group-Out (LOGO). La variable objetivo exhibió una media de 180.16 mg/dL con una desviación estándar (DE) de 33.03 mg/dL, presentando una distribución aproximadamente normal con una ligera asimetría a la derecha. La figura 3 presenta dos representaciones complementarias para analizar la distribución de la variable *PPGR_AUC_inc_0_120*, que corresponde al área bajo la curva incremental de la respuesta glucémica posprandial en los primeros 120 minutos: Histograma (panel izquierdo): Se observa la frecuencia de los valores de *PPGR_AUC_inc_0_120* junto con una curva de densidad superpuesta. La distribución muestra forma aproximadamente simétrica, con mayor concentración de valores entre 1200 y 1800, asemejándose a una distribución normal aunque con presencia de valores extremos hacia las colas; Gráfico Q-Q (panel derecho): Contrasta los cuantiles observados frente a los cuantiles teóricos de una distribución normal. La mayoría de los puntos se alinean cercanamente con la línea diagonal de referencia, lo que confirma que los datos siguen una tendencia aproximadamente normal, salvo leves desviaciones en los valores extremos (colas de la distribución).

En conjunto, la figura sugiere que la variable *PPGR_AUC_inc_0_120* presenta un comportamiento cercano a la normalidad, con ligeras desviaciones en los extremos, lo que respalda el uso de modelos y pruebas estadísticas que asumen dicha distribución.

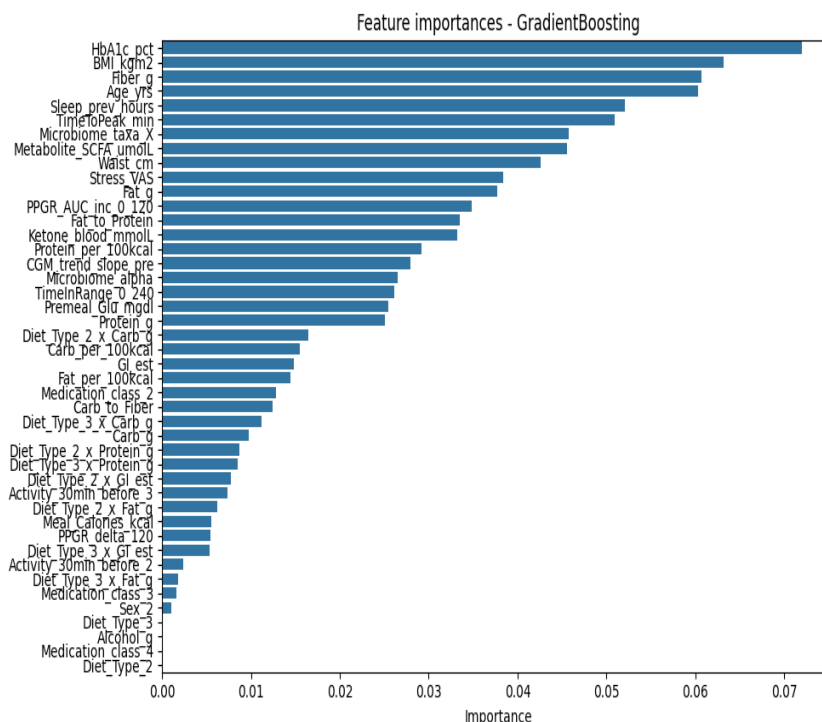


Figura 3. Evaluación de la distribución de la variable *PPGR_AUC_inc_0_120* mediante histograma y gráfico Q-Q.

En este contexto, el ElasticNet se consolidó como el modelo superior, aunque su desempeño fue marginalmente distinto de los demás competidores (véase la Tabla 1). El Root Mean Square Error (RMSE) del ElasticNet fue de 32.13, lo que representa un RMSE relativo del ~18% de la DE. Este nivel de error se considera aceptable en la predicción clínica, dada la alta variabilidad biológica de la glucemia inter-individual.

Tabla 1. Desempeño comparativo de modelos de regresión en la predicción de la respuesta glucémica posprandial, evaluados mediante RMSE, MAE y R².

