



UNIVERSIDAD PERUANA
CAYETANO HEREDIA

**PONTIFICIA UNIVERSIDAD CATÓLICA DEL
PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA**

**UNIVERSIDAD PERUANA CAYETANO
HEREDIA
FACULTAD DE CIENCIAS E INGENIERÍA**

***Modelo de clasificación de desnutrición aguda para
niños de 6 a 59 meses a partir de variables
climáticas en Perú***

Tesis para optar el Título Profesional de Ingeniero Biomédico

que presenta:

Luis Carlos Revilla Dominguez

Asesor:

Dr. Gabriel Carrasco Escobar

Lima, 2025

Jurado calificador

Presidente: Mg. Felix Alonso Agapito Cordova


Vocal: Paulo Camilo Alberto Vela Anton

Secretario: Mg. Umbert Lewis De La Cruz Rodriguez

Feedback Studio - Google Chrome
ev.turnitin.com/app/carta/es/?s=18ro=2700599549&lang-es&ro=103&u=1165482885

turnitin 1 de 1: Uigict FACI FAVEZ
Modelo de clasificación de desnutrición aguda para niños de...

Similitud 11% Marcas de alerta



PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ
FACULTAD DE CIENCIAS E INGENIERÍA

UNIVERSIDAD PERUANA CAYETANO HEREDIA
FACULTAD DE CIENCIAS E INGENIERÍA

Modelo de clasificación de desnutrición aguda para niños de 6 a 59 meses a partir de variables climáticas en Perú

Tesis para optar el Título Profesional de Ingeniero Biomédico que presenta:
Luis Carlos Revilla Domínguez

Asesor:
Dr. Gabriel Carrasco Escobar

Lima, 2025

Página 1 de 75 14135 palabras 81%

Informe estándar
Informe en inglés no disponible Más información

11% Similitud estándar Filtros

2 Exclusiones

Fuentes
Mostrar las fuentes solapadas

#	Internet	Similitud
1	hdl.handle.net	1%
13 bloques de texto 158 palabra que coinciden		
2	doczz.es	<1%
2 bloques de texto 73 palabra que coinciden		
3	www.frontiersin.org	<1%
2 bloques de texto 61 palabra que coinciden		
4	cdn.www.gob.pe	<1%

DEDICATORIA

A mi familia, por su constante comprensión y apoyo incondicional durante toda mi vida.

A Ariana, por confiar en que todo irá bien, darme ánimos y motivación.

A mi asesor, por ser un guía y un mentor durante esta etapa de mi carrera profesional.

AGRADECIMIENTOS

Para empezar, quisiera agradecer a mi familia, especialmente a mi mamá, quien ha estado conmigo en cada una de mis decisiones, me ha apoyado y ha apostado por mí desde el primer día.

A mi asesor, Gabriel Carrasco, quien me dio la oportunidad de trabajar a su lado con un increíble equipo en Innovalab y dándome la confianza que necesitaba para encontrar mi camino en la investigación.

A mi mejor amigo, Fabian, por demostrarme que siempre se puede salir adelante.

Finalmente agradezco a Ariana, quien durante todos estos años ha sido un pilar fundamental en mi vida, con quien comparto todos mis sueños y aspiraciones.

ÍNDICE

RESUMEN.....	1
ABSTRACT	2
I. INTRODUCCIÓN.....	3
1.1. <i>Presentación e importancia del tema</i>	3
1.2. <i>Justificación de la investigación</i>	5
1.3. <i>Antecedentes</i>	7
II. HIPÓTESIS Y OBJETIVOS.....	11
2.1. <i>Hipótesis</i>	11
2.2. <i>Objetivos</i>	11
III. MATERIALES Y MÉTODOS.....	12
3.1. <i>Materiales</i>	12
3.2. <i>Método</i>	13
3.3. <i>Población y Muestra</i>	19
IV. RESULTADOS.....	20
4.1. <i>Análisis descriptivo</i>	20
4.2. <i>Validación cruzada</i>	23
4.3. <i>Validación en ventanas de tiempo</i>	26
4.4. <i>Análisis por región</i>	30
4.5. <i>Comparación con modelo sociodemográfico</i>	32
V. DISCUSIÓN	34
5.1. <i>Análisis descriptivo</i>	34
5.2. <i>Análisis comparativo de resultados</i>	35
5.3. <i>Generalización de modelos en ventanas de tiempo</i>	36
5.4. <i>Interpretabilidad de resultados</i>	38
5.5. <i>Análisis por región</i>	39
5.6. <i>Comparación con modelo sociodemográfico</i>	40
5.7. <i>Limitaciones</i>	41
VI. CONCLUSIÓN.....	42
VII. REFERENCIAS	44
ANEXOS	48

RESUMEN

La desnutrición aguda es una condición de salud que afecta a una relevante cantidad de niños menores de 5 años. Anualmente se obtienen datos climatológicos y sociodemográficos que en la literatura han demostrado tener asociación con la desnutrición aguda infantil. Este estudio propone un modelo predictivo para la presencia de la desnutrición aguda en niños de 6 meses a 5 años en el Perú usando variables climáticas. La población y muestra de estudio seleccionada fueron los niños de entre 6 meses y 5 años en el Perú, durante el periodo de 2014 a 2019 elegibles para la ENDES INEI. Se emplearon modelos de Machine Learning como Random Forest y XGBoost con validación cruzada y evaluación por ventanas de tiempo. En el entrenamiento se obtuvo una precisión y AUC en validación de hasta 93% y 78% respectivamente, aunque en la evaluación se redujeron al 92% y 69%. Se aplicó un análisis desagregado por departamentos, donde la selva obtuvo los mejores resultados en sensibilidad (87% de promedio para los años de evaluación). Asimismo, se comparó un modelo base utilizando solo variables sociodemográficas, donde los resultados fueron comparables a los obtenidos por predictores climáticos. Como conclusión, se comprueba la capacidad de modelos basados en algoritmos de ML para utilizar datos secundarios climáticos y predecir la presencia de desnutrición aguda con un enfoque diferenciado en las regiones del Perú. Futuros estudios deberían incluir la temporalidad en el análisis, lo que favorece una herramienta de toma de decisiones en salud pública.

Palabras clave: machine learning, desnutrición aguda

ABSTRACT

Acute malnutrition is a health condition that affects a significant number of children under 5 years of age. Climatological and sociodemographic data are obtained annually and have been shown in the literature to be associated with acute malnutrition in children. This study proposes a predictive model for the presence of acute malnutrition in children aged 6 months to 5 years in Peru using climatic variables. The study population and sample selected were children aged 6 months to 5 years in Peru, during the period from 2014 to 2019 eligible for the ENDES INEI. Machine Learning models such as Random Forest and XGBoost were used with cross-validation and evaluation by time windows. In training, accuracy and AUC in validation were obtained up to 93% and 78%, respectively, although in evaluation they were reduced to 92% and 69%. A disaggregated analysis by department was applied, where the jungle obtained the best results in sensitivity (87% on average for the years of evaluation). Likewise, a base model was compared using only sociodemographic variables, where the results were comparable to those obtained by climatic predictors. In conclusion, the ability of models based on ML algorithms to use secondary climatic data to predict the presence of acute malnutrition with a differentiated approach in the regions of Peru is proven. Future studies should include temporality in the analysis, which would favor a decision-making tool in public health.

Keywords: machine learning, acute malnutrition, acute malnutrition

I. INTRODUCCIÓN

1.1. Presentación e importancia del tema

La desnutrición, un desafío crítico para la salud pública, afecta negativamente el bienestar y el potencial de crecimiento de millones de niños en todo el mundo. Globalmente, se identifican tres tipos de desnutrición al utilizar índices antropométricos: aguda, crónica y global. La desnutrición aguda, a diferencia de las otras dos, es una afección que está relacionada con la disminución de la masa corporal respecto a la estatura, lo cual puede deberse a los efectos secundarios de una enfermedad infecciosa, o por una baja ingesta de calorías en la alimentación del niño. A nivel de América Latina y el Caribe la prevalencia de la desnutrición aguda es del 1,3% (aproximadamente 700 mil niños y niñas menores de cinco años) (1), y específicamente en Perú esta proporción ha sido del 1,8% en los últimos dos años, de acuerdo con el Ministerio de Salud del Perú (MINSA) (2).

El impacto económico, a nivel nacional para el país, ha sido de 10, 500 millones de dólares los últimos años, equivalente al 4.6% del Producto Bruto Interno (PBI) (2). Esto se debe a que la desnutrición aguda no solo disminuye la capacidad intelectual y de aprendizaje, sino que limita el crecimiento y el desarrollo físico, llegando incluso a la posibilidad de desencadenar enfermedades cardiovasculares debido a las deficiencias metabólicas. Pese a que existen opciones de solución viables a través de políticas públicas y estrategias, organismos como el Banco Mundial han sugerido dirigir los recursos del MINSA para identificar los puntos críticos de esta patología y prevenir los casos de desnutrición (3). Al mismo tiempo, se ha observado como factores de riesgo determinantes a la pobreza, la falta de acceso a alimentos nutritivos, educación nutricional, y vivir en zonas rurales y en la región Selva, entre otros (4,5).

Se han estudiado a lo largo de los años y en varios países diversos factores sociodemográficos como la pobreza, el acceso limitado a los alimentos, la falta de agua potable o saneamiento, demostrando que contribuyen a la desnutrición aguda. Sin embargo, existen otros factores de riesgo influyentes pero no tan considerados a la hora de analizar la desnutrición aguda como la precipitación pluvial, temperaturas extremas o mínimas (6). En este contexto, la Encuesta Demográfica y

de Salud Familiar (ENDES), realizada todos los años por el Instituto Nacional de Estadística e Informática (INEI), se encarga de recopilar variables sociodemográficas e individuales, centrada en indicadores infantiles, incluyendo mediciones que permiten aproximar la presencia de la desnutrición aguda (7).

Además de estos factores, se destaca la estacionalidad como un término clave en el estudio de la desnutrición, la cual se refiere a la variabilidad de los resultados nutricionales correlacionada con las estaciones. Esta variabilidad, influenciada por cambios en las condiciones medioambientales tales como la temperatura, las precipitaciones y la vegetación, es significativa. Investigaciones al respecto indican que la fluctuación en la incidencia de malnutrición aguda a lo largo del año puede superar incluso las variaciones observadas a través de los años (6). En Etiopía, las variaciones estacionales y agroecológicas afectan significativamente la prevalencia de anemia y desnutrición entre las madres lactantes (8). Las diferencias regionales también son evidentes en la prevalencia de anemia en Bangladesh (9).

El aprendizaje de máquina o Machine Learning (ML), comprendido como una rama de la inteligencia artificial que desarrolla algoritmos para poder aprender información y realizar predicciones o inferencias con datos nuevos, puede ser utilizado para construir modelos que analicen factores relevantes en desnutrición aguda para reconocer patrones y predecir el riesgo de desnutrición en un niño específico (4). En Perú, para el análisis de la desnutrición aguda se han empleado modelos predictores de estadística clásica como la regresión logística a partir de características y factores sociodemográficos, de salud y económicos (4). No hay a la fecha modelos de inteligencia artificial que consideren los factores medioambientales como predictores relevantes. Mientras tanto, en continentes como África, donde también existe una gran variedad de climas, varios países han empleado modelos longitudinales estadísticos de prueba t de Student de resolución temporal mensual, modelos longitudinales de regresión lineal de resolución temporal mensual, y modelos transversales con efectos fijos de regresión lineal de resolución temporal mensual (4,10).

Ante el panorama del uso de variables ambientales en modelos de ML, Google Earth Engine (GEE) aparece como una plataforma de análisis geoespacial

basada en la nube que permite acceder a una gran cantidad de datos y herramientas para el análisis de imágenes satelitales y otros tipos de datos geoespaciales, incluyendo imágenes satelitales de alta resolución, datos de vegetación, temperaturas superficiales, precipitaciones, y otros indicadores medioambientales cruciales. A través de GEE, se puede acceder a datos climatológicos a nivel mundial, como lo es el caso de TerraClimate, que recoge datos mensuales de temperatura, precipitación, vapor; o la NASA LP DAAC at the USGS EROS Center para índices de vegetación del suelo. Estos datos climáticos pueden ser de gran utilidad en el monitoreo de variables climáticas y en la prevención de picos de desnutrición en niños al permitir la obtención de datos precisos y en tiempo real sobre la temperatura, la humedad y la precipitación en diferentes regiones. Además, estos datos pueden ser analizados junto con información sobre la vegetación y la calidad del suelo para obtener una visión más completa de las condiciones ambientales en las que viven los niños (11,12).

1.2. Justificación de la investigación

La desnutrición aguda es un problema grave en el Perú y afecta especialmente a los niños menores de 5 años. Inclusive a día de hoy la Organización para la Cooperación y el Desarrollo Económicos (OCDE) continúa señalando como un problema mayor en salud pública a la desnutrición aguda y aboga por medidas para prevenirla y tratarla. Por ello, el desarrollar un modelo predictivo para su presencia permitiría identificar de manera temprana a los niños en riesgo, lo que podría mejorar su atención y tratamiento. Asimismo, podría ayudar en la toma de decisiones de gestores de salud pública para priorizar la asignación de recursos y esfuerzos para prevenir y tratar la desnutrición aguda en las zonas donde se identifique un mayor riesgo.

Desde lo teórico, el estudio se basa en el enfoque de la epidemiología y la salud pública, así como un indicador de desarrollo humano en el que se busca identificar los factores medioambientales que pueden estar relacionados con la presencia de una afección como la desnutrición aguda, y de qué manera estos factores pueden estar asociados con el riesgo de su aparición. Además, el empleo de algoritmos de ML para elaborar modelos predictivos se alinea con la tendencia contemporánea de

incorporar técnicas de inteligencia artificial en el sector salud. Esta aproximación no solo optimiza la identificación de factores de riesgo y la predicción de eventos de salud, sino que también promueve una medicina más preventiva y personalizada. Al analizar vastos conjuntos de datos, incluidos aquellos obtenidos de GEE, los modelos de inteligencia artificial pueden revelar patrones complejos y correlaciones no evidentes para el análisis humano directo. El uso de datos satelitales permite el acceso a información sin necesidad de participar en una adquisición primaria de los datos, como entrevistas a las personas, lo que reduce la carga logística para poder predecir o tomar decisiones en intervenciones de salud pública

Metodológicamente, el estudio se basa en la adquisición de imágenes satelitales de variables climáticas y el análisis de datos secundarios del INEI ENDES 2014-2019, un conjunto de datos representativo a nivel nacional sobre la situación de la desnutrición aguda en los niños menores de 5 años en el Perú. Se optó por un enfoque a nivel nacional que no se limita a un solo centro de salud, lo cual proporciona una muestra más amplia y representativa de las diversas condiciones geográficas y climáticas del Perú. Se utilizarán algoritmos de ML para identificar los factores ambientales que están relacionados con la presencia de la desnutrición aguda y desarrollar un modelo predictivo.

