

Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), Noviembre-Diciembre 2025,
Volumen 9, Número 6.

https://doi.org/10.37811/cl_rcm.v9i6

MODELADO PREDICTIVO DEL ENVEJECIMIENTO USANDO MACHINE LEARNING: CUANTIFICACIÓN DEL RIESGO MULTIFACTORIAL DE DIABETES, ESTRÉS Y DISFUNCIÓN FAMILIAR PARA JÓVENES Y ADULTOS EN LA UNIVERSIDAD DE CUNDINAMARCA

**PREDICTIVE MODELING OF AGING USING MACHINE
LEARNING: QUANTIFICATION OF MULTIFACTORIAL RISK
FOR DIABETES, STRESS, AND FAMILY DYSFUNCTION IN
YOUNG ADULTS AND ADULTS AT THE UNIVERSITY OF
CUNDINAMARCA**

Elkin Oswaldo Forero Soto
Universidad de Cundinamarca Docente

María Victoria Rojas Martínez
Universidad de Cundinamarca Docente

Adriana Hernández Bustos
Universidad de Cundinamarca Docente

Héctor Fabio Cruz Cuellar
Universidad de Cundinamarca Docente

Diego Armando López Vargas
Universidad de Cundinamarca Docente

DOI: https://doi.org/10.37811/cl_rcm.v9i6.22062

Modelado Predictivo del Envejecimiento usando Machine Learning: Cuantificación del Riesgo Multifactorial de Diabetes, Estrés y Disfunción Familiar para jóvenes y adultos en la Universidad de Cundinamarca

Elkin Oswaldo Forero Soto¹eforero@ucundinamarca.edu.co<https://orcid.org/0000-0001-8388-5139>Universidad de Cundinamarca Docente
Colombia**María Victoria Rojas Martínez**mariavictoriarojas@ucundinamarca.edu.co<https://orcid.org/0000-0003-1637-4800>Universidad de Cundinamarca Docente
Colombia**Adriana Hernández Bustos**ahernandezb@ucundinamarca.edu.co<https://orcid.org/0000-0001-8146-5413>Universidad de Cundinamarca Docente
Colombia**Héctor Fabio Cruz Cuellar**hfabiocruz@ucundinamarca.edu.co<https://orcid.org/0000-0001-9532-9260>Universidad de Cundinamarca Docente
Colombia**Diego Armando López Vargas**diegoalopez@ucundinamarca.edu.co<https://orcid.org/0000-0003-3765-8351>Universidad de Cundinamarca Docente
Colombia

RESUMEN

En la actualidad el proceso de envejecimiento cobra un valor significativo y los actuales avances de las técnicas de *Machine Learning* hacen que la aplicación de modelos como la regresión logística permitan diagnosticar posibles problemas en adultos mayores. Este estudio se propuso validar la eficacia de la Regresión Logística Múltiple para cuantificar el riesgo multifactorial asociado al proceso de envejecimiento, centrándose en el impacto de variables sociodemográficas y de estilo de vida sobre la probabilidad de desarrollar Diabetes, Estrés y Disfunción Familiar en jóvenes y adultos universitarios. La metodología se implementó mediante un programa en computador desarrollado en Python, donde se emplearon métricas de rendimiento como el Área Bajo la Curva (AUC) para evaluar la capacidad de discriminación del modelo, y se recurrió a los Odds Ratios (OR) para establecer la fuerza de asociación de cada predictor. Los resultados demostraron consistentemente que la Regresión Logística es una metodología de alto rendimiento (con $AUC > 0.93$) y clínicamente interpretable para pronosticar estas condiciones. Estos hallazgos reafirman la utilidad del *Machine Learning* para la detección y estratificación temprana del riesgo, proporcionando una herramienta esencial para la formulación de políticas de salud pública proactivas en el ámbito universitario.

Palabras clave: regresión logística; envejecimiento; diabetes; estrés; disfunción familiar

¹ Autor principal

Correspondencia: eforero@ucundinamarca.edu.co

Predictive Modeling of Aging using Machine Learning: Quantification of Multifactorial Risk for Diabetes, Stress, and Family Dysfunction in Young Adults and Adults at the University of Cundinamarca

ABSTRACT

The aging process has gained significant value in contemporary society, and current advancements in Machine Learning techniques make the application of models such as logistic regression essential for diagnosing potential problems in older adults. This study aimed to validate the efficacy of Multiple Logistic Regression for quantifying the multifactorial risk associated with the aging process, focusing on the impact of socio-demographic and lifestyle variables on the probability of developing Diabetes, Stress, and Family Dysfunction in young adults and adults attending university. The methodology was implemented using a computer program developed in Python, where performance metrics like the Area Under the Curve (AUC) were employed to evaluate the model's discrimination capacity, and Odds Ratios (OR) were used to establish the strength of association for each predictor. The results consistently demonstrated that Logistic Regression is a high-performance methodology (with $AUC > 0.93$) and is clinically interpretable for prognosticating these conditions. These findings reaffirm the utility of Machine Learning for the early detection and risk stratification, providing an essential tool for formulating proactive public health policies within the university setting.