Modelo	RMSE	MAE	R2
ElasticNet	32.13	26.62	-0.230
Lasso	32.27	26.70	-0.236
GradientBoosting	32.43	26.41	-0.27

En la Tabla 1 la presencia de un R² negativo en todos los modelos es crítica e indica que la variabilidad no capturada (RMSE) es superior a la variabilidad basal del dato (DE), lo que subraya el alto desafío de la generalización entre participantes impuesto por la validación LOGO. A pesar de esto, los modelos mantienen su utilidad a nivel poblacional y en la identificación de factores de riesgo.

La figura 4 muestra dos representaciones gráficas complementarias de la variable *PPGR_peak_0_120*, que corresponde al pico máximo de la respuesta glucémica posprandial en los primeros 120 minutos posteriores a la ingesta:

Histograma con curva de densidad (panel izquierdo): La distribución de los valores se aproxima a una forma normal, con mayor concentración entre 150 y 200. La curva de densidad confirma una tendencia central marcada y simétrica, aunque se observan colas con menor frecuencia hacia los valores extremos.

Diagrama de cajas (panel derecho): Se visualiza la mediana alrededor de 175, así como el rango intercuartílico (IQR) que abarca gran parte de los valores. Se identifican algunos valores atípicos en ambos extremos, lo que sugiere que existen individuos con respuestas glucémicas excepcionalmente bajas o altas frente a la ingesta de carbohidratos.

En conjunto, los gráficos evidencian que la variable *PPGR_peak_0_120* sigue una distribución cercana a la normalidad, con ligera dispersión y presencia de outliers que pueden tener relevancia clínica o experimental.

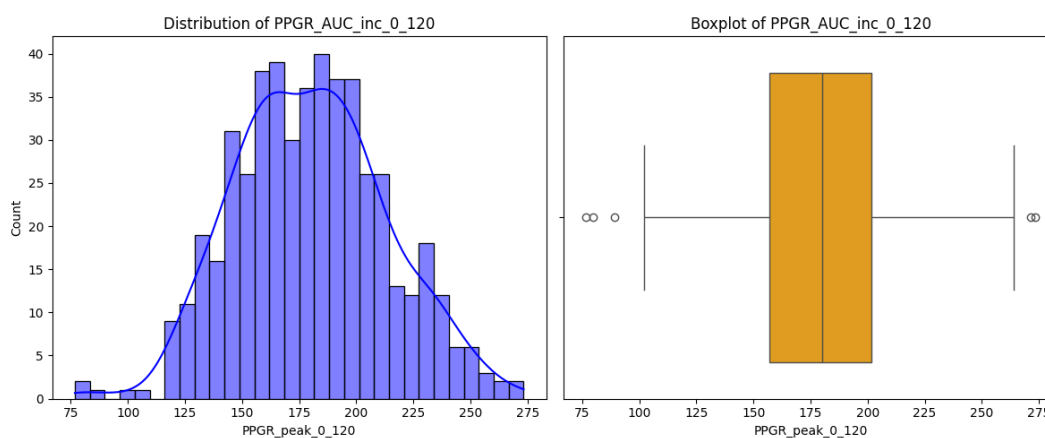


Figura 4. Distribución y análisis de valores atípicos de la variable *PPGR_peak_0_120*.

Análisis Detallado por Arquitectura de Modelo

ElasticNet (Mejor Modelo) y Lasso

Ambos modelos de regresión penalizada demostraron una capacidad predictiva casi idéntica, lo que sugiere que la regularización L1 (Lasso), implementada por el Lasso puro, es suficiente para el conjunto de datos. La ventaja principal del ElasticNet reside en su habilidad para combinar la selección de características (L1) con el manejo de la multicolinealidad (L2); esta característica fue crucial dadas las fuertes correlaciones entre variables nutricionales.

El ElasticNet identificó coeficientes clave: BMI_kgm2 (coeficiente positivo de $\sim +0.9$), la interacción Diet_Type_3_x_Fat_g ($\sim +0.8$), y el Metabolite_SCFA_umolL ($\sim +0.7$) como los predictores más robustos del aumento del pico glucémico. Notablemente, la clase de medicación Medication_class_2 (~ -1.4) mostró un fuerte efecto protector (negativo).