Científicamente, el estudio busca contribuir al conocimiento en el campo de la ingeniería biomédica, la epidemiología y la salud pública en el Perú, al desarrollar un modelo predictivo que permita identificar tempranamente a los niños en riesgo de desnutrición aguda. Además, el uso de algoritmos de aprendizaje automático permitirá explorar la relación entre los factores ambientales y la desnutrición aguda, lo que podría generar nuevas hipótesis y áreas de investigación.

Desde el punto de vista social, la desnutrición aguda es un problema grave que afecta la calidad de vida de los niños y su desarrollo futuro. Desarrollar un modelo predictivo para su presencia permitiría identificar de manera temprana a los niños en riesgo y mejorar su atención y tratamiento, lo que podría tener un impacto positivo en la salud y el bienestar de la población infantil del Perú. Además, el

estudio podría contribuir a informar políticas y programas de salud pública para prevenir y tratar la desnutrición aguda en el país.

1.3. Antecedentes

Con el propósito de abordar la compleja interacción entre variabilidad climática y salud, Berrang-Ford *et al.* realizaron una revisión sistemática que cubre el período de 2013 a 2020. Esta investigación reveló una marcada tendencia hacia estudios de impacto climático, especialmente en lo que respecta a estacionalidad, olas de calor extremo y variabilidad del tiempo. Critican la escasez de análisis en países de ingresos medios y bajos, los más afectados por el cambio climático, resaltando la promesa del ML para dilucidar estas dinámicas y sugerir estrategias mitigadoras (13).

Avanzando en esta línea de investigación, Nkiruka *et al.* arrojaron luz sobre cómo ciertos ecosistemas y patrones climáticos (precipitación, humedad, temperatura atmosférica, radiación y presión atmosférica) fomentan la proliferación de la malaria, introduciendo un innovador modelo de ML enfocado en prever su incidencia mediante el estudio de variaciones climáticas no estacionales. Este modelo destaca por su capacidad para ofrecer una herramienta preventiva contra la epidemia mediante la detección temprana, marcando un paso adelante en la gestión de la salud pública con alta precisión a partir del modelo XGBoost en países como Burkina Faso (97%), Camerún (94%), Mali (98%) y Nigeria (95%) (14). Paralelamente, la investigación de Singh *et al.* contribuye a este cuerpo de trabajo con un algoritmo de redes neuronales destinado a modelar brotes de malaria en India, utilizando variables meteorológicas (humedad relativa, temperatura y precipitación pluvial) procesadas por separado o en conjunto, aplicando técnicas como el análisis probabilístico de componentes principales (P²CA). Su estudio resalta el notable rendimiento del modelo como sistema de alerta temprana, basándose en métricas como el estadístico R y el error cuadrático (RMSE) con valores de 0.99 y 1.76 respectivamente, superando a otros algoritmos como Support Vector Regression (R = 0.44, RMSE = 92) y Random Forest (R= 0.28, RMSE = 89.57) (15).

En consonancia con estos esfuerzos, Kamana et al. Se propusieron mejorar la predicción de la malaria en China, aprovechando la sensibilidad de la enfermedad a las condiciones climáticas. Mediante el uso del modelo long short-term memory sequence-to-sequence (LSTMSeq2Seq) para los casos mensuales de malaria desde 2004 a 2016, demostraron una correlación significativa entre la incidencia de malaria y factores climáticos, logrando reducir el RMSE de las predicciones entre un 19,05% y un 33,93%, entre un 18,4% y un 33,59%, entre un 17,6% y un 26,67% y entre un 13,28% y un 21,34%, para *P. falciparum*, *P. vivax*, *P. malariae* y otros plasmodios, respectivamente, en comparación con otros modelos como XGBoost. (16).

Adicionalmente, Ochida *et al.* centraron su estudio en el dengue, evaluando el riesgo actual y futuro de brotes en Nueva Caledonia a la luz del cambio climático. Al identificar factores climáticos locales decisivos, su trabajo ofrece un enfoque metodológico replicable para otros contextos, facilitando la adaptación de estrategias de salud ante escenarios climáticos cambiantes (17). En un esfuerzo complementario, Mishra *et al.* exploraron la predicción de la propagación del dengue utilizando un modelo de regresión lineal doble, que considera variables climáticas altamente correlacionadas con la enfermedad. Sus hallazgos apuntan hacia una metodología predictiva robusta, abriendo nuevas avenidas para la prevención del dengue (18).

Asimismo, Deléglise *et al.* extendieron la aplicación del ML al ámbito de la seguridad alimentaria en África, un continente asolado por el hambre. A través de un análisis de datos heterogéneos que incluye condiciones climáticas y uso del suelo, demostraron cómo modelos avanzados pueden prever situaciones de inseguridad alimentaria con precisión ($R^2 = 0.47$ para el Food Consumption Score y $R^2 = 0.43$ para el Household Dietary Diversity Score), ofreciendo pistas vitales para el diseño de intervenciones en salud pública (19). Esta cadena de investigaciones subraya el valor inmenso del ML no solo en la comprensión, sino también en la mitigación de los impactos del cambio climático en la salud global. En cuanto a la desnutrición y ML, Siy *et al.* examinaron los factores de riesgo, comparando cuatro algoritmos en términos de precisión, sensibilidad y

especificidad. Entre todos los algoritmos, Random Forest tuvo la mejor precisión (78.55%), subrayando el potencial de estas técnicas para orientar intervenciones nutricionales, basado en gastos de los hogares, edad de los niños y de los hogares, inseguridad alimentaria y diversidad alimentaria (20). Concentrándose en la relación entre el entorno y el crecimiento infantil, Usman *et al.* identificaron la vulnerabilidad climática como un factor significativo en el retraso del crecimiento infantil, utilizando el modelo Extreme Gradient Boosting (XGBoost). Este análisis reveló la capacidad predictiva de los modelos, encontrando que un aumento de un punto porcentual (p.p.) de aumento de la pobreza puede traducirse en un aumento de 0.18 p.p. del retraso del crecimiento infantil. Además, la alta vulnerabilidad climática exacerbó el retraso del crecimiento infantil en un punto porcentual destacando la importancia de estos hallazgos para formular políticas de salud pública (21).

Explorando las posibilidades de predicción mediante Random Forest, Browne *et al.*, utilizaron este algoritmo, obteniendo una precisión que se compara favorablemente con los métodos de Deep Learning. Los autores resaltan la aplicabilidad de estos modelos para elaborar estimaciones de un conjunto de medidas de prevalencia de la desnutrición y la pobreza (potencialmente correlacionadas) utilizando datos georreferenciados, gratuitos, de libre acceso y actualizados periódicamente (22). Centrando la atención en el medio oriente, Talukder *et al.* aplicaron varios modelos, entre ellos Random Forest, para predecir la desnutrición en menores de cinco años en Bangladesh, resultando ser el más efectivo con una precisión del 68.51%, una sensibilidad del 94.66% y una especificidad del 69.76%, destacando su aplicación potencial en la identificación de desnutrición infantil (10). Siguiendo una dirección similar, Chilyabanyama *et al.* se centraron en el retraso del crecimiento en Zambia, utilizando modelos como Random Forest, que alcanzó una precisión del 79% en la predicción del retraso del crecimiento en niños menores de cinco años (23). Expandiendo el alcance geográfico a Etiopía Mekonnen *et al.* en su estudio encontraron que Random Forest fue el modelo más efectivo con un área bajo la curva (AUC) de 0.76, en comparación a modelos como Ridge Regression (0.70) y Lasso Regression (0.71).

Asimismo, la sensibilidad, especificidad y precisión de Random Forest fue de 52%, 82% y 72%, superando a todos los otros algoritmos evaluados (24).

II. HIPÓTESIS Y OBJETIVOS

2.1.Hipótesis

Las variables climáticas, como la precipitación, la temperatura y los índices de vegetación, permiten predecir la presencia o ausencia de desnutrición aguda en niños de 6 meses a 5 años en el Perú.

Variable dependiente

Presencia o ausencia de desnutrición aguda

Variables independientes

Precipitación

Temperatura

Índice de Vegetación de Diferencia Normalizada

Educación de la madre

Edad del niño

Índice de riqueza del hogar

Disponibilidad diaria del agua

Intervalo de nacimiento

Lugar de residencia

2.2.Objetivos

General

Desarrollar un modelo predictivo para la desnutrición aguda en niños de 6 meses a 5 años en el Perú, utilizando variables climáticas del periodo 2014-2019, mediante el uso de algoritmos de Machine Learning sobre el conjunto de datos ENDES 2014-2019 del INEI.

Específicos

Identificar las variables climáticas más influyentes para la desnutrición aguda en niños de 6 a 59 meses en el Perú, realizando un análisis nacional estratificado por región (costa, sierra, selva) y por grupos de edad, con el fin de determinar las diferencias en el impacto de dichas variables.

Comparar las métricas de evaluación del modelo basado en predictores climáticos con un modelo de línea base que utilice variables sociodemográficas reportadas en la literatura.

Evaluar la capacidad de generalización del modelo predictivo a través de una validación cruzada por años y estratificación regional.

III. MATERIALES Y MÉTODOS

3.1. Materiales

- Los datos utilizados de identificación por peso, talla, edad y conglomerado de los niños de 6 a 59 meses fueron obtenidos del servidor en línea del INEI, los cuales son parte de las bases de datos del ENDES. Estos datos abarcan un período de tiempo específico, desde 2014 hasta 2019.
- Los datos de variables climáticas como el Índice de vegetación de diferencia normalizada (NDVI), la precipitación, la temperatura, etc. fueron obtenidos con una resolución mensual a partir del conjunto de datos online Terraclimate y el centro de archivo y distribución de datos geoespaciales NASA LP DAAC at the USGS EROS Center.
- Repositorio de Github, donde se actualizaron los avances y se registraron cada uno de los scripts, modelos, conjuntos de datos, entre otros.
- Hardware y software: Se utilizó un equipo de computadora con ciertas especificaciones técnicas (AMD Ryzen 5 3600 6-Core Processor 3.59 GHz, memoria RAM de 16.0 GB) para llevar a cabo el procesamiento de los datos y la ejecución de los entrenamientos de modelos predictivos. Asimismo, el sistema operativo a utilizar es Windows 10 y RStudio versión 4.2.3.

3.2.Método

El presente estudio de investigación, desde el punto de vista metodológico, puede clasificarse como observacional analítico, correlacional y predictivo. Es un estudio de tipo analítico, ya que su principal objetivo fue analizar la relación entre las variables climáticas obtenidas a través de GEE y la presencia de desnutrición aguda en niños de seis meses a cinco años de edad en el Perú. Además, este estudio es observacional debido a que se basa en la observación de variables en su escenario natural sin manipulaciones experimentales. En cuanto a su enfoque correlacional, pretende establecer la relación estadística entre las variables independientes (variables climáticas) y la variable dependiente (presencia de desnutrición aguda). Finalmente, este estudio es predictivo en el sentido de que tuvo por objetivo desarrollar un modelo de aprendizaje automático para predecir la presencia de desnutrición aguda en niños a partir de variables climáticas, con el fin de identificar y prevenir tempranamente esta afección.

3.2.1. Método de la tesis

- El presente estudio se centró en el desarrollo de un modelo predictivo mediante técnicas de ML capaz de utilizar datos de variables climáticas y sociodemográficas, específicamente longitud y latitud por conglomerado, combinados con datos climáticos de TerraClimate y MODIS 006 MOD13Q1, para predecir de manera precisa y confiable la desnutrición aguda en la población infantil.

Recopilación y preparación de datos

- Datos Sociodemográficos y Climáticos: Se extrajeron variables sociodemográficas de peso, talla, edad, educación de la madre, índice de riqueza del hogar, disponibilidad diaria del agua, región, intervalo de nacimiento, lugar de residencia. La longitud y latitud reportadas en INEI ENDES se utilizaron para la vinculación y extracción de datos climáticos con resolución mensual de precipitación, temperatura (máxima y mínima), y el NDVI. No se incluyó de forma explícita una variable que indicase la presencia del fenómeno de El Niño, pues las alteraciones climáticas que este fenómeno genera en la precipitación, la temperatura, entre otros, ya se ven reflejadas en las propias variables satelitales mensuales utilizadas en el

modelo. Las variables climáticas de estadística descriptiva incluyeron el promedio anual de NDVI, la diferencia estacional de NDVI, el promedio anual de precipitación, y las temperaturas anual máxima, mínima, y la diferencia entre estas (TMAX, TMIN, TGAP). Además, se realizó un feature engineering para analizar la cantidad de meses (Ecuación 1 y 2), cantidad de meses consecutivos (Ecuación 3 y 4) y grupos de meses consecutivos (Ecuación 5 y 6) con condiciones climáticas extremas .

$$M_{high_p} = \sum_{m=1}^N I(X_m > Q_p) \quad (1)$$

$$M_{low_p} = \sum_{m=1}^N I(X_m < Q_p) \quad (2)$$

$$C_{high_p} = \max(\sum_{j=1}^k I(X_{m+j-1} > Q_p)) \quad (3)$$

$$C_{low_p} = \max(\sum_{j=1}^k I(X_{m+j-1} < Q_p)) \quad (4)$$

$$G_{high_p} = \sum_{m=1}^N I(X_m > Q_p \wedge X_{m-1} \leq Q_p) \quad (5)$$

$$G_{low_p} = \sum_{m=1}^N I(X_m < Q_p \wedge X_{m-1} \geq Q_p) \quad (6)$$

Donde:

X_m es el valor de la variable climática en el mes m .

Q_p es el percentil o umbral p de los valores mensuales

k es la longitud de una secuencia consecutiva de meses cumpliendo la condición

N es el número de meses que ingresan al cálculo

- Desbalance de Datos: Dado que la desnutrición aguda representa una fracción pequeña del total de la población estudiada (menos del 5% de las observaciones), se presenta un desbalance significativo. Para abordar este problema y mejorar la capacidad de predicción del modelo, se utilizó la técnica de SMOTE (Synthetic Minority Over-sampling Technique) (25). SMOTE es una técnica de sobre-muestreo que permite equilibrar las clases minoritarias generando ejemplos sintéticos basados en los ejemplos existentes de la clase minoritaria (niños con desnutrición aguda). Se

estableció un balance del 15% y 85% en ambas clases para reducir el sesgo del modelo hacia la clase mayoritaria.

- Manejo de Datos Faltantes y No Plausibles: Se revisó el origen y la proporción de los datos perdidos en cada una de las variables del del ENDES INEI incluidas en el análisis, así como aquellas que sirvieron como llaves para unir las variables climáticas. De acuerdo a su naturaleza y si la presencia de datos faltantes era significativa (Missing Not At Random, Missing At Random o Missing Completely At Random), se realizó la imputación de datos respectiva para cada variable (26). Para los datos no plausibles, se validó el diccionario de datos del ENDES INEI para encontrar los “casos marcados” y retirarlos de no representar una porción significativa de las observaciones.
- Manejo de Datos Temporales: Se aplicó una técnica de agregación de datos temporales para estandarizar la resolución temporal entre las variables climáticas, variables sociodemográficas y la medición de desnutrición. Esta agregación de datos temporales permitió resaltar tendencias a largo plazo, facilitando el análisis y la modelación predictiva.