Keywords: Machine Learning; Logistic Regression; Aging; Diabetes; Stress; Family Dysfunction

*Artículo recibido 8 noviembre 2025
Aceptado para publicación: 15 diciembre 2025*



INTRODUCCIÓN

La trayectoria del envejecimiento en contextos académicos y laborales requiere un enfoque proactivo en la salud de la población joven y adulta. El envejecimiento prematuro o patológico representa una amenaza al rendimiento y la productividad, manifestándose no solo a través del deterioro físico, sino también mediante la fragilidad de la estabilidad mental y la disfunción en el soporte social. La investigación demuestra que el aislamiento social y la soledad tienen un impacto significativo en la salud mental, predisponiendo a los individuos a condiciones como la depresión según Hoogendijk et al., (2020). Este estudio aborda la necesidad de una perspectiva más integral utilizando *Machine Learning* para cuantificar el riesgo de otros factores.

Por otra parte, investigaciones recientes han demostrado que la salud percibida juega un papel crucial en la calidad de vida, influenciada por factores sociodemográficos y económicos, y es fundamental para asegurar el bienestar de la población. Esto se traduce en la necesidad de implementar intervenciones adaptadas que promuevan estilos de vida saludables y un enfoque proactivo hacia la salud mental y física entre los jóvenes y adultos, ya que una percepción positiva de salud está relacionada con un mejor bienestar general (Salazar-Martínez, 2022). Debido a lo anterior, es imperativo validar modelos predictivos que integren técnicas de Machine Learning, buscando generar dinámicas que permitan entender la salud y calidad de vida de esta población.

La relevancia de este estudio radica en su potencial para abordar problemas críticos de salud pública al aplicar modelos predictivos en el contexto del envejecimiento de la población universitaria. A medida que los jóvenes y adultos enfrentan diversos desafíos, desde la carga académica hasta los efectos del sedentarismo, comprender la relación entre estos factores y la salud es fundamental. La capacidad de los modelos de regresión logística y otros métodos de *Machine Learning* para predecir problemas de salud puede facilitar intervenciones tempranas y dirigidas, que a su vez pueden mejorar la calidad de vida y el rendimiento académico de los estudiantes Gzar et al. (2022). Con un enfoque multidimensional que integra variables sociodemográficas y de estilo de vida, este estudio ofrece un modelo valioso para la formulación de políticas educativas y de salud que prioricen el bienestar integral de la población universitaria.



La investigación contemporánea ha demostrado que las técnicas más avanzadas pueden superar a los modelos tradicionales, lo que resalta la necesidad de innovaciones en el análisis de la salud (Sihombing et al., 2023). Implementar estos enfoques no solo contribuirá a mejorar la salud de los estudiantes, sino que también proporcionará un marco útil para la recolección y el análisis de datos en tiempo real, lo cual es vital en un contexto que continúa evolucionando debido a factores externos como la pandemia de COVID-19. Este estudio, por lo tanto, tiene implicaciones más amplias para la salud pública, sugiriendo un camino hacia una atención más personalizada y eficaz para los jóvenes adultos en entornos académicos.

En otros estudios la predicción de resultados de salud complejos ha avanzado significativamente gracias a la integración del *Machine Learning* en la epidemiología y la geriatría, superando las limitaciones de los modelos estadísticos tradicionales (He et al., 2025). La investigación global se ha enfocado en predecir la trayectoria del envejecimiento exitoso (*Successful Aging*), que se define bajo un marco multidimensional que incluye la ausencia de enfermedad crónica, alta función cognitiva y compromiso social activo (Zhang et al., 2025). Estudios recientes han empleado algoritmos avanzados como XGBoost y Redes Neuronales Artificiales (ANN) para construir modelos predictivos, identificando factores cruciales como el estado de las enfermedades crónicas (hipertensión, enfermedad renal) como los predictores más influyentes del envejecimiento exitoso (Zhang et al., 2025). Estos modelos, al integrar variables sociodemográficas, clínicas y de estilo de vida, han logrado altas métricas de rendimiento (AUC cercanos a 0.78), demostrando que las técnicas de *Machine Learning* son herramientas valiosas para la detección temprana y la gestión de la salud poblacional (Zaccheus et al., 2024).

En América Latina, la aplicación de técnicas de inteligencia artificial y *Machine Learning* en el contexto del envejecimiento, aunque incipiente, está creciendo, enfocándose principalmente en desafíos de enfermedades crónicas y el deterioro cognitivo. Dada la limitada disponibilidad de bases de datos estandarizadas en la región, los estudios a menudo se centran en evaluar los riesgos de salud en poblaciones específicas (Navarro et al., 2021). Por ejemplo, estudios de corte transversal en Colombia han explorado la relación entre condiciones demográficas, salud física (como la Diabetes) y salud mental (síntomas depresivos) en adultos mayores urbanos, destacando la necesidad de intervenciones que



atiendan la complejidad multifactorial del bienestar en entornos de rápido envejecimiento (Gómez-Restrepo et al., 2020). La relevancia regional de este estudio radica en que, al emplear el modelado predictivo con Regresión Logística, contribuye a cerrar la brecha metodológica existente en la cuantificación del riesgo y se alinea con las recomendaciones de integrar la IA y el *Machine Learning* para optimizar la atención y rehabilitación de los adultos mayores en países como México (Jarillo-Silva, 2025).

Este estudio se fundamenta en el paradigma de la Inferencia Bayesiana y la Teoría de la Clasificación de *Machine Learning*. La Regresión Logística es inherentemente un modelo de clasificación que estima la probabilidad de que una observación pertenezca a una clase binaria (ej., tener Diabetes, Sí/No), basándose en una función sigmoide (Wu et al., 2021). Su eficacia como algoritmo de *Machine Learning* reside en su interpretabilidad, ya que los Odds Ratios (OR) obtenidos permiten cuantificar el impacto multiplicativo de los factores de riesgo con claridad clínica (Zhang et al., 2025). La Teoría Bayesiana complementa este enfoque al ofrecer un marco para actualizar las probabilidades de riesgo a medida que se incorporan nuevos datos, siendo especialmente útil en la predicción de resultados complejos y evolutivos como la mortalidad asociada a la Diabetes, donde los modelos de Edad-Período-Cohorte Bayesiana han demostrado una alta capacidad predictiva (Wu et al., 2021).