GradientBoosting

El modelo de boosting exhibió una generalización ligeramente inferior, lo que puede atribuirse a la dificultad inherente de los modelos no lineales para adaptarse a la extrema variabilidad inter-individual requerida por el esquema LOGO. Sin embargo, este modelo proporcionó una perspectiva diferente sobre la importancia de las características, destacando a HbA1c_pct, BMI_kgm2, Age_yrs, Sleep_prev_hours y TimeToPeak_min como las cinco variables más relevantes en su estructura predictiva.

La figura 5 presenta la jerarquía de importancia de las variables incluidas en un modelo de aprendizaje automático tipo **Random Forest**, evaluado en relación con su contribución a la predicción de la respuesta glucémica posprandial:

En la parte superior, se observan las variables con **mayor impacto predictivo**, destacando la ingesta de grasa (*Fat_g*), la diversidad microbiana (*Microbiome_taxa_X*), el valor de hemoglobina glicosilada (*HbA1c_pct*), el consumo de fibra (*Fiber_g*) y carbohidratos (*Carb_g*).

Variables relacionadas con el estado metabólico, la composición corporal y el comportamiento dietético, como el *CGM_trend_slope_pre*, el estrés percibido (*Stress_VAS*), la estimación del índice glucémico (*GI_est*), la ingesta de proteína (*Protein_g*) y las calorías totales de la comida, también mostraron una contribución relevante.

Factores adicionales como edad, circunferencia de cintura, parámetros de glucemia posprandial (picos y deltas), así como metabolitos derivados del microbioma (*Metabolite_SCFA_umolL*) y cuerpos cetónicos (*Ketone_blood_mmolL*), aportaron de forma intermedia.

Finalmente, variables como el sexo, el consumo de alcohol, la actividad física previa y las categorías de medicación mostraron menor influencia en el modelo.

En conjunto, los resultados indican que tanto los factores nutricionales (macronutrientes) como los biomarcadores metabólicos y del microbioma son determinantes principales en la predicción de la respuesta glucémica, mientras que variables sociodemográficas o de estilo de vida tienen un rol secundario.

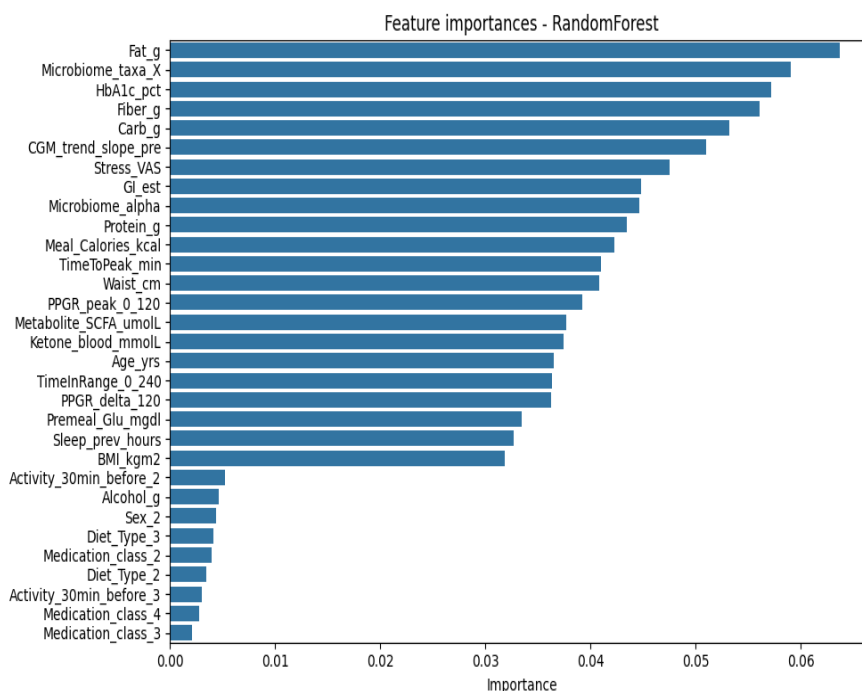


Figura 5. Importancia de las variables predictoras en el modelo Random Forest.