3.2.2. Análisis exploratorio de los datos

- Se revisó la correlación entre variables predictoras y se visualizaron sus distribuciones temporales respecto a la desnutrición aguda. Un muestreo aleatorio jerárquico permitió evaluar la homogeneidad en las distribuciones de las variables en observaciones de niños con desnutrición aguda, originalmente menos del 5% del conjunto de datos.

3.2.3. Modelado Predictivo

- Comparación de enfoque sociodemográfico y climático: Para evaluar el impacto de las variables climáticas en la capacidad predictiva del modelo, se compararon dos enfoques distintos: uno que utiliza únicamente predictores climáticos y otro que incorpora sólo variables

sociodemográficas. Las variables sociodemográficas incluyen edad, educación de la madre, índice de riqueza del hogar, disponibilidad diaria de agua, intervalo de nacimiento y lugar de residencia, reportadas en la literatura previamente para el modelado de la desnutrición aguda (10).

- Algoritmos Utilizados: Para la construcción del modelo predictivo, se optó por utilizar los algoritmos de clasificación Random Forest y XGBoost, los cuales fueron seleccionados principalmente por la capacidad de interpretar los resultados e importancia de variables de cada modelo, así como sus características específicas para manejar conjuntos de datos con ciertas particularidades, como un número moderado de observaciones superior a las 10,000, el desbalance de clases y la complejidad de las relaciones entre variables.

Random Forest

- Random Forest es un algoritmo de ensamble basado en la combinación de múltiples árboles de decisión, lo que le permite capturar patrones complejos en los datos sin riesgo significativo de sobreajuste (27). Este modelo es altamente eficiente para manejar variables de diferentes tipos, como las variables climáticas y sociodemográficas, y es particularmente útil cuando se cuenta con datos desbalanceados, como en el caso de la desnutrición aguda, o con datos en diferentes escalas que en otros algoritmos requerirían una normalización en la base de datos. Aunque puede llegar a ser más difícil de interpretar que modelos más sencillos como la regresión logística, Random Forest permite analizar la importancia de las variables predictoras, lo cual entrega información sobre cómo toma las decisiones el modelo.

XGBoost

- XGBoost es un algoritmo de boosting que se basa en la optimización secuencial de modelos, donde cada árbol posterior corrige los errores del árbol anterior (28). Su eficacia radica en su capacidad para ajustar modelos de forma eficiente y precisa, lo que le permite

ofrecer un alto rendimiento. A diferencia de Random Forest y otros modelos de Machine Learning, ofrece una variedad de hiperparámetros que ajustan el modelo de acuerdo a las necesidades específicas del problema de regresión o clasificación. Pese a tener configuraciones que ayudan a regularizar el uso de variables, como L1 y L2, XGBoost tiene un mayor riesgo al sobreajuste de los datos en comparación a Random Forest.

- Validación y Ajuste de Hiperparámetros: Debido a la resolución temporal anual de la medición de la desnutrición aguda en ENDES INEI, se implementó una validación por ventanas de tiempo, utilizando datos de los años 2014 a N para entrenamiento y evaluando el modelo en el año N+1 (29). Asimismo, la validación por ventanas de tiempo se escogió por la capacidad de simular un escenario real de predicción de desnutrición a partir de información previa. En esta línea, se realizó una validación cruzada estratificada en el tiempo con ajuste de hiperparámetros en 5 folds de 10 modelos cada uno, entrenando un total de 50 modelos por ventana de tiempo, con un enfoque en sensibilidad, especificidad, y el score F1. Los hiperparámetros ajustados fueron: número mínimo de muestras por hoja (*min_n*), número de árboles (*trees*), tasa de aprendizaje (*learning_rate*), número de variables consideradas por división (*mtry*), profundidad del árbol (*tree_depth*). Además, se analizó la importancia de las variables en dos grupos, enfocándose en variables descriptivas y en el conteo de meses en condiciones climáticas críticas.

3.2.4. Evaluación de modelos

- Se evaluaron las métricas de sensibilidad (Ecuación 7), especificidad (Ecuación 8), donde *TP* son los Verdaderos Positivos (True Positive), *TN* son los Verdaderos Negativos (True Negative), *FP* son los Falsos Positivos (False Positive), y *FN* son los Falsos Negativos (False Negative). Asimismo, se calculó la métrica del área bajo la curva (ROC-AUC) para medir la capacidad del modelo para discriminar entre clases positivas y

negativas (Ecuación 9). El coeficiente de concordancia Kappa se empleó para medir el acuerdo entre la clasificación del modelo predictivo y el ground truth, usando la precisión observada p_o y la precisión esperada p_e (Ecuación 10). La importancia de las variables se determinó para identificar aquellas más significativas en los modelos.

$$\frac{TP}{TP + FN} \quad (7)$$

$$\frac{TN}{TN + FP} \quad (8)$$

$$AUC = \int_0^1 TPr(FPr)dFPr \quad (9)$$

$$\kappa = \frac{p_o - p_e}{1 - p_e} \quad (10)$$

3.2.5. Análisis por región

- Dada la variabilidad de resultados entre regiones, se realizó un análisis específico para cada departamento de Perú, revisando veinticuatro mapas en las tres regiones naturales para obtener métricas de sensibilidad y precisión detalladas. Adicionalmente, se realizó una segmentación de las edades de los niños de acuerdo a la recomendación de la OMS para la vigilancia nutricional. La edad se categorizó en cinco bloques (6–11, 12–23, 24–35, 36–47 y 48–59 meses) a fin de capturar hitos biológicos y de alimentación diferenciados.

3.2.6. Herramientas

- El lenguaje de programación R versión 4.4.3 y la colección de paquetes Tidy Models versión 1.3.0 fueron instrumentales para el modelado predictivo, proporcionando estructuras organizadas para la preparación de datos, selección de variables, ajuste de modelos, y evaluación de rendimiento.

3.2.7. Validación metodológica

- La sensibilidad y precisión del modelo fueron métricas clave para validar su exactitud, mientras que la confiabilidad de las predicciones se analizó mediante el AUC y el coeficiente Kappa.

3.3.Población y Muestra

De acuerdo a lo descrito anteriormente, se establece como población de estudio a los niños de entre seis meses y cinco años en el Perú, durante el periodo de 2014 a 2019.

La muestra estuvo conformada por los niños de seis meses a cinco años que hayan sido encuestados en ENDES INEI 2014-2019 y que cumplan con los criterios de inclusión y exclusión establecidos para la encuesta.

IV. RESULTADOS

4.1. Análisis descriptivo

Un total de 81 119 observaciones fueron analizadas entre los años 2014 al 2019, para los niños de 0 a 59 meses. La Figura 1 muestra la correlación entre variables, donde destacan aquellas con un valor superior a 0.6 ($p < 0.01$), como fue en el caso de la variable **diferencia estacional de la precipitación y precipitación promedio**, que describen la diferencia estacional y el promedio de precipitación mensual por cada conglomerado. Otras variables como **temperatura máxima y temperatura mínima** también tuvieron una correlación mayor a 0.8 ($p < 0.01$).

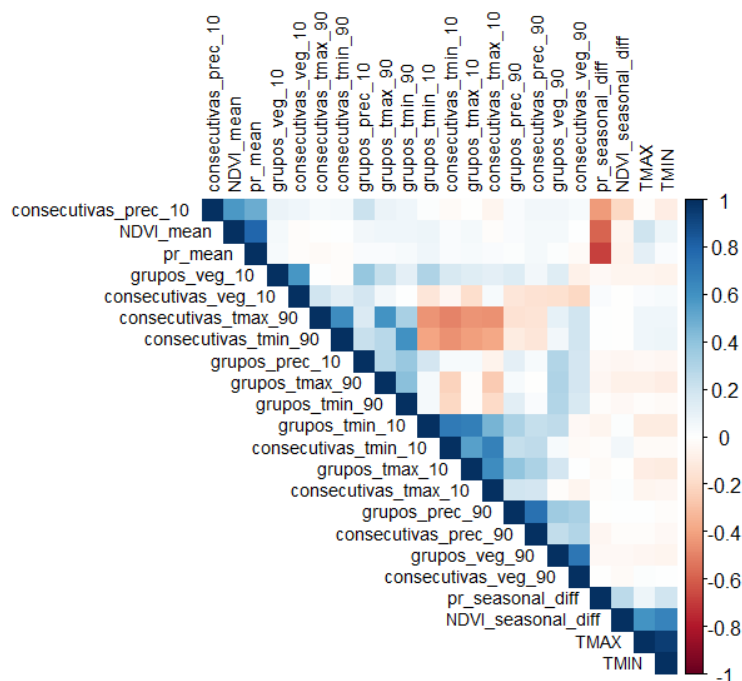


Figura 1. Correlación entre variables de estadística descriptiva. Fuente:

Elaboración propia.

Al analizar la distribución con el muestreo proporcional al 5% de los registros (Figura 2), se observó un sesgo negativo en los datos de **precipitación promedio** y el **NDVI**, y se puede apreciar que en todos los años a excepción de 2017, hay más concentración de valores precipitación media baja, lo que indicaría un sesgo a la derecha. En 2017, la precipitación media aumenta, lo cual podría estar relacionado con el fenómeno de El Niño.



Figura 2. Distribución de la variable media de precipitación en una muestra de 200 niños. Fuente: Elaboración propia.

Por su parte, la Figura 3 muestra el promedio mensual de NDVI, que debido a la naturaleza de su fórmula concentra la mayoría de sus valores en el centro del histograma, lo cual indica que esta variable tiene una tendencia a la distribución normal.

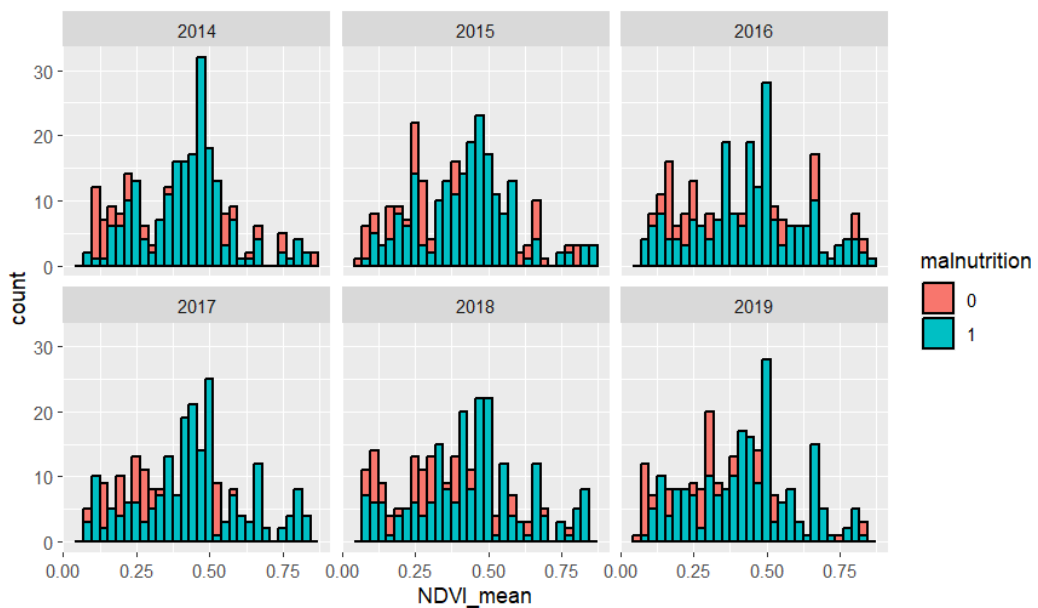


Figura 3. Distribución de la variable promedio de NDVI en una muestra de 200 niños. Fuente: Elaboración propia.

En el gráfico de cajas y bigotes de la Figura 4 se evidencia que algunos años manifiestan diferencias en las medianas, como en la variable de diferencia entre temperatura máxima y mínima. Se puede apreciar una divergencia más pronunciada en el año 2014 y 2019 en las medianas para la variable **TGAP**, dependiendo del estado de desnutrición. Sin embargo, el rango intercuartílico muestra diferencias menos marcadas en años como 2015, 2016 y 2018.

Las variables climáticas de diferencia estacional (verano e invierno) de la precipitación y NDVI no siguen una distribución normal de los datos (**Anexos 1**). Para cada variable climática se encontró diferencia estadísticamente significativas en las medianas de los grupos con y sin desnutrición, para un nivel de significancia del 5% según la prueba de U de Mann-Whitney.

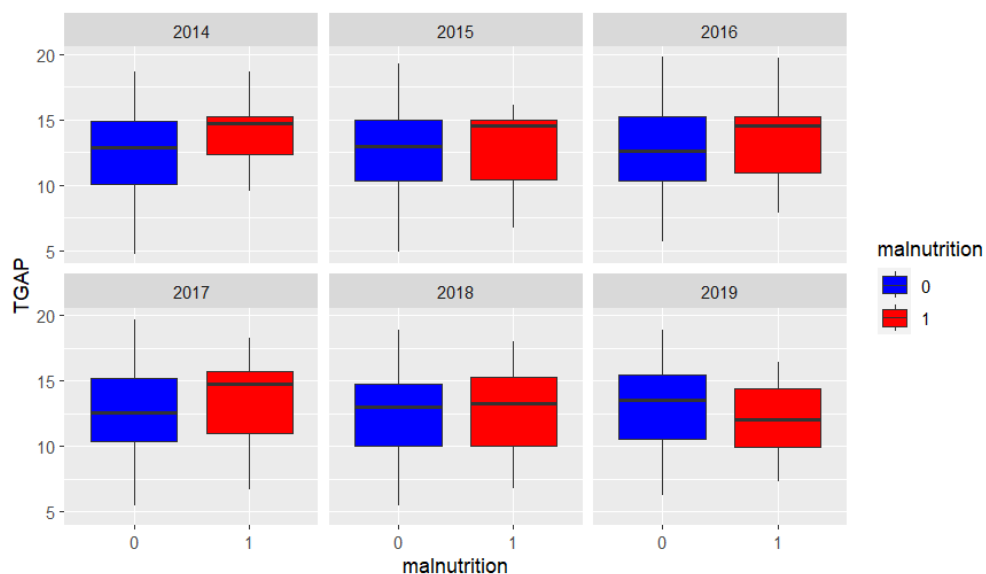


Figura 4. Gráfico de cajas y bigotes de la variable de diferencia de temperatura máxima y mínima. Fuente: Elaboración propia.

Respecto a las variables sociodemográficas, se encontró que habían datos faltantes en las variables de educación de la madre, intervalo de nacimiento y disponibilidad del agua. Se identificó que no habían casos “marcados” pero sí que los datos faltantes respondían al tipo Missing Not At Random. Para realizar la imputación de datos, se buscaron variables asociadas en las encuestas del hogar y de la madre del ENDES INEI. La Figura 8 muestra la distribución categórica de 5 de las 6 variables sociodemográficas consideradas en el estudio. No se presentó un desbalance

significativo entre las clases a nivel general, por lo que se continuó con la categorización establecida previamente.