El objetivo principal de este trabajo es validar el modelo de regresión logística mediante la integración de técnicas de *Machine Learning* para determinar el impacto de las variables socioeconómicas y hábitos de vida que pueden afectar la calidad de vida para las personas mayores, proporcionando un análisis exhaustivo que cuantifique el impacto de la diabetes, el estrés y la disfunción familiar, evaluando la capacidad predictiva del modelo mediante métricas de alto rendimiento.

Si bien es cierto que estudio no ahonda sus resultados en la calidad del envejecimiento, si aporta un análisis de la relación que tienen los factores socioeconómicos con las variables como diabetes, estrés y disfunción familiar.

Para este estudio se trabajó con la población de jóvenes y adultos la Universidad de Cundinamarca en su sede Girardot, una ciudad con 326 metros a nivel del mar, con clima cálido y tropical, con aproximadamente 150.000 habitantes y separada de Bogotá por 120 kilómetros



METODOLOGÍA

Gracias al uso intensivo de datos numéricos, modelos matemáticos y estadísticos para probar el método de regresión logística multifactor, este estudio es predominantemente de enfoque cuantitativo.

Por otro lado, el uso de la regresión logística busca establecer relaciones entre las características de los individuos y la probabilidad de un resultado específico, lo cual determina que esta investigación es de tipo predictivo y su diseño es observacional porque los investigadores no manipulan ninguna variable o intervención en los participantes, por el contrario, observan y miden las variables sociodemográficas, hábitos de vida y los resultados en Diabetes, Estrés y Disfunción Familiar, como se presentan en la población de la Universidad de Cundinamarca.

En la fase de diseño se decidió que la población a analizar estaría conformada en dos grandes grupos, estudiantes y administrativos, los primeros deberían ser estudiantes de primer semestre de ingeniería de software, enfermería e ingeniería ambiental y los segundos todo el personal administrativo de la Universidad de Cundinamarca Sede Girardot.

Luego de esto se diseñó un instrumento de recolección de datos basados en el test de Findrisc, Apgar familiar, tamización de depresión y estrés y también se tuvo en cuenta la caracterización sociodemográfica de los individuos de la muestra.

Estos datos se recopilaron y se limpiaron en un micrositio web desarrollado en Python donde se almacena toda la data de este estudio. Esta plataforma comunica los datos a una aplicación de escritorio denominada "Relopen" (Regresion Logistica Pronostico envejecimiento) que utiliza librerías de Python como: pandas, numpy, sklearn y statsmodels entre otros. Con esta aplicación se realizó un análisis exploratorio de datos, basado en la estadística descriptiva con el fin de comprender el comportamiento de la información recolectada y determinar la presencia de *outlier* en estos datos, luego fueron codificadas las variables categóricas como: rol (estudiantes o administrativo), sexo, estrato y zona residencial.

Luego de esto y teniendo en el análisis exploratorio de datos se construyó un modelo predictivo de datos, basados en la regresión logística multifactor, para en este tipo de investigación, donde se busca establecer relaciones entre las características de los individuos y la probabilidad de un resultado específico, como es el caso del envejecimiento (Li et al. 2025).



El uso de la regresión logística como modelo predictivo se debe a su alta capacidad de establecer relaciones entre características individuales y la probabilidad de resultados binarios, que para este estudio seria: desarrollar o no diabetes (score_neurodiab), tener una buena o mala dinámica familiar (score_dinafami) y tener o no tener estrés (score_estres). Todos estos valores fueron calculados de acuerdo a la tabla 1.

Tabla 1

Criterios para cálculos de Riesgo binomial para diabetes, dinámica familiar y estrés

Criterio	Verdadero	Falso
'score_neurodiab' >= 13	Si tiene riesgo de diabetes	No tienen riesgo de diabetes
'score_dinafami' <= 12	Posibilidades de tener una buena dinámica familiar	Posibilidades de tener una mala dinámica familiar
'score_estres' >= 14	Si tiene posibilidades de tener estrés	No tiene posibilidades de tener estrés

Nota: Criterios definido por los investigadores para el modelamiento usando la regresión logística

Evaluación del modelo de regresión logística:

Para la regresión logística como lo menciona Costa y Borges (2024) las métricas como Accuracy y el AUC son claves para evaluar la capacidad del modelo especialmente para determinar el desbalance entre las variables. Por lo que para este estudio fue evaluado mediante las siguientes métricas:

Sensibilidad: Para corroborar la validación del modelo de regresión logística también se incluyó el concepto de Curva ROC que permite ver la sensibilidad del modelo a medida que aumenta la tasa de Falsos Positivos y Curva AUC o área bajo la curva que representa la capacidad de discriminación del modelo, es decir, la capacidad que tiene este modelo de distinguir entre casos positivos y negativos

Interpretación: La aplicación método de regresión logística ha sido validado en numerosas investigaciones por su habilidad para generar Odds Ratios que simplifican la interpretación clínica de los resultados, incluyendo su fácil adaptación en el contexto de análisis multifactoriales para la evaluación en salud pública.