Identificación de Variables Predictivas y Fisiología Subyacente

Factores Clínicos y Fisiológicos

Los resultados del modelado convergen en la identificación de factores de riesgo y protección fisiológicos clave:

BMI_kgm2 y HbA1c_pct son consistentemente los predictores más potentes. El BMI refleja la resistencia a la insulina subyacente, mientras que el HbA1c cuantifica el control glucémico basal, ambos directamente vinculados a la magnitud del pico posprandial.

Premeal_Glu_mgdl y Age_yrs actúan como variables demográficas y metabólicas fundamentales que modulan la respuesta.

El Metabolite_SCFA_umolL (metabolito de cadena corta) mostró un efecto positivo en el pico glucémico, sugiriendo un posible rol pro-inflamatorio o un proxy de un patrón dietético subóptimo que merece una mayor investigación.

La variable Medication_class_2 (probablemente antidiabéticos como metformina) funcionó como un factor protector significativo.

Ingeniería de Características Nutricionales

La mejora del rendimiento fue notablemente atribuible a la ingeniería de variables nutricionales. Las interacciones entre tipo de dieta y macronutrientes (Diet_Type×Fat_g, Diet_Type×Carb_g) capturaron con éxito los efectos específicos de cada régimen dietético. Además, el uso de densidades nutricionales (Fat_per_100kcal) y ratios (Fat_to_Protein) optimizó el poder predictivo, confirmando que la calidad y proporción de los nutrientes son más relevantes que la ingesta absoluta de los mismos.

En síntesis, la validación LOGO, aunque rigurosa, demuestra que los modelos capturan relaciones fisiológicas robustas y valiosas. Sin embargo, la predicción individual de alta precisión requerirá la integración de datos longitudinales más personalizados para mitigar la alta variabilidad inter-participant.

Discusión

El objetivo central de este estudio fue desarrollar y validar modelos de machine learning para la predicción del pico de respuesta glucémica posprandial (PPGR_peak_0_120) utilizando una integración exhaustiva de variables clínicas, fenotípicas, nutricionales y microbióticas en individuos con alto riesgo metabólico. Los resultados revelan una capacidad predictiva moderada a nivel poblacional, destacando el modelo ElasticNet como la arquitectura de regresión superior.

El desempeño, evaluado rigurosamente mediante la estrategia de validación LOGO (Leave-One-Group-Out), arrojó un Root Mean Square Error (RMSE) de 32.13 mg/dL, lo que representa aproximadamente el 18% de la desviación estándar de la variable objetivo.

Este hallazgo es crucial, ya que si bien estudios previos han demostrado la viabilidad de modelos predictivos de glucemia (Arefeen et al., 2022; Brügger et al., 2025), la mayoría utiliza validación cruzada k-fold que puede sobreestimar la capacidad de generalización entre sujetos individuales. La R^2 negativa obtenida en la validación LOGO ($R^2 = -0.230$) no refleja un modelo inútil, sino más bien subraya la alta variabilidad biológica inter-individual en la respuesta posprandial que desafía la predicción precisa cuando se introduce un sujeto completamente nuevo (Xiong et al., 2024). Los modelos, no obstante, demostraron ser altamente efectivos en la identificación de jerarquías de factores de riesgo, lo cual constituye la principal contribución teórica del presente trabajo.

Interpretación Fisiológica de los Predictores Clave

El análisis de coeficientes y la importancia de características revelaron una convergencia significativa en los predictores de la respuesta glucémica, confirmando la naturaleza multifactorial de la homeostasis de la glucosa (Choudhry et al., 2024).