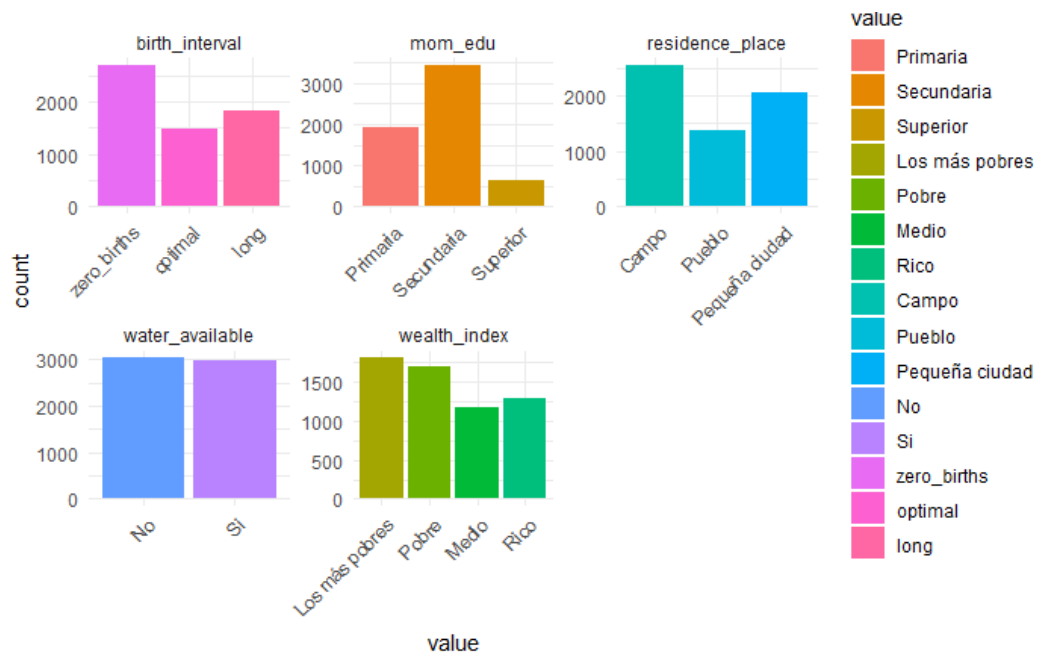


Figura 5. Gráfico de barras para las variables sociodemográficas de intervalo de nacimiento (birth_interval), educación de la madre (mom_edu), lugar de residencia (residence_place), disponibilidad de agua (water_available) e índice de riqueza del hogar (wealth_index). Fuente: Elaboración propia.

4.2. Validación cruzada

4.2.1. Random Forest

La evaluación cuantitativa de hiperparámetros mediante validación cruzada para los modelos de Random Forest demostró métricas interesantes para el grupo de variables climáticas descriptivas y de conteos de meses en condiciones climáticas críticas. La Tabla 1 es una muestra de los hiperparámetros encontrados en la validación cruzada para los años de entrenamiento 2014 y 2015 para evaluar en la ventana de tiempo del año 2016.

Tabla 1. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas y conteos de meses en condiciones climáticas críticas. 2014 a 2016. Fdescriptivas uente: Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.92</i>
<i>Model 01</i>	<i>3</i>	<i>1000</i>	<i>5</i>	<i>Kappa</i>	<i>0.39</i>
				<i>ROC AUC</i>	<i>0.71</i>
				<i>Sensitivity</i>	<i>0.75</i>

La precisión y la sensibilidad a nivel nacional fue de 92% y 87% respectivamente, indican que la capacidad de predecir entre clases es alta. Asimismo, se encontró un moderado coeficiente Kappa de 0.39 y un AUC de 0.72. Mientras tanto, la Tabla 4 muestra los hiperparámetros que mejores resultados consiguieron en la validación cruzada para la ventana de tiempo 2019, entrenando con todos los años anteriores. En ella, pese a que aumenta el AUC de 0.73, disminuyó ligeramente la precisión, y el coeficiente Kappa (94% y 38% respectivamente). En general, no hubo una variación notoria entre los años de estudio para las ventanas de tiempo de 2017 y 2018 (**Anexos 2**). En contraste, la validación cruzada para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas demostró un menor coeficiente Kappa, AUC y sensibilidad para cada ventana de tiempo, siendo el máximo 15%, 60% y 73% respectivamente (**Anexos 2**). A partir de estos resultados, se decidió optar por el set de variables climáticas descriptivas y de conteo de meses en condiciones climáticas críticas para el entrenamiento en el conjunto de datos completo y la posterior evaluación por ventanas de tiempo.

Tabla 2. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2019. Fuente: Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.94</i>
<i>Model 03</i>	<i>8</i>	<i>344</i>	<i>5</i>	<i>Kappa</i>	<i>0.38</i>
				<i>ROC AUC</i>	<i>0.73</i>
				<i>Sensitivity</i>	<i>0.78</i>

4.2.2. XGBoost

En XGBoost, la validación cruzada denota una baja en el AUC en comparación al Random Forest, como se muestra en la Tabla 3 y Tabla 4. A pesar de mantener un valor alto de precisión (95% para ambas ventanas de tiempo) y de sensibilidad (mayor o igual a 95%), el coeficiente Kappa fue más bajo para ambas configuraciones, indicando un ligero acuerdo: la de variables climáticas descriptivas y conteo de meses en condiciones climáticas críticas, así como variables climáticas, de meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas.

Tabla 3. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2016. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 18</i>	<i>10</i>	<i>6</i>	<i>0.1</i>	<i>Kappa</i>	<i>0.01</i>
				<i>ROC AUC</i>	<i>0.51</i>
				<i>Sensitivity</i>	<i>0.95</i>

Tabla 4. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2018. Fuente:

Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 20</i>	<i>27</i>	<i>8</i>	<i>0.2</i>	<i>Kappa</i>	<i>0.02</i>
				<i>ROC AUC</i>	<i>0.62</i>
				<i>Sensitivity</i>	<i>0.97</i>

Esta línea se mantuvo para el resto de las configuraciones encontradas a partir de la validación cruzada, lo cual limitó la capacidad posterior de XGBoost para el entrenamiento en el conjunto de datos completo y evaluar en las ventanas de tiempo (**Anexos 2**). Pese a variar la tasa de aprendizaje en diversas configuraciones en cuanto a la cantidad de decimales el resultado no necesariamente condujo a mejoras en las métricas, puesto que no se superó el 62% del AUC o el 3% de coeficiente Kappa.

4.3. Validación en ventanas de tiempo

Una vez obtenidos los mejores hiperparámetros para cada grupo de variables en los distintos años de validación, se verificó que en la importancia de variables del algoritmo de Random Forest, en algunas variables (grupos de meses consecutivos en condiciones climáticas críticas, meses consecutivos en condiciones climáticas críticas) para los umbrales del 90% y 10% esta importancia tomó valores negativos, mientras que las variables promedio de NDVI, diferencia estacional de NDVI,

precipitación promedio, diferencia estacional de precipitación y TGAP muestran consistentemente una importancia positiva y alta en todos los años (Figura 9 y 10). Tanto para el algoritmo de Random Forest como para XGBoost, la evolución de las métricas de precisión y sensibilidad no fue homogénea, en el sentido que los valores variaron entre los conjuntos de datos entrenados y evaluados para el periodo 2014 a 2019.

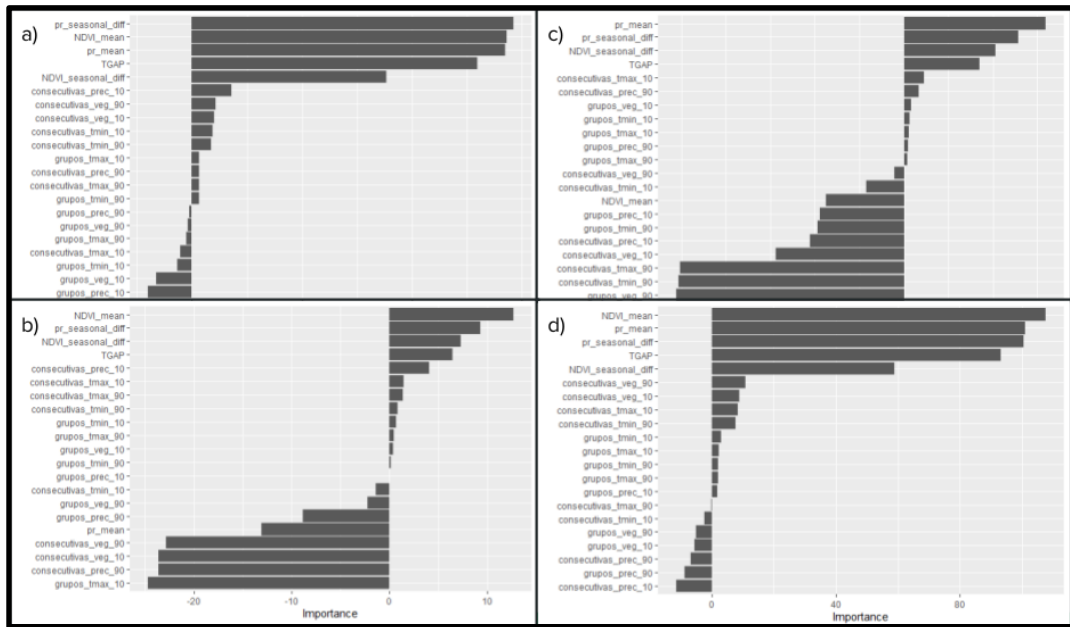


Figura 6. Gráfico de importancia de variables (estadística descriptiva, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas) por grupo de años entrenados. a) 2014 a 2016. b) 2014 a 2017. c) 2014 a 2018. d)

2014 a 2019. Fuente: Elaboración propia.

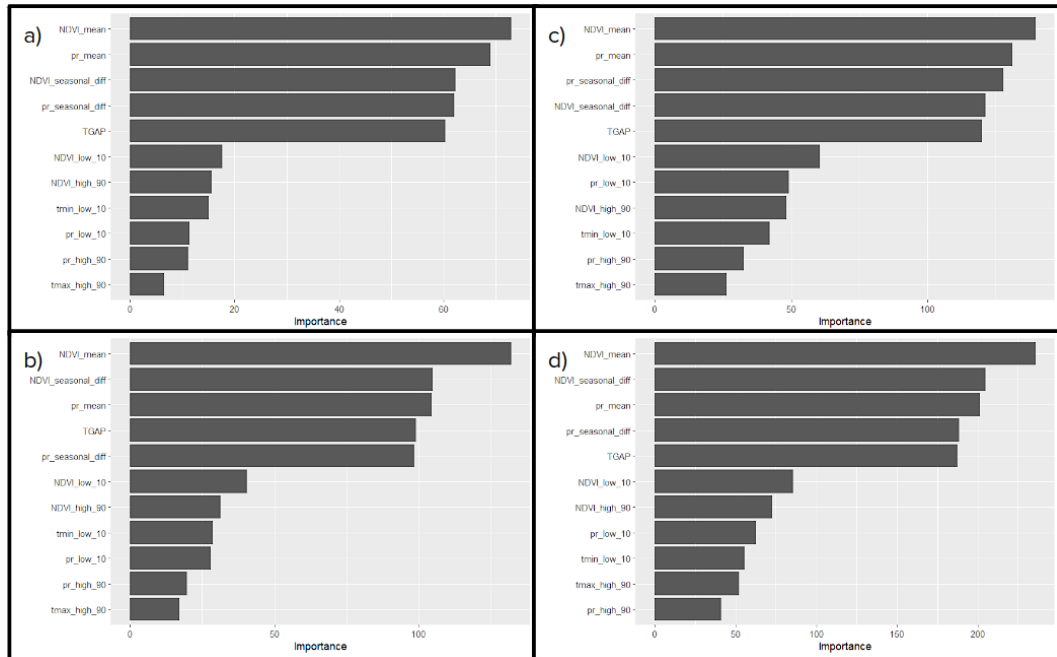


Figura 7. Gráfico de importancia de variables (estadística descriptiva y los umbrales directos de condiciones climáticas críticas) en evaluación de ventanas de tiempo por grupo de años entrenados para Random Forest. a) 2014 a 2016. b) 2014 a 2017. c) 2014 a 2018. d) 2014 a 2019. Fuente: Elaboración propia.

En el caso de Random Forest, la métrica de precisión, que cuantifica la proporción total de predicciones correctas, experimentó una disminución del promedio de 92% (Figura 11) a 90% (Figura 12) al aplicar esta técnica de validación. Aún más relevante desde una perspectiva clínica, la sensibilidad, o recall, sufrió una disminución más marcada, pasando de 82% a un máximo de 65%.

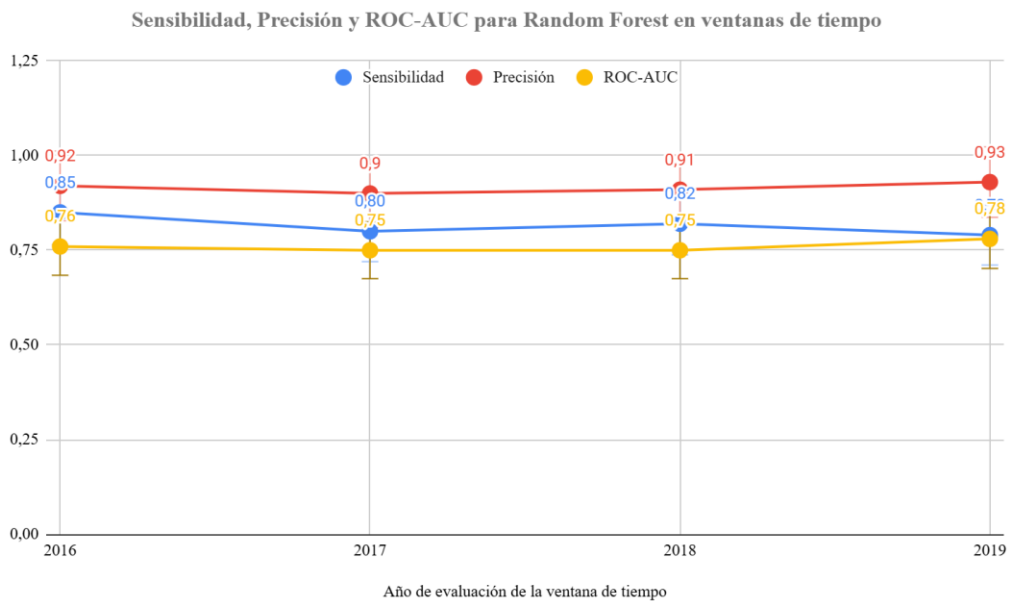


Figura 8. Evolución de las métricas de evaluación del modelo a través de las ventanas de tiempo para los datos de validación en Random Forest. Fuente: Elaboración propia.

En la validación inicial, Random Forest alcanzó un AUC del 76% para los datos sólo del 2014-2015, aumentando ligeramente al 78% para los datos del 2014-2018 (Figura 11). Estos valores, aunque incrementales, subrayan la necesidad de mejora en la discriminación del modelo entre clases. Sin embargo, para la evaluación el Random Forest tuvo un AUC del 69% como máximo en los datos del 2014-2019 (Figura 12).