Para analizar la magnitud de estos efectos se aplicaron medidas como Accuracy y el área bajo la curva AUC importantes para determinar la capacidad del modelo para predecir los efectos en diabetes, estrés



y disfunción familiar. También se empleó la razón de Momios reconocida como Odds Ratios (OR) para cuantificar la fuerza que tiene una variable en la presencia o ausencia del riesgo de tener o no diabetes, estrés y disfunción familiar.

Las variables independientes son aquellas que se presume que tienen un efecto sobre el envejecimiento. Estas pueden incluir aspectos demográficos como la edad, el sexo, el nivel socioeconómico, y el estado civil, así como factores de estilo de vida como el nivel de actividad física, la dieta, el consumo de alcohol, y el tabaquismo (Filho et al., 2023). Además, se podrían incluir factores psicológicos como el estrés percibido y el estado de salud mental, que también son influyentes en el proceso de envejecimiento y deterioro cognitivo; sin embargo, no se cuenta con una referencia específica en este contexto que valide esta afirmación (Gutiérrez-Monsalve et al., 2021).

Por otro lado, las *variables dependientes* en este contexto son aquellas que representan los resultados del envejecimiento y que se desean predecir mediante los modelos estadísticos y de machine learning. Estas pueden incluir indicadores de salud como la calidad de vida, la presencia de enfermedades crónicas (por ejemplo, diabetes o hipertensión), la funcionalidad física, o puntajes en pruebas de cognición (González et al. (2025), Muscedere et al. (2025), Ray et al. (2025).). Estrategias de *Machine Learning* pueden mejorar la identificación de patrones en estas variables dependientes que, a su vez, facilitan una comprensión más profunda de los factores que contribuyen al envejecimiento y su impacto en la salud de la población.

La relevancia de estas variables es fundamental, ya que entender cómo diferentes factores independientes pueden predecir cambios en las variables dependientes permitirá desarrollar modelos más precisos que no solo describen el fenómeno del envejecimiento, sino que también pueden ofrecer pronósticos valiosos que informen tratamientos, políticas de salud y programas de intervención dirigidos a mejorar la salud y la calidad de vida de los jóvenes y adultos en el contexto universitario (Sota, 2022). Lo anterior establece un camino claro hacia la implementación de estrategias de predicción más efectivas, una meta clave para el desarrollo de este estudio.

RESULTADOS Y DISCUSIÓN

El modelado predictivo del riesgo de Diabetes, Estrés, y Disfunción Familiar se realizó mediante Regresión Logística Múltiple, basándose en la interpretación a través de los Odds Ratios (OR). Este



modelo predictivo "Sin Fuga" se integraron variables como: IMC_PUNTOS (índice de masa corporal), EDAD_PUNTOS (edad) y DIFICUVIDA_PUNTOS (Con qué frecuencia ha podido controlar las dificultades de su vida) demostraron una buena capacidad predictiva significativamente superior a los modelos base.

El modelo de predicción de Diabetes SIN FUGA alcanzó un Área Bajo la Curva (AUC) de 0.9351, indicando un alto grado de pronóstico, tal como lo muestra la Tabla 2. El modelo de Estrés SIN FUGA mostró un AUC de 0.9332, lo que también confirma una excelente capacidad de discriminación. El modelo de Dinámica Familiar SIN FUGA obtuvo un AUC de 0.9362, siendo el de mayor rendimiento, lo que sugiere que las variables utilizadas son altamente predictivas del resultado. Estos resultados validan la idoneidad de la Regresión Logística y los *Odds Ratios* para el análisis de la proyección del envejecimiento, al abordar la probabilidad de experimentar resultados binarios

Tabla 2

Resultados del modelo de regresión logística múltiple diabetes, estrés y Dinámica Familiar sin Fuga

Resultados Modelo de Regresión Logística Múltiple Diabetes SIN FUGA	Resultados ESTRES Modelo de Regresión Logística Múltiple SIN FUGA	Resultados DINAMICA FAMILIAR Modelo de Regresión Logística Múltiple SIN FUGA
Variables Predictoras (X): - IMC_PUNTOS - EDAD_PUNTOS - SEXO_MASCULI - ESTRATO - ZONA_URBANA Intercepto (b): -2.3838	Variables Predictoras (X): - DIFICUVIDA_PUNTOS - ENFADADO_PUNTOS - SEXO_MASCULI - ESTRATO - ZONA_URBANA Intercepto (b): -0.0650	Variables Predictoras (X): - FAMILIAFEC_PUNTOS - SEXO_MASCULI - ESTRATO - ZONA_URBANA Intercepto (b): 0.4763
Pendientes (w) para cada característica: - IMC_PUNTOS : 0.9977 - EDAD_PUNTOS : 0.8749 - SEXO_MASCULI : -0.6758 - ESTRATO : -0.2024 - ZONA_URBANA : -0.9681 Odds Ratio (OR) para cada característica: - IMC_PUNTOS : 2.7120 - EDAD_PUNTOS : 2.3987 - SEXO_MASCULI : 0.5088 - ESTRATO : 0.8168 - ZONA_URBANA : 0.3798	Pendientes (w) para cada característica: - DIFICUVIDA_PUNTOS : 1.5735 - ENFADADO_PUNTOS : 0.3542 - SEXO_MASCULI : -0.1391 - ESTRATO : 0.0532 - ZONA_URBANA : -0.2473 Odds Ratio (OR) para cada característica: - DIFICUVIDA_PUNTOS : 4.8233 - ENFADADO_PUNTOS : 1.4250 - SEXO_MASCULI : 0.8702 - ESTRATO : 1.0546 - ZONA_URBANA : 0.7809	Pendientes (w) para cada característica: - FAMILIAFEC_PUNTOS : -1.4399 - SEXO_MASCULI : -0.3418 - ESTRATO : 0.2450 - ZONA_URBANA : 0.4969 Odds Ratio (OR) para cada característica: - FAMILIAFEC_PUNTOS : 0.2369 - SEXO_MASCULI : 0.7105 - ESTRATO : 1.2776 - ZONA_URBANA : 1.6436

Nota: Resultado generados por el software Relopen en el componente "regrelogi_diabetes v2.py", "regrelogi_estres.py", "regrelogi_dinafami.py"

Para la cuantificación del Riesgo para Diabetes y Estrés se analizaron los *Odds Ratios* (OR) permitiendo identificar los factores de riesgo más influyentes en cada resultado binario.