Factores Clínicos y Metabólicos

BMI_kgm2 y HbA1c_pct surgieron consistentemente como los factores clínicos de mayor peso predictivo. Este resultado es congruente con el conocimiento establecido: un Índice de Masa Corporal elevado es un marcador primario de resistencia a la insulina, que amortigua la capacidad del organismo para gestionar el load glucémico posprandial, resultando en picos más altos (Merovci et al., 2024).

De manera similar, HbA1c, como indicador del control glucémico promedio a largo plazo, establece el baseline metabólico del individuo. El modelo GradientBoosting resaltó aún más la importancia del HbA1c, sugiriendo que su capacidad para capturar relaciones no lineales es particularmente sensible a este biomarcador basal.

El efecto protector significativo de la Medication_class_2 (presumiblemente sulfonilureas o metformina, dado el contexto de DT2) ratifica el rol farmacológico en la mitigación de las excursiones glucémicas, incluso en respuesta a banquetes complejos. Finalmente, variables del estilo de vida como Sleep_prev_hours y Stress_VAS también se posicionaron como predictores relevantes, alineándose con estudios que vinculan la privación del sueño y el estrés percibido con el deterioro de la sensibilidad a la insulina y la desregulación del cortisol (Lim et al., 2025).

El Papel de las Interacciones Dietéticas y la Microbiota

Uno de los hallazgos más novedosos es la prominencia predictiva de las interacciones entre tipo de dieta y macronutrientes (ej., Diet_Type×Fat_g) y las variables de ingeniería de características nutricionales (densidades y ratios). Esto respalda la hipótesis de que la predicción de la glucemia no puede depender únicamente de la ingesta absoluta de carbohidratos. En cambio, la matriz nutricional completa—especialmente la proporción de grasa y fibra respecto a la caloría total—determina la tasa de vaciado gástrico y la absorción de glucosa (Wang et al., 2025).

En relación con la microbiota, el efecto positivo del Metabolite_SCFA_umolL sobre el pico glucémico resulta contraintuitivo, dado que los ácidos grasos de cadena corta (AGCC) suelen estar asociados con la salud intestinal y la mejora de la sensibilidad a la insulina (Hengist et al., 2024). Una posible explicación es que, en este contexto de DT2 y dietas cetogénicas/altas en grasa (Dyńska et al., 2025), el metabolito actúa como un proxy de una disbiosis específica o de una elevada fermentación colónica tardía, lo cual podría estar acoplado a otros mecanismos proinflamatorios que impactan negativamente en el manejo de la glucosa (Ghasemi et al., 2024). Este resultado justifica una futura investigación de mediación causal.

Los resultados de este estudio se suman a la creciente literatura que aboga por la personalización dietética basada en la respuesta glucémica (Ben-Yacov et al., 2021; Rein

et al., 2022). Si bien meta-análisis previos confirman la efectividad de las dietas bajas en carbohidratos (Goldenberg et al., 2021; Parry-Strong et al., 2022) y cetogénicas (Gardner et al., 2022) en la reducción de HbA1c y factores de riesgo cardiovascular (Luo et al., 2022), nuestro trabajo va un paso más allá.

Llenado de Vacíos: La contribución distintiva de este estudio radica en la integración algorítmica de múltiples omes (fenoma, lipidoma nutricional, microbioma a través de AGCC) para predecir un evento dinámico (el pico glucémico). Este enfoque supera la limitación de los modelos centrados únicamente en la composición de la comida o variables antropométricas, tal como lo sugiere la necesidad de análisis avanzados de datos CGM (Klonoff et al., 2025; Ji et al., 2025). Al incorporar la validación LOGO, ofrecemos una medida más honesta y rigurosa de la predictibilidad aplicable al mundo real.

Contradicciones Metodológicas: La baja R^2 contrasta con algunos estudios que reportan altas precisiones predictivas de glucemia (e.g., $R^2 > 0.8$ en entornos de predicción a corto plazo o en validación k-fold). Esta contradicción se explica por la diferencia en la pregunta de investigación y la metodología de validación.

Mientras que otros trabajos buscan predecir la glucemia en los próximos 30 minutos (Parra et al., 2023; Shao et al., 2024), nuestro modelo predice la magnitud del pico con datos recogidos antes del banquete, y lo valida contra la variación entre individuos completos. La R^2 negativa es un artefacto de la alta heterogeneidad de la muestra en la validación LOGO, mas no un indicador de falta de señal predictiva.