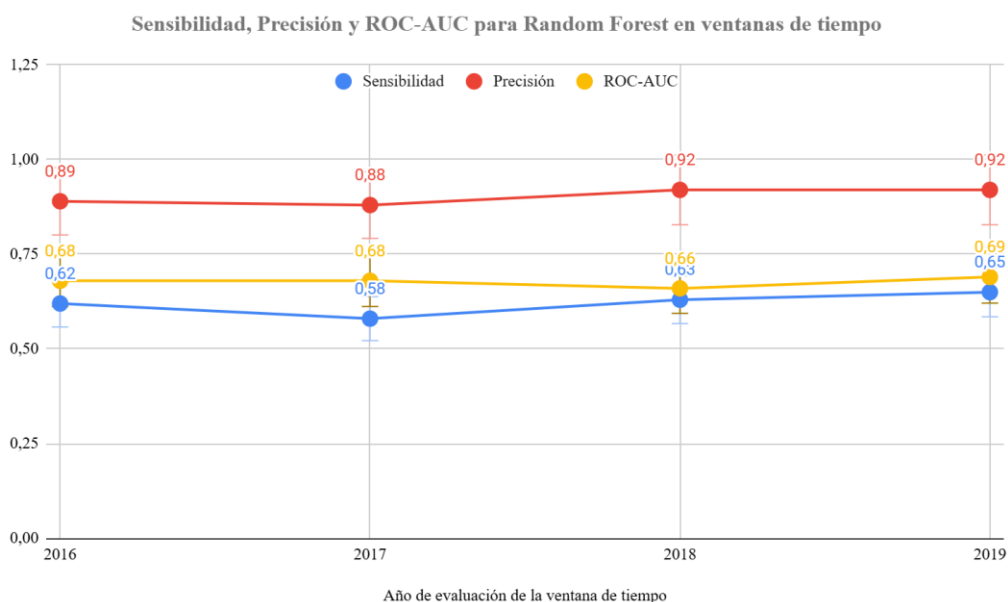


Figura 9. Evolución de las métricas de evaluación del modelo a través de las ventanas de tiempo para los datos de evaluación en Random Forest. Fuente: Elaboración propia.

4.4. Análisis por región

Los resultados de sensibilidad por departamento a través de los años se pueden apreciar en la Tabla 5 para el modelo de Random Forest, el cual tuvo los mejores resultados por departamento. A nivel de región la sensibilidad más baja se encontró en departamentos de la costa del Perú (Ica, Lambayeque, Tacna, etc.). En cuanto a la sierra, los valores de sensibilidad variaron con un mínimo de 51% en 2017 hasta 72% en 2019. Los departamentos de la selva del Perú mostraron la mayor sensibilidad a la hora de evaluar las ventanas de tiempo con los modelos entrenados, obteniendo hasta un 90% en esta métrica (Figura 13).

Los valores detallados de sensibilidad para cada grupo etario y departamento (Tabla 20, **Anexos 2**) muestran que, en términos generales, las métricas se mantienen estables a lo largo de los cinco bloques de edad considerados (6-11, 12-23, 24-35, 36-47 y 48-59 meses). Aunque las variaciones absolutas son pequeñas, se observa una tendencia consistente a la reducción de la sensibilidad conforme aumenta la edad: los intervalos de 6-23 meses suelen presentar los valores más altos, mientras

que los niños de 48-59 meses exhiben una ligera disminución en la mayoría de los departamentos y años.

Tabla 5. Sensibilidad por departamento a través de los años 2016 a 2019 usando las variables climáticas de estadística descriptiva y conteo de meses en condiciones climáticas críticas. Fuente: Elaboración propia

Departamento	2016	2017	2018	2019
<i>Amazonas</i>	0,88	0,85	0,75	0,78
<i>Ancash</i>	0,29	0,50	0,29	0,20
<i>Apurimac</i>	0,94	0,70	0,60	1,00
<i>Arequipa</i>	0,23	0,30	0,35	0,33
<i>Ayacucho</i>	0,79	0,36	0,81	1,00
<i>Cajamarca</i>	0,46	0,46	0,64	0,40
<i>Cusco</i>	0,84	0,41	0,23	1,00
<i>Huancavelica</i>	0,40	0,50	0,88	0,67
<i>Huánuco</i>	0,71	0,70	0,89	1,00
<i>Ica</i>	0,20	0,25	0,25	0,20
<i>Junín</i>	0,24	0,38	0,86	0,38
<i>Libertad</i>	0,63	0,40	0,53	0,55
<i>Lambayeque</i>	0,45	0,35	0,29	0,36
<i>Lima</i>	0,44	0,20	0,25	0,40
<i>Loreto</i>	1,00	0,95	0,87	0,90
<i>Madre de Dios</i>	0,75	0,80	0,83	0,70
<i>Moquegua</i>	0,50	0,60	0,47	0,50
<i>Pasco</i>	0,25	0,40	0,90	0,60
<i>Piura</i>	0,23	0,29	0,23	0,38
<i>Puno</i>	0,47	0,70	0,50	0,67
<i>San Martín</i>	0,90	0,88	0,87	0,87
<i>Tacna</i>	0,20	0,20	0,20	0,20
<i>Tumbes</i>	0,67	0,42	0,43	0,56
<i>Ucayali</i>	0,90	0,95	1,00	1,00

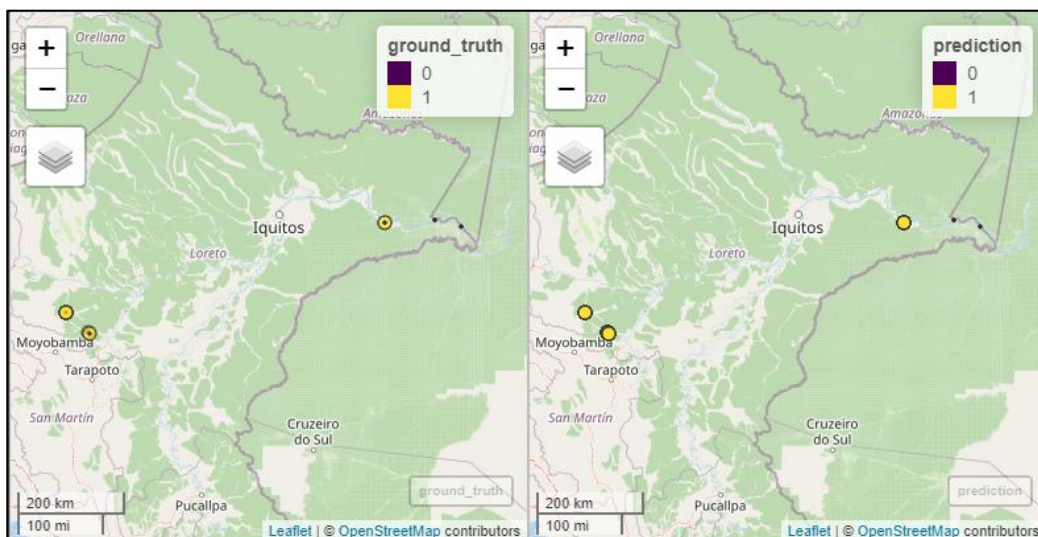


Figura 10. Visualización de casos de desnutrición aguda en el departamento de Loreto en 2016. A la izquierda se observa el ground truth, mientras que a la derecha están las predicciones del modelo de Random Forest. Fuente: Elaboración propia.

4.5. Comparación con modelo sociodemográfico

Se generó un modelo con predictores sociodemográficos para cada una de las ventanas de tiempo de evaluación, utilizando los mismos hiperparámetros que los que fueron ajustados por la validación cruzada para el modelo Random Forest. Los resultados de sensibilidad por departamento a lo largo de los años se presentan en la Tabla 6. A nivel regional, los valores más bajos de sensibilidad se registraron en los departamentos de la costa del Perú, con valores que oscilaron entre 36% y 68%, reflejando una capacidad variada de detección en esta zona. En la sierra, la sensibilidad mostró variaciones significativas, con un mínimo de 52% en 2018 y un máximo de 68% en 2017, lo que indica una mejora progresiva en el desempeño del modelo para esta región. Por su parte, la selva presentó los valores de sensibilidad más altos a lo largo del tiempo, alcanzando hasta un 73% en 2017 y manteniéndose por encima del 50% en todos los años evaluados.

Para evaluar si el desempeño del modelo sociodemográfico varía dentro del propio rango etario analizado (6-59 meses), se construyó la Tabla 21 (**Anexos 2**). El patrón resultante describe sensibilidades comparativamente más bajas en los extremos (6-11 m y 48-59 m), alcanzan sus máximos entre los 12-23 m y 24-35 m, y descienden

nuevamente en los grupos mayores. Por ejemplo, en Huánuco en 2017 la métrica pasa de 0,70 (6-11 m) a 0,95 (24-35 m) y baja a 0,71 (48-59 m); un comportamiento semejante se observa en Cajamarca y Ayacucho. Las diferencias absolutas rara vez superan los 0,10-0,15 puntos, lo que indica que, a pesar de esta variación con la edad, el modelo mantiene un rendimiento estable.

Tabla 6. Sensibilidad por departamento a través de los años 2016 a 2019 usando las variables sociodemográficas. Fuente: Elaboración propia

Departamento	2016	2017	2018	2019
<i>Amazonas</i>	0.33	0.73	0.46	0.42
<i>Ancash</i>	0.60	0.63	0.52	0.60
<i>Apurimac</i>	0.72	0.75	0.50	1.00
<i>Arequipa</i>	0.30	0.13	0.22	0.18
<i>Ayacucho</i>	0.68	0.72	0.43	0.67
<i>Cajamarca</i>	0.78	0.85	0.71	0.70
<i>Cusco</i>	0.43	0.70	0.74	0.50
<i>Huancavelica</i>	0.49	0.68	0.67	1.00
<i>Huanuco</i>	0.73	0.92	0.58	0.58
<i>Ica</i>	0.30	0.17	0.32	0.00
<i>Junin</i>	0.44	0.67	0.37	0.88
<i>La Libertad</i>	0.29	0.77	0.60	1.00
<i>Lambayeque</i>	0.30	0.41	0.46	0.50
<i>Lima</i>	0.19	0.56	0.11	0.30
<i>Loreto</i>	0.73	1.00	0.81	1.00
<i>Madre de Dios</i>	0.35	0.25	0.31	0.67
<i>Moquegua</i>	0.20	0.38	0.22	0.46
<i>Pasco</i>	0.56	0.54	0.57	0.00
<i>Piura</i>	0.50	0.66	0.50	0.23
<i>Puno</i>	0.64	0.65	0.38	0.25
<i>San Martin</i>	0.50	0.70	0.69	1.00
<i>Tacna</i>	0.17	0.53	0.25	0.39

<i>Tumbes</i>	<i>0.32</i>	<i>0.55</i>	<i>0.32</i>	<i>0.65</i>
<i>Ucayali</i>	<i>1.00</i>	<i>1.00</i>	<i>0.52</i>	<i>0.00</i>

V. DISCUSIÓN

La desnutrición aguda infantil continúa siendo un reto crítico en salud pública, particularmente en contextos como el Perú, donde las condiciones climáticas pueden agravar esta problemática. En este estudio, se desarrollaron y evaluaron modelos predictivos basados en algoritmos de aprendizaje automático, como Random Forest y XGBoost, para predecir la desnutrición aguda en niños de 0 a 59 meses entre 2014 y 2019, utilizando exclusivamente variables climáticas. Los resultados mostraron un rendimiento consistente en la validación, con métricas de precisión superiores al 90% y sensibilidades de hasta 85% en años de entrenamiento. Sin embargo, se observó una disminución en la generalización a conjuntos nuevos de datos, reflejando limitaciones inherentes a los modelos basados únicamente en factores climáticos. Además, el análisis por región evidenció que las zonas de la selva y la sierra presentaron una sensibilidad significativamente mayor en comparación con la costa, destacando la influencia de factores climáticos diferenciados entre regiones, especialmente en áreas con alta vulnerabilidad climática como la selva peruana.

5.1. Análisis descriptivo

La identificación de correlaciones altas entre variables climáticas específicas, como la diferencia estacional y el promedio de precipitación mensual, junto con las temperaturas máximas y mínimas, enfatiza la estrecha interacción entre el clima y el bienestar nutricional de los niños.

Los casos de desnutrición tienden a aumentar durante períodos de precipitación media baja y reducción en la vegetación resalta la vulnerabilidad de las comunidades a los cambios medioambientales. Esto es particularmente notable en años específicos, como 2015, 2016 y 2018, donde se acentúan estas condiciones. Sin embargo, la ausencia de patrones consistentes en las variaciones anuales y entre

los grupos de niños sugiere la influencia de otros factores no climáticos o la necesidad de un enfoque más granular para capturar cómo las condiciones ambientales específicas de cada región del Perú afectan la nutrición infantil. Esto resalta la complejidad de la desnutrición aguda como un fenómeno multifactorial, donde los determinantes climáticos son significativos, pero no exclusivos.

5.2. Análisis comparativo de resultados

Este estudio comprobó que el diseño de un modelo predictivo puede identificar la presencia de desnutrición aguda en niños de seis meses a cinco años en el Perú, a partir de variables climáticas como predictores principales, alineándose con lo propuesto por Berrang-Ford *et al.* a la hora de analizar el impacto climático y la estacionalidad en la salud para países de ingresos medios y bajos (13). De esta manera, los valores de precisión (92%) y sensibilidad (85%) en validación, es decir en el mismo año de entrenamiento, son comparables a lo propuesto por Siy *et al.*(20), quienes usaron Naive Bayes, K-Nearest Neighbors, Support Vector Machine y Random Forest (precisión de 79% y la sensibilidad de 90% como máximo) para predecir la inseguridad alimentaria en una ciudad y provincia rural de Filipinas. La utilización de la precipitación y la temperatura se había demostrado como un relevante set de predictores para enfermedades infecciosas como la malaria, sin incluir el factor estacional, tal como lo demostró Nikiruka *et al.* (14), con resultados comparables a los de este estudio en precisión (98% como máximo en países africanos como Mali y 95% como máximo en el Perú) respecto al uso de XGBoost para modelar una enfermedad no transmisible como la desnutrición aguda. Sin embargo, aplicando la estacionalidad, los resultados del XGBoost empeoraron, tal y como lo demostró el estudio de Kamana *et al.*, quienes identificaron que el factor de la temporalidad en los datos climáticos impacta en el rendimiento del modelo, puesto que el RMSE era mayor que otros modelos más avanzados como el LSTMSeq2Seq, empleados principalmente para tareas de series temporales (16).

Por su parte, en comparación a otros estudios que buscaron predecir la desnutrición en niños menores de 5 años (10), se tuvieron resultados similares en cuanto a la sensibilidad de sus modelos y el presentado en esta investigación (94,66% y 85%

respectivamente) para Random Forest, pero en términos de precisión bajaron considerablemente su rendimiento (68% y 92%). Asimismo, este estudio, identificó como predictores principales a las variables climáticas, variables importantes tanto en estadística descriptiva, condiciones climáticas críticas y conteos de meses bajo estas condiciones.