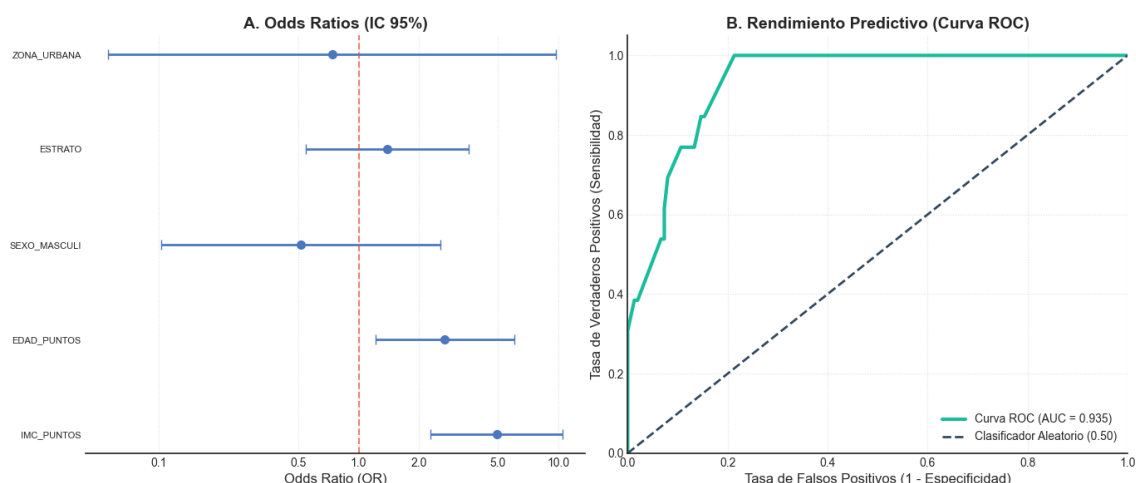
Es así que el riesgo de la diabetes esta influenciada por: El IMC_PUNTOS (OR = 2.71) y la EDAD_PUNTOS (OR = 2.40) quienes fueron los predictores más fuertes. Un aumento de 1 punto en el IMC se asoció con un aumento del 171.20% en la probabilidad relativa de tener diabetes

aproximadamente 2.71 veces), mientras que un aumento de 1 punto en la edad implicó un aumento del 139.87% aproximadamente 2.40 veces.

Grafica 1

Modelo Predictivo Diabetes Sin Fuga

Modelo Predictivo Diabetes SIN FUGA Odds Ratio (IC 95%)



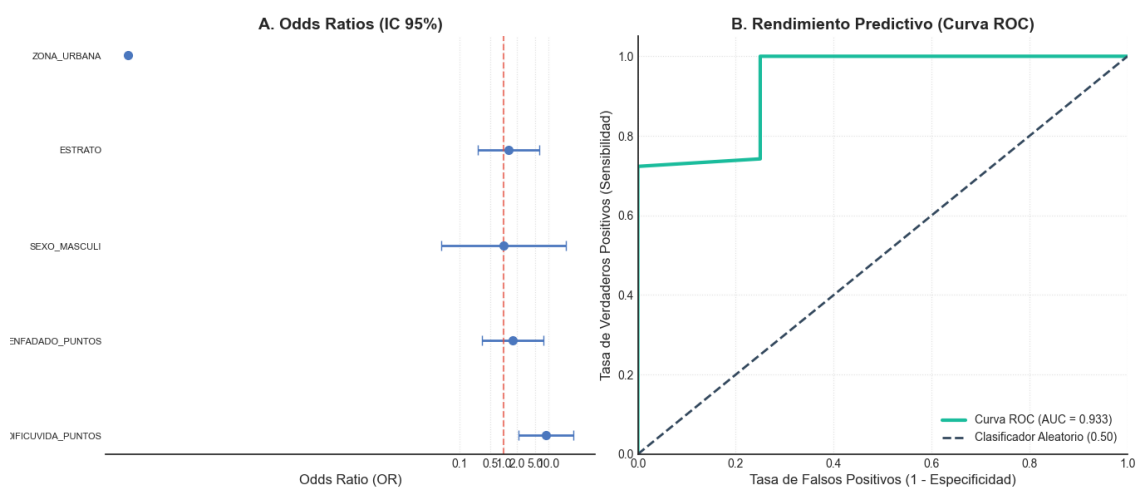
Nota: Resultado generados por el software Relopen

Por otro lado, el riesgo psicológico en cuanto al estrés esta referida para: La DIFICUVIDA_PUNTOS (OR = 4.82) es el predictor más fuerte y significativo. Un aumento de 1 punto en la dificultad de vida percibida incrementó las OR de tener Estrés en un 382.33% aproximadamente 4.82 veces. El ENFADADO_PUNTOS (OR = 1.43) también fue un factor de riesgo moderado.