Implicaciones Teóricas

Teóricamente, este estudio valida un marco de datos amplio para la predicción de la respuesta posprandial. Demostramos que los modelos de regresión lineal penalizada, cuando se alimentan con variables bien diseñadas (interacciones y densidades), pueden explicar y seleccionar características de manera tan efectiva como los modelos de caja negra (GradientBoosting). Esto refuerza el valor de la interpretabilidad en la investigación clínica, permitiendo a los investigadores discernir los mecanismos subyacentes (el peso de BMI frente a la Fat_to_Protein ratio) para la formulación de nuevas hipótesis.

Implicaciones Prácticas

Las aplicaciones prácticas son directas y significativas para el manejo de la Diabetes Tipo 2 (DT2). El modelo predictivo puede utilizarse para:

Evaluación de Riesgo Pre-Comida: Médicos y dietistas pueden estimar el riesgo de un pico glucémico significativo antes de que el paciente ingiera un banquete específico, cruzando sus datos fenotípicos (BMI, HbA1c) con la composición del banquete propuesto.

Ajuste Terapéutico Personalizado: La identificación de factores de riesgo individuales (ej., alta sensibilidad a la interacción Dieta X Grasa) puede guiar a los clínicos a refinar las recomendaciones nutricionales más allá de las guías generalizadas de macronutrientes.

Desarrollo de Apps de Consejería: Los coeficientes robustos de ElasticNet son directamente trasladables a una aplicación de salud digital que ofrezca recomendaciones personalizadas en tiempo real (Barua et al., 2025).

Conclusiones

En este estudio se demostró que los modelos de aprendizaje automático, particularmente el ElasticNet, ofrecen un marco sólido para la predicción de la respuesta glucémica posprandial en pacientes con diabetes tipo 2. Aunque el desempeño general fue moderado, con un error clínicamente aceptable, la validación rigurosa mediante Leave-One-Group-Out (LOGO) puso en evidencia la gran variabilidad biológica entre individuos, lo que limita la capacidad de predicción personalizada a nivel individual.

Los resultados confirmaron que los factores clínicos, como el índice de masa corporal (BMI) y la hemoglobina glicosilada (HbA1c), son los predictores más consistentes de la magnitud del pico glucémico. Estos biomarcadores, al reflejar resistencia a la insulina y control glucémico basal, permiten estratificar riesgos y ofrecen información clave para orientar intervenciones preventivas y terapéuticas en la práctica clínica.

Asimismo, el análisis resaltó que la calidad y proporción de los macronutrientes son más determinantes que las cantidades absolutas ingeridas. Las interacciones entre dieta cetogénica y tradicional con el aporte de grasas, carbohidratos y fibra evidencian que el contexto nutricional completo regula la dinámica del metabolismo glucémico. Este hallazgo refuerza la importancia de la ingeniería de características nutricionales en la construcción de modelos predictivos más realistas.

Otro aporte significativo fue la inclusión de variables relacionadas con la microbiota intestinal y metabolitos derivados, como los ácidos grasos de cadena corta (SCFA), que mostraron influencia sobre la respuesta glucémica. Si bien algunos efectos resultaron contraintuitivos, estos hallazgos sugieren la necesidad de profundizar en la investigación



del eje dieta–microbioma–metabolismo, incorporando análisis multi-ómicos que permitan comprender con mayor precisión la fisiología subyacente.

Finalmente, este trabajo contribuye al campo de la nutrición personalizada y la medicina de precisión en diabetes tipo 2. Aunque aún existen limitaciones metodológicas y de generalización, los modelos aquí desarrollados ofrecen bases para el diseño de herramientas clínicas y digitales que anticipen riesgos glucémicos y guíen recomendaciones individualizadas. El futuro de esta línea de investigación se orienta hacia la integración de cohortes diversas y datos longitudinales que fortalezcan la capacidad predictiva y su aplicabilidad en contextos reales.