En línea con lo anterior, Nchimunya agregó en su estudio comparativo de modelos de Machine Learning al XGBoost, y entre sus métricas de evaluación aplicó el área bajo la curva, obteniendo un valor de 0,59 para su mejor modelo, el Random Forest (23). Es notorio entonces que en esta investigación logró igualar y superar en Random Forest el valor de ROC-AUC, con un 76% para la ventana de tiempo de 2019 en los datos de validación, además que usando el XGBoost, se llegó a valores de hasta 61% para esta misma métrica. En el ámbito de la seguridad alimentaria, Deléglise *et al.* usó variables como la precipitación, las altas temperaturas y otros datos meteorológicos con el fin de predecir indicadores como el Food Consumption Score (FCS) (19), demostrando que, así como en este estudio, las variables climáticas son estadísticamente significativas a la hora de componer modelos de inteligencia artificial para solucionar problemas de salud pública. También Usman *et al.* corresponde a esta idea de la vulnerabilidad climática como un factor significativo (21), siendo que, sin utilizar el cálculo de importancia de variables de Random Forest, comprobaron que este factor climático exacerba el desarrollo infantil por hasta un punto porcentual.

5.3. Generalización de modelos en ventanas de tiempo

Para la construcción del modelo predictivo, un total de 50 modelos entrenados y evaluados para cada ventana de tiempo, favoreciendo la generalización de los modelos y el uso de los mejores hiperparámetros, que se ajustaron para maximizar métricas de interés a la sensibilidad, precisión, AUC y coeficiente kappa. Este último representó un problema que es importante destacar, ya que si bien no es la única medida estadística que va a determinar qué tan bien el modelo encaja con los datos, sí da un acercamiento más robusto a la concordancia y el efecto del azar en los modelos como XGBoost, donde es más probable que exista un sobreajuste, ya sea debido a la cantidad de variables como a la representatividad indirecta de las

variables predictoras. El hecho de que diferentes configuraciones de modelos muestran variaciones en el AUC y la métrica kappa refleja la complejidad inherente al modelado de fenómenos de salud pública como la desnutrición. La mejora en el AUC a lo largo de las configuraciones de modelos sugiere una discriminación más efectiva entre clases (niños desnutridos y no desnutridos), lo cual es esencial para la implementación de programas de intervención temprana y precisos.

Consecuentemente a la utilización de ventanas de tiempo, la importancia de las variables de grupos de meses y meses en condiciones climáticas críticas en los modelos varió, teniendo desde valores positivos en algunos años, hasta valores muy negativos. En los modelos donde la importancia de las variables tomó valores negativos, la explicación se basa en que aquellas variables carecen de relevancia, y es más el azar lo que provoca que puedan aportar en la predicción e influir positivamente en el modelo. En términos simples, esas variables actúan mejor cuando las observaciones se permutan aleatoriamente, cuando debería ser que empeoran el rendimiento del modelo basado en el error. Por su otro lado, las variables de estadística descriptiva se mantuvieron consistentemente como las más relevantes, y por ende con valores positivos durante los cuatro años de evaluación en ventanas de tiempo.

Otra observación pertinente en relación a la generalización de los modelos es la agregación de los datos para los modelos de Machine Learning. Esto puede ser un problema debido a la pérdida de información que podría significar sumarizar, en este caso específico 12 meses, una cantidad de variables a solo una. Para ello, los conteos de meses consecutivos y grupos de meses consecutivos se aplicaron; sin embargo, en los gráficos de importancia de variables, presentaron un menor aporte a la clasificación de desnutrición aguda respecto a las variables de estadística descriptiva. Estos resultados sugieren que la complejidad de los datos meteorológicos responde mejor a una estacionalidad con resolución temporal más pequeña, es decir, analizando el fenómeno en unidades de tiempo menores que el año calendario.

La pregunta de investigación planteada inicialmente estuvo alineada con la sensibilidad a la hora de predecir los verdaderos positivos en desnutrición aguda, debido a que es una afección que representa una amenaza a la salud pública todos

los años, especialmente para los años de vida saludables potenciales perdidos. Con esto en consideración, y debido al desbalance en la proporción en la población general y la muestra tomada del ENDES INEI, la especificidad resulta ser una métrica interesante para una herramienta de descarte.

5.4. Interpretabilidad de resultados

La capacidad predictiva observada mediante variables exclusivamente climáticas puede explicarse desde varias perspectivas epidemiológicas y ambientales. En primer lugar, las variables climáticas seleccionadas, tales como la precipitación y las temperaturas extremas, poseen una relación directa y significativa con la producción agrícola y, por ende, con la disponibilidad alimentaria, especialmente en zonas rurales del Perú que dependen fuertemente de la agricultura de subsistencia. De esta manera, estas variables ambientales pueden actuar como determinantes inmediatos en el acceso a alimentos y la seguridad alimentaria familiar.

Adicionalmente, cambios en los patrones climáticos están estrechamente relacionados con la incidencia de enfermedades infecciosas y parasitarias, consideradas causas indirectas relevantes de desnutrición infantil aguda. Estos eventos infecciosos aumentan los requerimientos energéticos y reducen la capacidad de absorción nutricional en niños afectados. Por lo tanto, las variables climáticas también funcionan como indicadores indirectos del riesgo epidemiológico que influye significativamente en el estado nutricional infantil.

En este contexto, la elección específica de algoritmos como XGBoost y Random Forest resulta especialmente adecuada debido a su inherente capacidad interpretativa. Estos modelos permiten medir claramente la importancia relativa de cada variable predictora, facilitando la interpretación directa y aportando valor adicional en estudios epidemiológicos donde se maneja un elevado volumen de datos y clases desbalanceadas. Además, el uso futuro de técnicas avanzadas de interpretación local, tales como SHapley Additive exPlanations (SHAP) o Local Interpretable Model-Agnostic Explanations (LIME), podría fortalecer aún más esta

capacidad interpretativa, apoyando así decisiones clínicas y políticas públicas más informadas.

Finalmente, para evaluar la robustez y estabilidad predictiva, el índice Kappa desempeñó un papel clave al considerar el efecto del azar en las clasificaciones del modelo. La variabilidad observada en Kappa a través de distintas configuraciones y períodos temporales refleja tanto la complejidad inherente del fenómeno estudiado como la necesidad de interpretar cuidadosamente los resultados obtenidos. Los valores moderados de Kappa (0.21-0.60) encontrados sugieren que, aunque el modelo demuestra una capacidad predictiva adecuada, existe espacio considerable para mejorar la estabilidad y la consistencia de sus predicciones.

5.5. Análisis por región

En cuanto al análisis de sensibilidad por región, se encontró que para los departamentos de la selva la sensibilidad era más alta, para la sierra variaron considerablemente y para la costa se obtuvieron métricas de sensibilidad baja. Respecto a la región selva, en la Tabla 5 se aprecia que Madre de Dios, Ucayali y Loreto presentaron una tendencia de alta sensibilidad a lo largo de los años. En este sentido, el estudio de Sobrino et al. indicó que generalmente la desnutrición aguda afecta en mayor medida a la sierra y selva peruana (1), que puede ser explicado debido a la variabilidad climatológica de estas regiones en cuanto a la precipitación, las temperaturas extremas e incluso las actividades económicas principales de los centros poblados o comunidades donde se extrajeron los datos para el ENDES. Al haber muchas zonas cuya actividad económica es la ganadería, agricultura, acuicultura, entre otros, dependen en gran medida de los factores climáticos. Mientras tanto, en la costa del Perú y en las ciudades más grandes de esta región usualmente hay una menor cantidad de casos que puedan deberse a los factores como las variables independientes utilizadas para esta investigación. En general, el Perú contiene un gran porcentaje de población rural, y esta es la más afectada con los cambios en las variables climáticas que juegan un papel importante, aunque indirecto, en la alimentación y la salud de los niños.

Ante esta nueva perspectiva sobre la sensibilidad de los modelos con enfoque epidemiológico y ambiental para la prevención de la desnutrición aguda en niños,

el tener regiones donde las métricas sean mejores o peores, lleva a la disyuntiva sobre si utilizarlo a nivel nacional sería una posibilidad, o si por el contrario una herramienta como esta sirve para tomar decisiones en un área más focalizada, como puede ser el análisis por departamento, en el cual cada gestor de salud regional tendría la potestad de emplear métodos como este para dirigir de una manera más efectiva los recursos destinados a combatir la desnutrición aguda en niños.

Asimismo, al desagregar la sensibilidad regional por grupos etarios (Tablas 20 y 21, **Anexos 2**) se observa que el efecto de la edad es discreto pero diferenciado según el tipo de modelo. Con variables climáticas (Tabla 20) la sensibilidad permanece prácticamente estable en todos los departamentos, especialmente en la selva, y solo muestra una ligera caída de 3 a 5 puntos porcentuales en los niños de 48-59 meses, lo que confirma que el componente ambiental predice con similar eficacia a lo largo del primer quinquenio de vida. Esto podría explicarse debido a que los primeros meses de vida constituyen la etapa de mayor vulnerabilidad nutricional: el rápido crecimiento, las reservas corporales limitadas y la dependencia casi total de los cuidados incrementan la sensibilidad a variaciones ambientales. Por ello el modelo climático mantiene valores de sensibilidad elevados y bastante estables en los bloques etarios iniciales, reflejando un riesgo continuo asociado a las condiciones externas.

5.6. Comparación con modelo sociodemográfico

Los resultados obtenidos por el modelo sociodemográfico demuestran una tendencia similar en la diferencia de la sensibilidad para las regiones de la selva (60% y 87%), sierra (61% y 61%) y costa (51% y 39%) para el set sociodemográfico y el set climático respectivamente, habiendo utilizado como predictores a las variables descritas en el trabajo de Talukder para la predicción de la desnutrición en Bangladesh (10). El modelo basado en variables climáticas muestra una mayor contribución a la sensibilidad en la selva, lo que sugiere que factores como la precipitación, la vegetación y la temperatura podrían estar desempeñando un papel clave en la detección de la desnutrición en esta región. En cuanto al análisis segmentado por edad (Tabla 21), los valores son algo menores en los extremos (lactantes 6-11 m y pre-escolares 48-59 m), alcanzan su punto más alto entre 12-35

meses y descienden después. Pese a estas diferencias, las variaciones absolutas de sensibilidad entre grupos de edad son moderadas; en consecuencia, las conclusiones regionales previas podrían mantenerse independientemente del estrato etario analizado. Dado que las mediciones sociodemográficas solo se realizan una vez al año en ENDES INEI y las variables climáticas están disponibles con mayor frecuencia a través de bases de datos satelitales, estos resultados resaltan el potencial de un modelo basado exclusivamente en variables climáticas, al menos para regiones donde las variaciones climáticas pueden impactar en mayor medida.

5.7.Limitaciones

Las limitaciones del modelo deben tenerse en cuenta al interpretar los resultados. Se observó una reducción en la sensibilidad al aplicar el modelo a nuevos conjuntos de datos, reflejando desafíos en la generalización y estabilidad del modelo. Este estudio buscó proponer una solución práctica y viable en contextos donde se dispone de fuentes de información satelitales que son más accesibles que las bases de datos recogidas anualmente por ENDES INEI, aunque la dependencia exclusiva de variables climáticas puede implicar la exclusión de factores adicionales relevantes para comprender integralmente la desnutrición.

La cantidad de observaciones, el desbalance de clases y la temporalidad de los datos suponen un reto que se ha abordado en la presente investigación, y que a futuro deberá continuar con el análisis de otros métodos para mejorar la precisión predictiva, como las series temporales, siempre y cuando se pueda obtener una medición de la desnutrición aguda en una resolución temporal comparable a la que permiten las variables climáticas.

VI. CONCLUSIÓN

En este estudio se diseñó un modelo predictivo basado en variables climáticas, obtenidas de Google Earth Engine, para identificar la presencia de desnutrición aguda en niños menores de cinco años, utilizando datos del INEI ENDES entre 2014 y 2019. Las variables más relevantes fueron el índice NDVI, la precipitación media mensual y las temperaturas máximas, mínimas y su diferencia.

El análisis evidenció que las variables de estadística descriptiva, particularmente el promedio de NDVI y la precipitación media mensual, poseen mayor importancia predictiva que aquellas basadas en conteos de meses consecutivos bajo condiciones climáticas extremas. Al evaluar diferentes configuraciones temporales, se observó una mejora progresiva en las métricas de precisión, sensibilidad y AUC, alcanzando los mejores resultados al utilizar los años 2014 a 2018 para entrenamiento y evaluar en el año 2019.

Los resultados indican similitudes en los patrones regionales de sensibilidad al comparar el modelo climático con modelos sociodemográficos previos, aunque se destacan diferencias claras en la capacidad predictiva por región. Específicamente, el modelo basado en variables climáticas demostró una mejor capacidad predictiva en la región de la selva, sugiriendo que factores ambientales como la precipitación, la vegetación y las temperaturas extremas desempeñan un papel significativo en la predicción de la desnutrición infantil en estas áreas.

No obstante, es importante reconocer que el modelo presenta limitaciones significativas. Se observó una disminución considerable de la sensibilidad al extrapolar resultados a nuevos conjuntos de datos, especialmente en regiones como la costa, indicando desafíos en términos de generalización y robustez del modelo. La baja sensibilidad observada en la costa resalta la necesidad de enfoques específicos que puedan incorporar potencialmente otros factores más allá de las variables climáticas disponibles.

Además, aunque los resultados obtenidos fueron alentadores en términos de capacidad predictiva y utilidad práctica en contextos donde las fuentes tradicionales de datos son limitadas, el modelo desarrollado no es completamente robusto. Las variaciones moderadas observadas en métricas como el índice Kappa sugieren la

necesidad de interpretar los resultados con precaución y continuar trabajando en mejorar la estabilidad y consistencia predictiva.

Finalmente, este estudio refuerza la importancia de considerar explícitamente las limitaciones y capacidades del uso exclusivo de variables climáticas, especialmente en contextos vulnerables. Se subraya la necesidad de seguir explorando soluciones prácticas y efectivas en países de medianos y bajos ingresos, como el Perú, adoptando un enfoque diferenciado y adaptado a las particularidades climáticas y socioeconómicas de cada región.