Grafica 2

Modelo Predictivo Estrés Sin Fuga

Modelo Predictivo Estres SIN FUGA Odds Ratio (IC 95%)



Nota: Resultado generados por el software Relopen



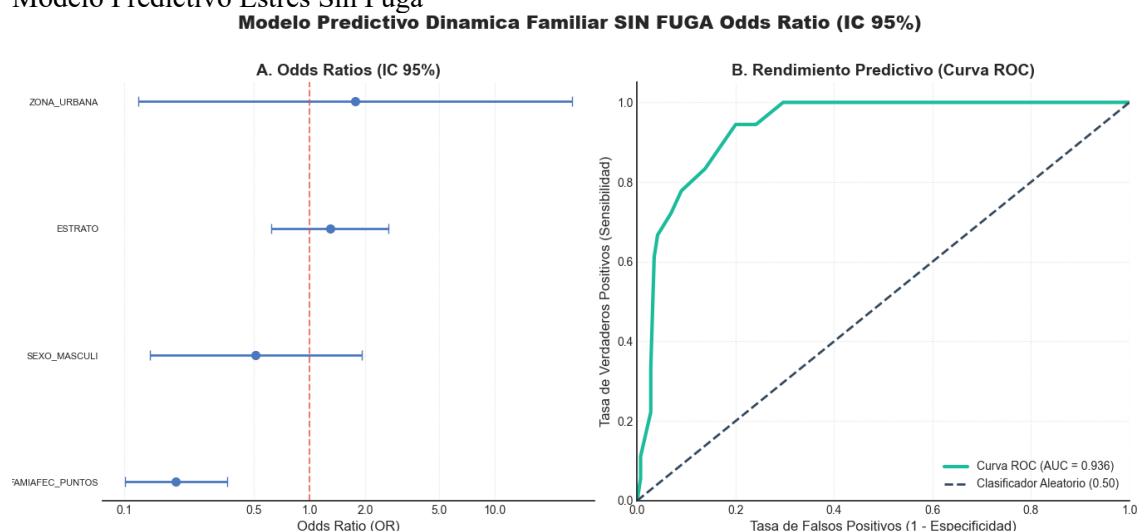
También se puede ver en los resultados como vivir en una Zona Urbana ($OR = 0.38$) y ser Hombre ($OR = 0.51$) mostraron un efecto protector contra la Diabetes, en cambio para el modelo de Estrés, la variable ESTRATO ($OR = 1.05$) tuvo un efecto muy bajo.

El impacto de la Dinámica Familiar y Factores Sociodemográficos reveló que la calidad del afecto familiar es el principal factor protector, mientras que las variables sociodemográficas (Zona Urbana y Estrato) actúan como factores de riesgo:

La variable Afecto Familiar tiene un aumento de 1 punto en la variable FAMIAFEC_PUNTOS (afecto familiar) y se asoció con una disminución del 76.31% en las OR de tener una Mala Dinámica Familiar.

Grafica 3

Modelo Predictivo Estrés Sin Fuga



Nota: Resultado generados por el software Relopen

Algunas variables de Riesgo Sociodemográficas como son: ZONA_URBANA ($OR = 1.64$) y ESTRATO ($OR = 1.28$) tuvieron los índices más altos de riesgo. Lo que se puede asociar que vivir en una Zona Urbana tiene con un aumento del 64.36% en las OR de disfunción familiar, siendo el factor de riesgo socio-demográfico más fuerte. Adicionalmente, un aumento de 1 unidad en el Estrato aumentó las OR en un 27.76%.

De manera general se puede establecer que el factor más determinante que apunta a un deterioro en la calidad del envejecimiento es el riesgo metabólico y físico, es representado principalmente por la

Diabetes, en donde se destaca por su valor dominante: El IMC ($OR \approx 2.71$) y la Edad ($OR \approx 2.40$) son en grupo, los mayores multiplicadores de las OR en la Diabetes.

Dado que la Diabetes acelera el daño vascular, cognitivo y renal, su predicción alta por estos factores implica un alto riesgo de envejecimiento patológico. Un factor clave en esta relación es el daño vascular asociado con la diabetes. Valverde et al. (2022) Este daño vascular incrementa el riesgo de eventos adversos como infartos y accidentes cerebrovasculares, que son comunes en las poblaciones ancianas con diabetes.

Adicionalmente se puede establecer que vivir en una Zona Urbana ($OR \approx 0.38$) y ser Hombre ($OR \approx 0.51$) para este estudio muestra un efecto protector contra la Diabetes, lo cual podría estar relacionado con diferencias en el acceso a la atención médica preventiva o comportamientos específicos dentro de la muestra.

Si bien se ha encontrado que el entorno urbano influye en los estilos de vida y en la prevención de la diabetes, factores como el acceso a servicios de salud, educación y oportunidades económicas suelen ser mejores en estas áreas en comparación con las zonas rurales (Wu et al., 2021).

En cuanto a la variable Estrés, subraya que las dificultades de vida percibidas y la gestión de emociones negativas son barreras significativas para un envejecimiento saludable.

En el Estrés las variables dominantes son: Dificultades de Vida ($OR \approx 4.82$) y Enfado ($OR \approx 1.43$) quienes en conjunto impulsan el Estrés. Existe evidencia con población envejecida que se señala que la inflamación sistémica y el envejecimiento vascular son componentes clave en el deterioro cognitivo, y que la inflamación crónica puede actuar como puente entre estrés, comorbilidades y cambio cognitivo, con implicaciones para la calidad de vida en la vejez según Ragonnaud y Biragyn (2021).