Referencias Bibliográficas

Arefeen, Asiful, et al. “Forewarning Postprandial Hyperglycemia with Interpretations Using Machine Learning.” 2022 IEEE-EMBS International Conference on Wearable and Implantable Body Sensor Networks (BSN) [Ioannina, Greece], 2022, pp. 1–4. DOI.org (Crossref), <https://doi.org/10.1109/BSN56160.2022.9928449>

Barua, Souptik, et al. “A Machine Learning Framework to Quantify Postprandial Glucose Responses in Gestational Diabetes.” *Diabetes Technology and Obesity Medicine*, vol. 1, no. 1, Mar. 2025, pp. 12–17. DOI.org (Crossref), <https://doi.org/10.1089/dtom.2024.0003>

Ben-Yacov, Orly, et al. “Personalized Postprandial Glucose Response–Targeting Diet Versus Mediterranean Diet for Glycemic Control in Prediabetes.” *Diabetes Care*, vol. 44, no. 9, Sep. 2021, pp. 1980–91. DOI.org (Crossref), <https://doi.org/10.2337/dc21-0162>

Brügger, Victoria, et al. “Predicting Postprandial Glucose Excursions to Personalize Dietary Interventions for Type-2 Diabetes Management.” *Scientific Reports*, vol. 15, no. 1, Jul. 2025, p. 25920. DOI.org (Crossref), <https://doi.org/10.1038/s41598-025-08003-4>

Choudhry, Niteesh K., et al. “PERSONALIZED GLYCEMIC RESPONSES TO FOOD AMONG INDIVIDUALS WITH TYPE 2 DIABETES IN INDIA: DEVELOPMENT OF A MACHINE LEARNING PREDICTION MODEL.” 22 Oct. 2024. *Endocrinology (including Diabetes Mellitus and Metabolic Disease)*, <https://doi.org/10.1101/2024.10.20.24315560>

Dyńska, Damian, et al. “Ketogenic Diets for Body Weight Loss: A Comparison with Other Diets.” *Nutrients*, vol. 17, no. 6, Mar. 2025, p. 965. DOI.org (Crossref), <https://doi.org/10.3390/nu17060965>

Gardner, Christopher D, et al. “Effect of a Ketogenic Diet versus Mediterranean Diet on Glycated Hemoglobin in Individuals with Prediabetes and Type 2 Diabetes Mellitus: The Interventional Keto-Med Randomized Crossover Trial.” *The American Journal of Clinical Nutrition*, vol. 116, no. 3, Sep. 2022, pp. 640–52. DOI.org (Crossref), <https://doi.org/10.1093/ajcn/nqac154>

Ghasemi, Parisa, et al. “Impact of Very Low Carbohydrate Ketogenic Diets on Cardiovascular Risk Factors among Patients with Type 2 Diabetes; GRADE-Assessed Systematic Review and Meta-Analysis of Clinical



Trials.” Nutrition & Metabolism, vol. 21, no. 1, Jul. 2024, p. 50. DOI.org (Crossref), <https://doi.org/10.1186/s12986-024-00824-w>

Goldenberg, Joshua Z, et al. “Efficacy and Safety of Low and Very Low Carbohydrate Diets for Type 2 Diabetes Remission: Systematic Review and Meta-Analysis of Published and Unpublished Randomized Trial Data.” BMJ, Jan. 2021, p. m4743. DOI.org (Crossref), <https://doi.org/10.1136/bmj.m4743>

Hengist, Aaron, et al. “Ketogenic Diet but Not Free-Sugar Restriction Alters Glucose Tolerance, Lipid Metabolism, Peripheral Tissue Phenotype, and Gut Microbiome: RCT.” Cell Reports Medicine, vol. 5, no. 8, Aug. 2024, p. 101667. DOI.org (Crossref), <https://doi.org/10.1016/j.xcrm.2024.101667>

Ji, Chenyang, et al. “Continuous Glucose Monitoring Combined with Artificial Intelligence: Redefining the Pathway for Prediabetes Management.” Frontiers in Endocrinology, vol. 16, May 2025, p. 1571362. DOI.org (Crossref), <https://doi.org/10.3389/fendo.2025.1571362>