VII. REFERENCIAS

1. World Health Organization, UNICEF, World Bank Group. Levels and trends in child malnutrition. [citado 16 de marzo de 2024]; Disponible en: <https://www.fao.org/3/cc2314es/online/sofi-statistics-rlc-2022/wasting-among-children.html>
2. Ministerio de Salud (MINSA), Programa Mundial de Alimentos, Naciones Unidas. El Costo de la Doble Carga de la Malnutrición: Impacto Social y Económico [Internet]. 2022 [citado 16 de marzo de 2024]. Disponible en: https://docs.wfp.org/api/documents/WFP-0000140902/download/?_ga=2.137069616.1643457884.1710643220-1640029338.1710643220
3. Alessandra Marini, Claudia Rokx, Paul Gallagher. Dando la talla: El éxito del Perú en la lucha contra la desnutrición crónica [Internet]. Grupo Banco Mundial; 2017 [citado 16 de marzo de 2024]. Disponible en: <https://documents1.worldbank.org/curated/en/891441505495680959/pdf/FINAL-Peru-Nutrition-Book-in-Spanish-Oct-11.pdf>
4. Sobrino M, Gutiérrez C, Cunha AJ, Dávila M. Desnutrición infantil en menores de cinco años en Perú: tendencias y factores determinantes. *Rev Panam Salud Publica*. 2014;
5. World Health Organization. Malnutrition [Internet]. 2024 [citado 16 de marzo de 2024]. Disponible en: <https://www.who.int/news-room/fact-sheets/detail/malnutrition>
6. Marshak A, Venkat A, Young H, Naumova EN. How Seasonality of Malnutrition Is Measured and Analyzed. *Int J Environ Res Public Health*. enero de 2021;18(4):1828.
7. Paredes Mamani RP. Efecto de factores ambientales y socioeconómicas del hogar sobre la desnutrición crónica de niños menores de 5 años en el Perú. *Rev Investig Altoandinas*. julio de 2020;22(3):226-37.
8. Roba KT, O'Connor TP, Belachew T, O'Brien NM. Seasonal variation in nutritional status and anemia among lactating mothers in two agro-ecological zones of rural Ethiopia: A longitudinal study. *Nutrition*. 1 de octubre de 2015;31(10):1213-8.
9. Zahangir MS, Hasan MM, Richardson A, Tabassum S. Malnutrition and non-

- communicable diseases among Bangladeshi women: an urban–rural comparison. *Nutr Diabetes*. marzo de 2017;7(3):e250-e250.
10. Talukder A, Ahammed B. Machine learning algorithms for predicting malnutrition among under-five children in Bangladesh. *Nutrition*. 1 de octubre de 2020;78:110861.
 11. Dlamini SN, Beloconi A, Mabaso S, Vounatsou P, Impouma B, Fall IS. Review of remotely sensed data products for disease mapping and epidemiology. *Remote Sens Appl Soc Environ*. 1 de abril de 2019;14:108-18.
 12. Yuan L, Pu R, Zhang J, Wang J, Yang H. Using high spatial resolution satellite imagery for mapping powdery mildew at a regional scale. *Precis Agric*. junio de 2016;17(3):332-48.
 13. Berrang-Ford L, Sietsma AJ, Callaghan M, Minx JC, Scheelbeek PFD, Haddaway NR, et al. Systematic mapping of global research on climate and health: a machine learning review. *Lancet Planet Health*. 1 de agosto de 2021;5(8):e514-25.
 14. Nkiruka O, Prasad R, Clement O. Prediction of malaria incidence using climate variability and machine learning. *Inform Med Unlocked*. 1 de enero de 2021;22:100508.
 15. Singh A, Mehra M, Kumar A, Niranjannaik M, Priya D, Gaurav K. Leveraging hybrid machine learning and data fusion for accurate mapping of malaria cases using meteorological variables in western India. *Intell Syst Appl*. 1 de febrero de 2023;17:200164.
 16. Kamana E, Zhao J, Bai D. Predicting the impact of climate change on the re-emergence of malaria cases in China using LSTMSeq2Seq deep learning model: a modelling and prediction analysis study. *BMJ Open*. 2022;e053922-e053922.
 17. Ochida N, Mangeas M, Dupont-Rouzeyrol M, Dutheil C, Forfait C, Peltier A, et al. Modeling present and future climate risk of dengue outbreak, a case study in New Caledonia. *Environ Health*. 20 de enero de 2022;21(1):20.
 18. Mishra VK, Tiwari N, Ajaymon SL. Dengue Disease Spread Prediction Using Twofold Linear Regression. En: 2019 IEEE 9th International Conference on Advanced Computing (IACC) [Internet]. 2019 [citado 16 de marzo de 2024]. p. 182-7. Disponible en: <https://ieeexplore.ieee.org/document/8971567>

19. Deléglise H, Interdonato R, Bégué A, Maître d'Hôtel E, Teisseire M, Roche M. Food security prediction from heterogeneous data combining machine and deep learning methods. *Expert Syst Appl.* 15 de marzo de 2022;190:116189.
20. Siy Van VT, Antonio VA, Siguin CP, Gordoncillo NP, Sescon JT, Go CC, et al. Predicting undernutrition among elementary schoolchildren in the Philippines using machine learning algorithms. *Nutr Burbank Los Angel Cty Calif.* abril de 2022;96:111571.
21. Usman M, Kopczewska K. Spatial and Machine Learning Approach to Model Childhood Stunting in Pakistan: Role of Socio-Economic and Environmental Factors. *Int J Environ Res Public Health.* enero de 2022;19(17):10967.
22. Browne C, Matteson DS, McBride L, Hu L, Liu Y, Sun Y, et al. Multivariate random forest prediction of poverty and malnutrition prevalence. *PLOS ONE.* 8 de septiembre de 2021;16(9):e0255519.
23. Chilyabanyama ON, Chilengi R, Simuyandi M, Chisenga CC, Chirwa M, Hamusonde K, et al. Performance of Machine Learning Classifiers in Classifying Stunting among Under-Five Children in Zambia. *Children.* julio de 2022;9(7):1082.
24. Fenta HM, Zewotir T, Muluneh EK. A machine learning classifier approach for identifying the determinants of under-five child undernutrition in Ethiopian administrative zones. *BMC Med Inform Decis Mak.* 24 de octubre de 2021;21(1):291.
25. Chawla NV, Bowyer KW, Hall LO, Kegelmeyer WP. SMOTE: Synthetic Minority Over-sampling Technique. *J Artif Intell Res.* 1 de junio de 2002;16:321-57.
26. Lee KJ, Carlin JB, Simpson JA, Moreno-Betancur M. Assumptions and analysis planning in studies with missing data in multiple variables: moving beyond the MCAR/MAR/MNAR classification. *Int J Epidemiol.* 1 de agosto de 2023;52(4):1268-75.
27. Breiman L. Random Forests. *Mach Learn.* 1 de octubre de 2001;45(1):5-32.
28. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. En: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [Internet].* 2016 [citado 22 de marzo de 2025]. p. 785-94. Disponible en: <http://arxiv.org/abs/1603.02754>

29. Maldonado S, López J, Iturriaga A. Out-of-time cross-validation strategies for classification in the presence of dataset shift. *Appl Intell.* 1 de marzo de 2022;52(5):5770-83.

ANEXOS

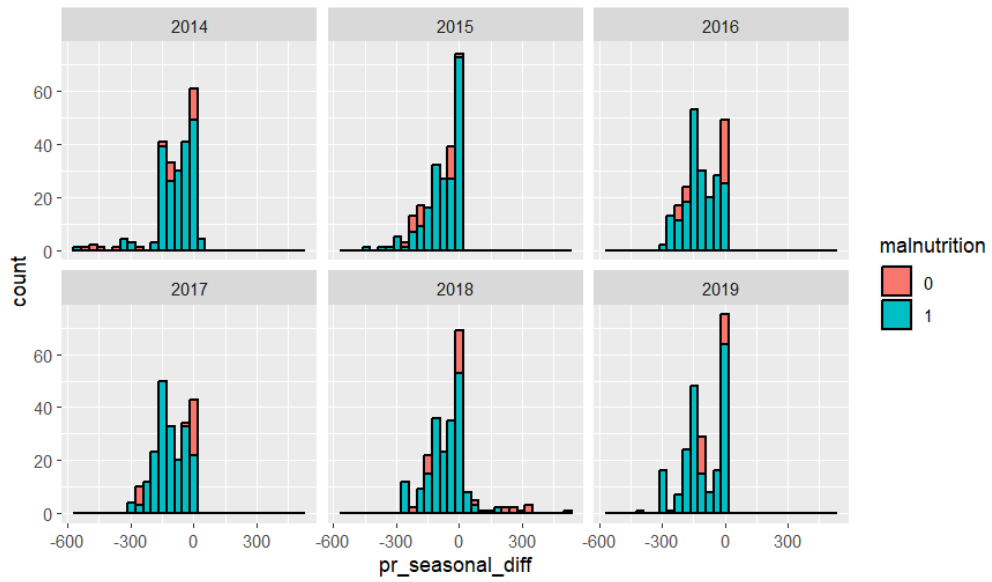


Figura 11. Distribución de la variable diferencia estacional de precipitación en una muestra de 200 niños. Fuente: Elaboración propia.

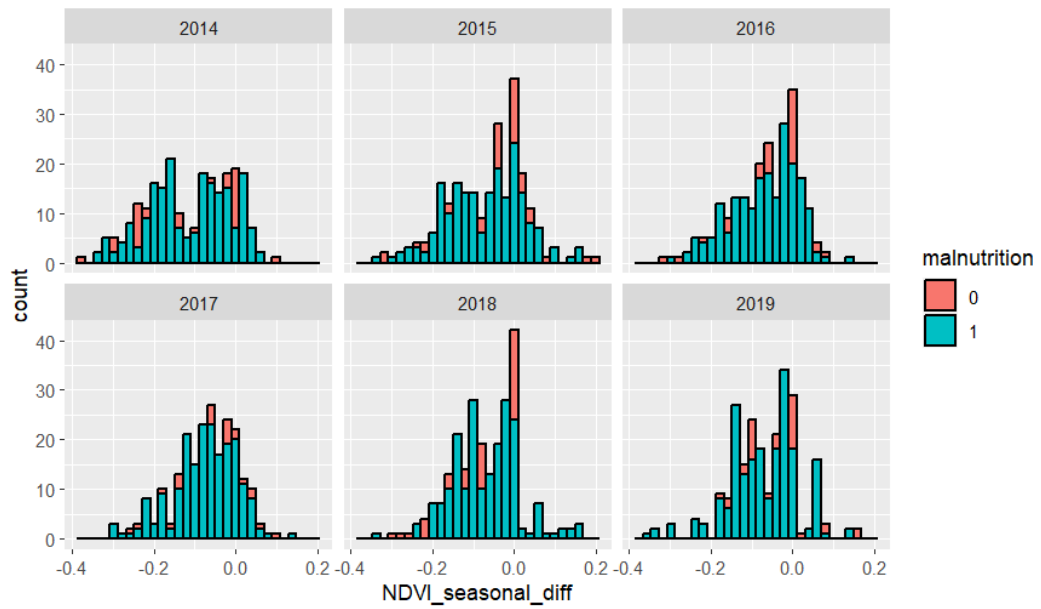


Figura 12. Distribución de la variable diferencia estacional de NDVI en una muestra de 200 niños. Fuente: Elaboración propia.

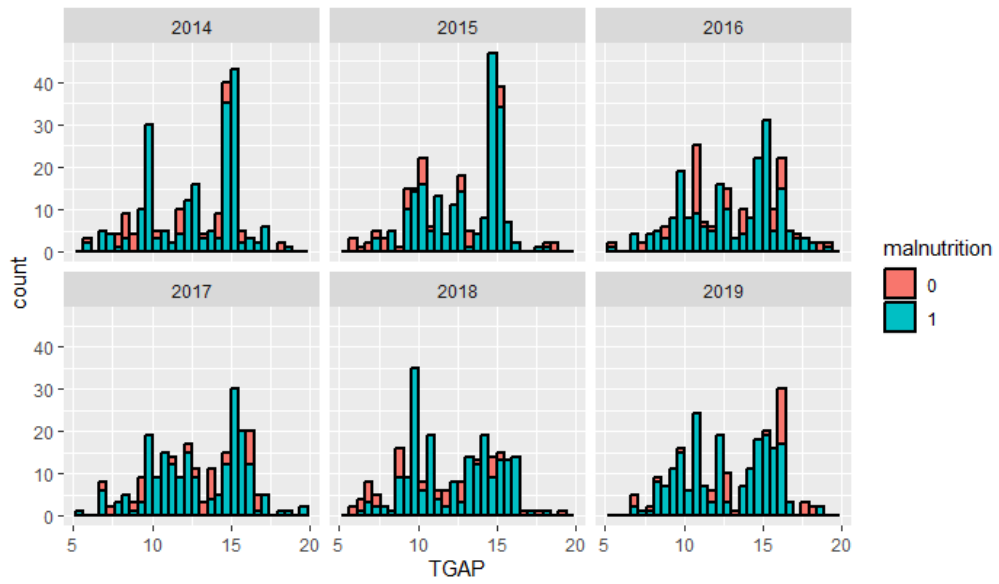


Figura 13. Distribución de la variable de diferencia de temperatura máxima y mínima en una muestra de 200 niños. Fuente: Elaboración propia.

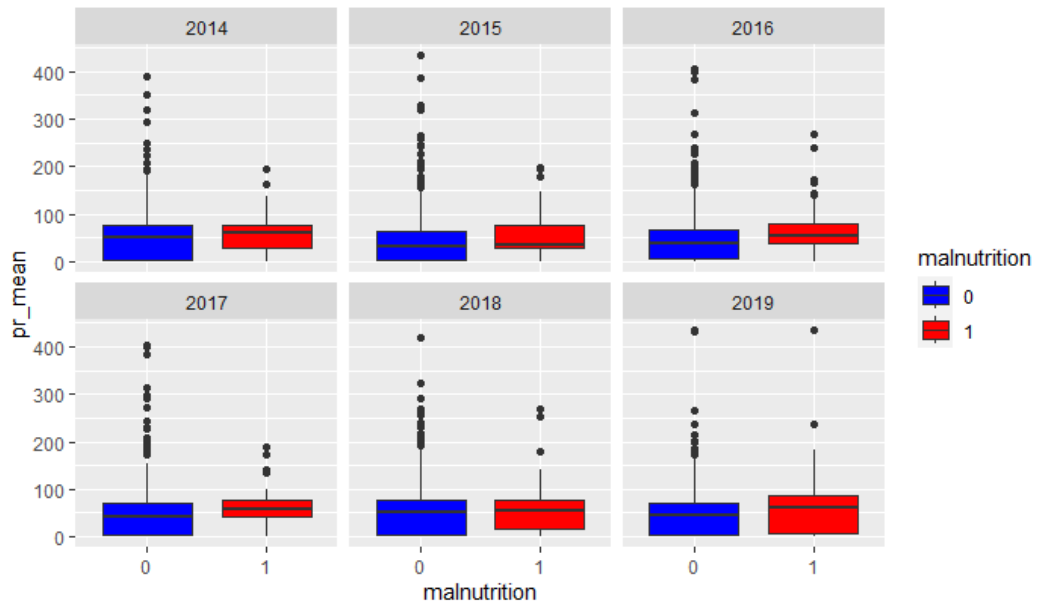


Figura 14. Gráfico de cajas y bigotes de la variable media de precipitación.

Fuente: Elaboración propia.