En términos más generales el estrés, se manifiesta a través de variables tales como las dificultades de vida y el enfado, elementos que pueden incrementar significativamente los niveles de estrés en las personas. Este vínculo ha sido documentado, resaltando cómo el estrés crónico puede llevar a serias consecuencias en la salud metabólica y psicológica, particularmente en poblaciones vulnerables Ortiz et al. (2022) Crielaard et al., (2021).



Para la matriz entorno Dinámica Familiar se destacan FAMIAFEC (afecto familiar), y SEXO_MASCULI donde se puede analizar que:

La variable FAMIAFEC_PUNTOS, por cada aumento de 1 punto en esta variable, los OR tiene una Mala Dinámica Familiar disminuyen en un 76.31% $((1 - 0.2369) \times 100\%)$. Esto sugiere que la funcionalidad o afecto familiar es el predictor más fuerte de una buena dinámica.

Por otro lado, según este estudio ser Hombre se asocia con una disminución del 28.95% en las OR de tener una Mala Dinámica Familiar, en comparación con ser Mujer.

Adicionalmente de acuerdo a los OR vivir en Zona Urbana o en Estratos más altos aumenta significativamente la Mala Dinámica Familiar.

Esto sugiere que la mayor complejidad o las presiones socioeconómicas y laborales asociadas a estos entornos pueden crear una disfunción dentro del núcleo familiar, erosionando el soporte vital necesario para un envejecimiento saludable. La relación entre las dificultades económicas y la dinámica familiar ha sido objeto de atención, ya que el estrés financiero tiende a incrementar las tensiones dentro del hogar. Un estudio señala que la presión económica puede agravar los conflictos familiares y afectar la calidad de las relaciones interpersonales, lo que a su vez repercute negativamente en el bienestar de los adultos mayores dependientes de su red familiar. Este fenómeno cuenta con respaldo de la literatura que documenta la influencia de las relaciones interpersonales y del apoyo social en el bienestar de los adultos mayores (Manchana, 2023).

CONCLUSIONES

Los investigadores validaron el modelo de *Machine Learning* de regresión logística en la población de la Universidad de Cundinamarca Girardot, con el fin de determinar la capacidad de predicción y el impacto que tiene los hábitos de vida y las variables socioeconómicas en la calidad de vida de los adultos mayores.

El uso de la Regresión Logística puesta a la luz de AUC y Odds Ratios ha demostrado ser una metodología de *Machine Learning* adecuada para el análisis de la proyección del envejecimiento, donde se aborda como la probabilidad de experimentar resultados binarios. Su aplicación en toda la población Universitaria puede ofrecer una base sólida para la formulación de políticas de salud pública dirigidas a mejorar la calidad de vida y el bienestar integral de la población universitaria.



Sin embargo, la aplicación del modelo de regresión logística debe reconocer una serie de desafíos. Uno de los principales retos radica en la dependencia de la linealidad en la relación entre las variables predictoras (X) y el logaritmo de las probabilidades (Log-Odds) del resultado. Si la verdadera relación subyacente entre los factores de riesgo IMC y Edad con el riesgo de Diabetes es intrínsecamente no lineal, el modelo podría ofrecer una subestimación o sobreestimación del riesgo en los extremos de las distribuciones de datos.

Adicionalmente, este estudio afronta como reto la trascendencia y aplicación a largo plazo del modelo propuesto. El objetivo es asegurar que el modelo se mantenga relevante y preciso, lo cual requiere aumentar la base de datos de la población universitaria y nutrir continuamente el modelo con una mayor cantidad de datos.

Para futuros avances de este estudio se recomienda implementar y comparar algoritmos de *Machine Learning* más robustos contra el desbalance de clases, como XGBoost o Random Forest, para potencialmente mejorar la Sensibilidad del modelo.

REFERENCIAS BIBLIOGRÁFICAS

- Christodoulou et al. "A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models" *Journal of clinical epidemiology* (2019). [10.1016/j.jclinepi.2019.02.004](https://doi.org/10.1016/j.jclinepi.2019.02.004) <https://pubmed.ncbi.nlm.nih.gov/30763612/>
- Costa, O. y Borges, L. (2024). Artificial Intelligence in primary care: Intelligent risk predicting platform for non-communicable chronic diseases. *Procedia Computer Science*. Volumen 239. Pages 2243-2250. ISSN 1877-0509. <https://doi.org/10.1016/j.procs.2024.06.415>
- Crielaard, L., Nicolaou, M., Sawyer, A., Quax, R., & Stronks, K. (2021). Understanding the impact of exposure to adverse socioeconomic conditions on chronic stress from a complexity science perspective. *BMC Medicine*, 19(1). <https://doi.org/10.1186/s12916-021-02106-1>
- Filho, C. C. F., Andrade, M. H. M. L., Nunes, J. A. R., Wipff, J. K., Hignight, D., Rios, E. F., ... & Hignight, K. (2023). Breeding for drought tolerance in perennial ryegrass (*lolium perenne* l.) and tall fescue (*lolium arundinaceum* [schreb.] darbysh.) by exploring genotype by environment by management interactions. *Grassland Research*, 2(1), 22-36. <https://doi.org/10.1002/glr2.12045>