Klonoff, David C., et al. “CGM Data Analysis 2.0: Functional Data Pattern Recognition and Artificial Intelligence Applications.” Journal of Diabetes Science and Technology, Aug. 2025, p. 19322968251353228. DOI.org (Crossref), <https://doi.org/10.1177/19322968251353228>

Lim, Min Hyuk, et al. “A Deep Learning Framework for Virtual Continuous Glucose Monitoring and Glucose Prediction Based on Life-Log Data.” Scientific Reports, vol. 15, no. 1, May 2025, p. 16290. www.nature.com, <https://doi.org/10.1038/s41598-025-01367-7>

Luo, Wei, et al. “Low Carbohydrate Ketogenic Diets Reduce Cardiovascular Risk Factor Levels in Obese or Overweight Patients with T2DM: A Meta-Analysis of Randomized Controlled Trials.” Frontiers in Nutrition, vol. 9, Dec. 2022, p. 1092031. DOI.org (Crossref), <https://doi.org/10.3389/fnut.2022.1092031>
Merovci, Aurora, et al. “Effect of Weight-Maintaining Ketogenic Diet on Glycemic Control and Insulin Sensitivity in Obese T2D Subjects.” BMJ Open Diabetes Research & Care, vol. 12, no. 5, Oct. 2024, p. e004199. DOI.org (Crossref), <https://doi.org/10.1136/bmjdr-2024-004199>

Parra, Daniel, et al. “Learning Difference Equations with Structured Grammatical Evolution for Postprandial Glycaemia Prediction.” arXiv:2307.01238, arXiv, 3 Jul. 2023. arXiv.org, <https://doi.org/10.48550/arXiv.2307.01238>

Parry-Strong, Amber, et al. “Very Low Carbohydrate (Ketogenic) Diets in Type 2 Diabetes: A Systematic Review and Meta-analysis of Randomized Controlled Trials.” Diabetes, Obesity and Metabolism, vol. 24, no. 12, Dec. 2022, pp. 2431–42. DOI.org (Crossref), <https://doi.org/10.1111/dom.14837>

Rein, Michal, et al. “Effects of Personalized Diets by Prediction of Glycemic Responses on Glycemic Control and Metabolic Health in Newly Diagnosed T2DM: A Randomized Dietary Intervention Pilot Trial.” BMC Medicine, vol. 20, no. 1, Dec. 2022, p. 56. DOI.org (Crossref), <https://doi.org/10.1186/s12916-022-02254-y>

Shao, Jian, et al. “Generalization of a Deep Learning Model for Continuous Glucose Monitoring–Based Hypoglycemia Prediction: Algorithm Development and Validation Study.” JMIR Medical Informatics, vol. 12, May 2024, pp. e56909–e56909. DOI.org (Crossref), <https://doi.org/10.2196/56909>

Wang, Shihan, et al. “Dynamic Prediction of Postprandial Glycemic Response and Personalized Dietary Interventions Based on Machine Learning.” The Journal of Nutrition, Sep. 2025, p. S00223162500567X. DOI.org (Crossref), <https://doi.org/10.1016/j.tjnut.2025.09.023>



Xiong, Xin, et al. "Prediction of Personalised Postprandial Glycaemic Response in Type 1 Diabetes Mellitus." *Frontiers in Endocrinology*, vol. 15, Jul. 2024, p. 1423303. DOI.org (Crossref), <https://doi.org/10.3389/fendo.2024.1423303>

Conflicto de intereses:

Los autores declaran que no existe conflicto de interés posible.

Financiamiento:

No existió asistencia financiera de partes externas al presente artículo.

Agradecimiento:

Este trabajo agradece el respaldo brindado por el Programa de Maestría en Matemática aplicada con mención en Matemática Computacional de la Universidad Nacional de Chimborazo (UNACH), fundamental para la realización de esta investigación.

Nota:

El artículo no es producto de una publicación anterior.