- Gómez-Restrepo, C., Tovar, J., & Hernández, M. (2020). Relación entre calidad del sueño y salud en adultos mayores de cinco ciudades colombianas. *Revista de Salud Pública*, 22(4), 1-15. Recuperado de <https://revistas.uis.edu.co/index.php/revistasaluduis/article/view/16435>
- González, J. C. A., Taub-Krivoy, A., Sierra-Peña, J. A., & Lizarazo, J. G. (2025). Determining prognostic factors in the treatment of primary hemifacial spasm: clinical outcomes and complications. a literature review. *World Neurosurgery*: X, 25, 100406. <https://doi.org/10.1016/j.wnsx.2024.100406>
- Gutiérrez-Monsalve, J. A., Garzón, J., & Cardona, Á. M. S. (2021). Factores asociados al rendimiento académico en estudiantes universitarios. *Formación Universitaria*, 14(1), 13-24. <https://doi.org/10.4067/s0718-50062021000100013>
- Gzar, D.A., Mahmood, A.M., Abbas, M.K. (2022). A comparative study of regression machine learning algorithms: Tradeoff between accuracy and computational complexity. *Mathematical Modelling of Engineering Problems*, Vol. 9, No. 5, pp. 1217-1224. <https://doi.org/10.18280/mmep.090508>
- He, N., Xu, X., Wang, H., & He, F. (2025). Revisiting successful aging through a machine learning approach to quantifying the influence of chronic diseases. *Scientific reports*, 15(1), 40206. <https://doi.org/10.1038/s41598-025-24154-w>
- Hoogendijk, E. O., Smit, A. P., Dam, C. v., Schuster, N. A., Breij, S. d., Holwerda, T. J., ... & Andrew, M. K. (2020). Frailty combined with loneliness or social isolation: an elevated risk for mortality in later life. *Journal of the American Geriatrics Society*, 68(11), 2587-2593. <https://doi.org/10.1111/jgs.16716>
- Jarillo-Silva, A., Domínguez-Ramírez, O., & Cruz-Tolentino, J. (2025). Uso de la tecnología en temas de rehabilitación y atención en el adulto mayor con deterioro cognitivo en México. *Revista de Gerontología y Geriatria*, 14(2), 45-60. https://www.researchgate.net/publication/395453356_Uso_de_la_tecnologia_en_temas_de_rehabilitacion_y_atencion_en_el_adulto_mayor_con_deterioro_cognitivo_en_Mexico
- Li, B., Li, W., & Wan, C. (2025). Cross-disciplinary risk prediction for muscle weakness and physical decline in older adults: a machine learning model integrating social determinants of health and



- clinical characteristics. *Journal of International Medical Research*, 53(9).
<https://doi.org/10.1177/03000605251379211>
- Manchana, V. (2023). Interpersonal relationships, subjective health, psychological well-being, and quality of life among older adults in south india: evidence from a population-based study. *Journal of Education and Health Promotion*, 12(1). https://doi.org/10.4103/jehp.jehp_735_22
- Navarro, D., Suárez, M., & López, R. (2021). Inteligencia artificial aplicada a la salud en América Latina: Luces y sombras. *Nueva Sociedad*, 294, 1-15. Recuperado de <https://periodico.unal.edu.co/uploads/2023/junio/PeriodicoUNAL-229.pdf>
- Ortiz, R., Gilgoff, R., & Harris, N. B. (2022). Adverse childhood experiences, toxic stress, and trauma-informed neurology. *JAMA Neurology*, 79(6), 539.
<https://doi.org/10.1001/jamaneurol.2022.0769>
- Ragonnaud, E. and Biragyn, A. (2021). Gut microbiota as the key controllers of “healthy” aging of elderly people. *Immunity & Ageing*, 18(1). <https://doi.org/10.1186/s12979-020-00213-w>
- Salazar-Martínez, R. (2022). Coaching Comunicativo Grupal en Adultos Mayores para el Envejecimiento Saludable. *Areté*, <https://arete.iberro.edu.co/article/view/art22107>
- Sihombing, P. R., Budiantono, S., Arsani, A. M., Aritonang, T. M., & Kurniawan, M. A. (2023). Comparison of Regression Analysis with Machine Learning Supervised Predictive Model Techniques . *Jurnal Ekonomi Dan Statistik Indonesia*, 3(2), 113-118.
<https://doi.org/10.11594/jesi.03.02.03>
- Sota, J. V. Q. (2022). Diseño de interfaces de sistemas interactivos utilizando técnicas de machine learning: una revisión del diseño y la usabilidad. *Interfases*, (016), 202-214.
<https://doi.org/10.26439/interfases2022.n016.6028>
- Valverde, D. M., Brenes, I. A., & Muñoz, M. P. (2022). Mecanismos de envejecimiento renal. *Revista Medica Sinergia*, 7(5), e804. <https://doi.org/10.31434/rms.v7i5.804>
- Wu, X., Du, J., Li, L., Cao, W., & Sun, S. (2021). Edad-Período-Cohorte Bayesiana Predicción de la mortalidad por diabetes tipo 2 Enfermedad renal diabética en China: Un estudio de modelado. *Frontiers in Endocrinology*, 12, 767263. <https://doi.org/10.3389/fendo.2021.767263>



Zaccheus J, Atogwe V, Oyejide A and Salau AO. (2024) Towards successful aging classification using machine learning algorithms <https://doi.org/10.12688/f1000research.138608.2>

Zhang, Z., Peng, C., Li, Z., Li, J., Li, Y., Pan, Y., Liu, R., & Chen, X. (2025). Development and validation of a successful aging prediction model for older adults in China based on health ecology theory.

Frontiers in Public Health, 13, Article 1595540. <https://doi.org/10.3389/fpubh.2025.1595540>