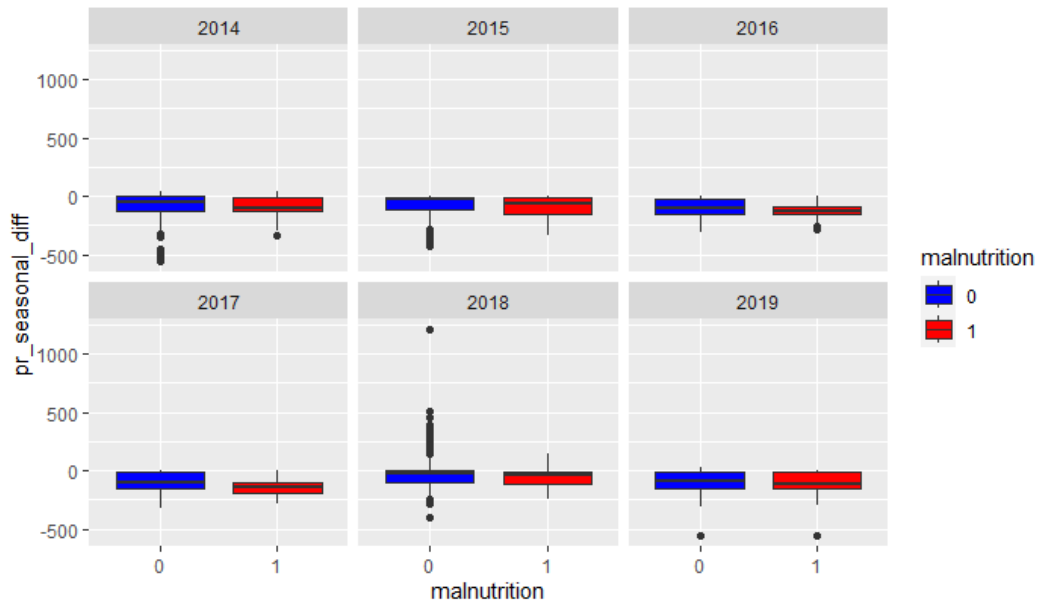


Figura 15. Gráfico de cajas y bigotes de la variable diferencia estacional de precipitación. Fuente: Elaboración propia.

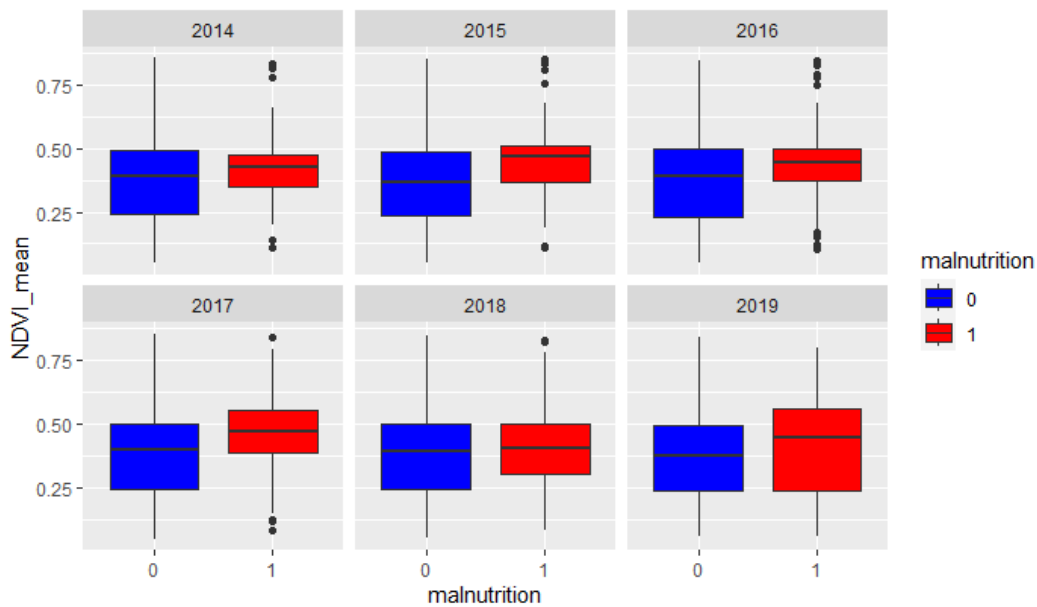


Figura 16. Gráfico de cajas y bigotes de la variable promedio de NDVI. Fuente: Elaboración propia.

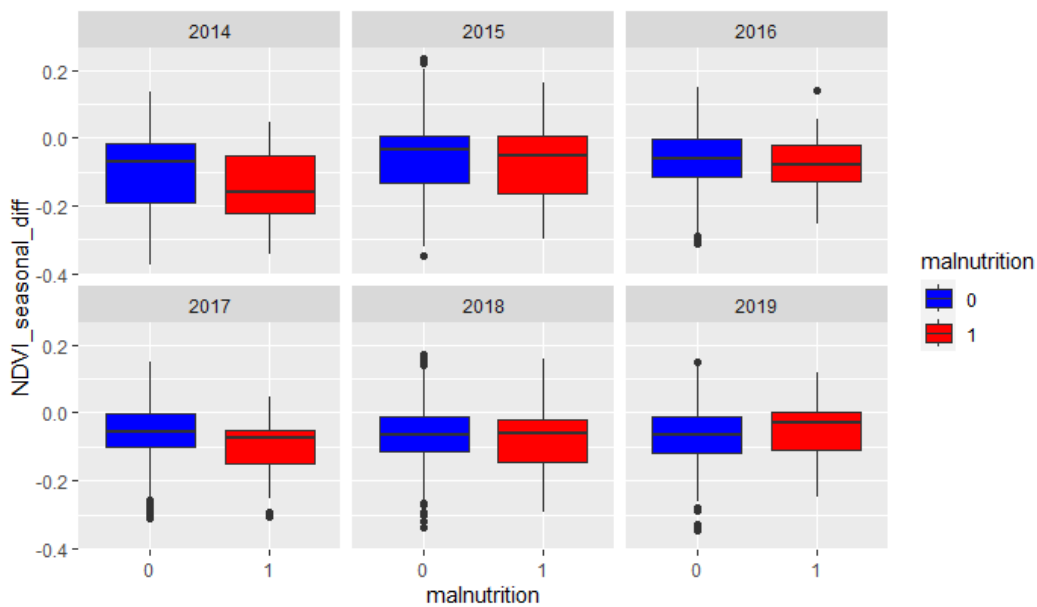


Figura 17. Gráfico de cajas y bigotes de la variable diferencia estacional del NDVI. Fuente: Elaboración propia.

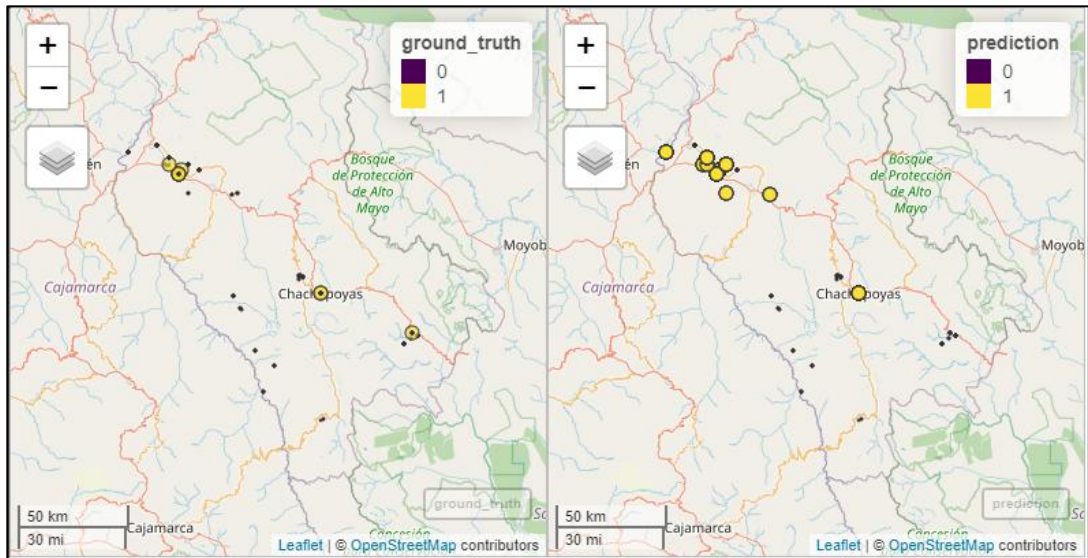


Figura 18. Visualización de casos de desnutrición aguda en el departamento de Amazonas en 2016. A la izquierda se observa el ground truth, mientras que a la derecha están las predicciones del modelo de Random Forest. Fuente: Elaboración propia.

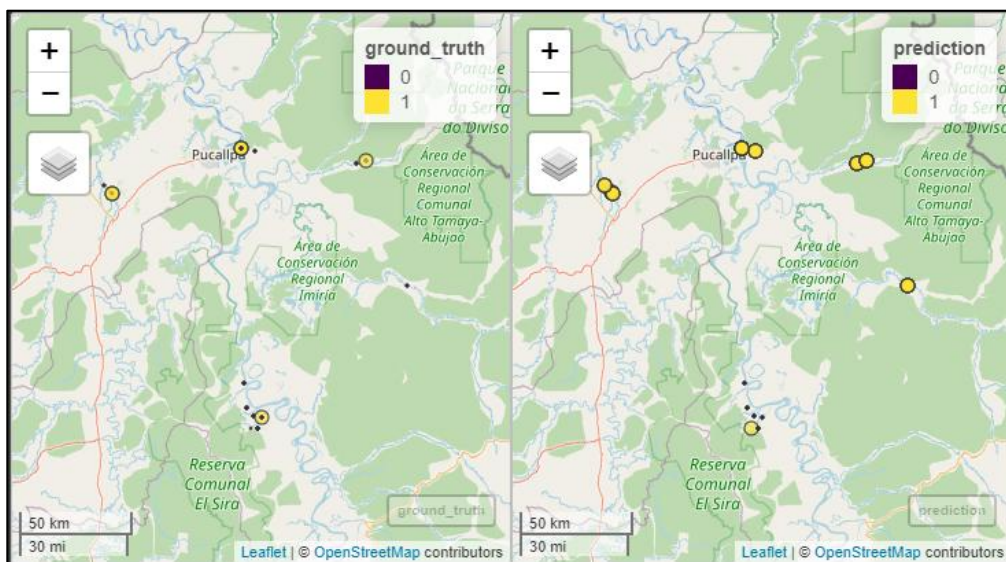


Figura 19. Visualización de casos de desnutrición aguda en el departamento de Ucayali en 2017. A la izquierda se observa el ground truth, mientras que a la derecha están las predicciones del modelo de Random Forest. Fuente: Elaboración propia.

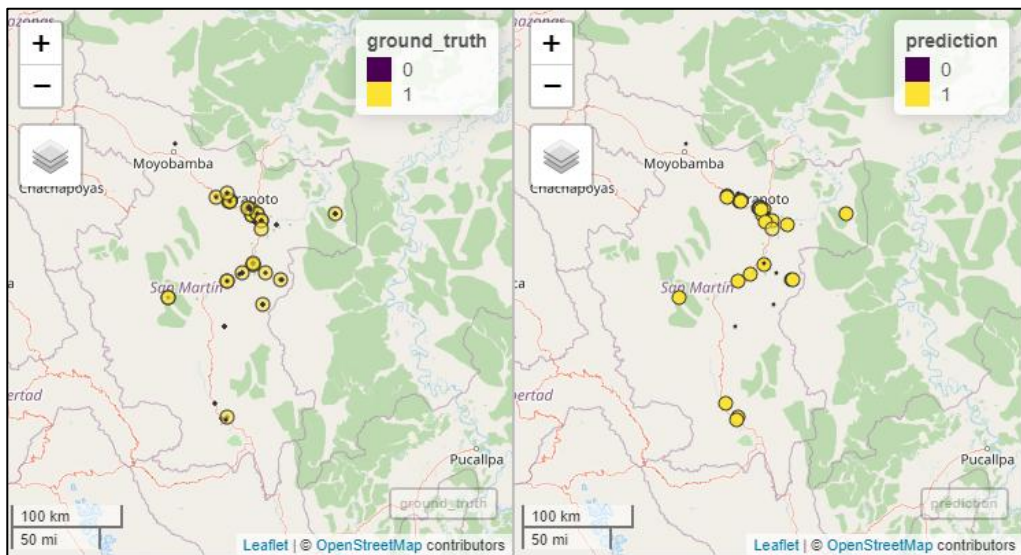


Figura 20. Visualización de casos de desnutrición aguda en el departamento de San Martín en 2018. A la izquierda se observa el ground truth, mientras que a la derecha están las predicciones del modelo de Random Forest. Fuente: Elaboración propia.

ANEXOS 2

Tabla 7. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2017. Fuente: Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 10</i>	<i>2</i>	<i>500</i>	<i>10</i>	<i>Kappa</i>	<i>0.36</i>
				<i>ROC AUC</i>	<i>0.69</i>
				<i>Sensitivity</i>	<i>0.69</i>

Tabla 8. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2018. Fuente: Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.91</i>
<i>Model 20</i>	<i>3</i>	<i>965</i>	<i>2</i>	<i>Kappa</i>	<i>0.38</i>
				<i>ROC AUC</i>	<i>0.70</i>
				<i>Sensitivity</i>	<i>0.68</i>

Tabla 9. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2016. Fuente:

Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>

<i>Model 03</i>	5	1500	5	<i>Kappa</i>	0.09
				<i>ROC AUC</i>	0.57
				<i>Sensitivity</i>	0.66

Tabla 10. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2017. Fuente:

Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
<i>Model 09</i>	2	300	20	<i>Accuracy</i>	0.97
				<i>Kappa</i>	0.07
				<i>ROC AUC</i>	0.55
				<i>Sensitivity</i>	0.63

Tabla 11. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2018. Fuente:

Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.92</i>
<i>Model 07</i>	<i>6</i>	<i>150</i>	<i>1</i>	<i>Kappa</i>	<i>0.15</i>
				<i>ROC AUC</i>	<i>0.60</i>
				<i>Sensitivity</i>	<i>0.73</i>

Tabla 12. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de Random Forest para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2019. Fuente:

Elaboración propia

<i>config</i>	<i>mtry</i>	<i>trees</i>	<i>min_n</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 04</i>	<i>3</i>	<i>500</i>	<i>7</i>	<i>Kappa</i>	<i>0.10</i>
				<i>ROC AUC</i>	<i>0.59</i>
				<i>Sensitivity</i>	<i>0.59</i>

Tabla 13. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2016. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 18</i>	<i>10</i>	<i>6</i>	<i>0.1</i>	<i>Kappa</i>	<i>0.01</i>
				<i>ROC AUC</i>	<i>0.51</i>
				<i>Sensitivity</i>	<i>0.95</i>

Tabla 14. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2017. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.96</i>
<i>Model 19</i>	<i>5</i>	<i>4</i>	<i>0.3</i>	<i>Kappa</i>	<i>0.01</i>
				<i>ROC AUC</i>	<i>0.55</i>
				<i>Sensitivity</i>	<i>0.99</i>

Tabla 15. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2018. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 01</i>	<i>2</i>	<i>10</i>	<i>0.004</i>	<i>Kappa</i>	<i>0.02</i>
				<i>ROC AUC</i>	<i>0.59</i>
				<i>Sensitivity</i>	<i>0.90</i>

Tabla 16. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas descriptivas y conteos de meses en condiciones climáticas críticas. 2014 a 2019. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 15</i>	<i>7</i>	<i>14</i>	<i>0.05</i>	<i>Kappa</i>	<i>0.02</i>
				<i>ROC AUC</i>	<i>0.55</i>
				<i>Sensitivity</i>	<i>0.95</i>

Tabla 17. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2016. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.97</i>
<i>Model 02</i>	<i>20</i>	<i>2</i>	<i>0.02</i>	<i>Kappa</i>	<i>0.01</i>
				<i>ROC AUC</i>	<i>0.58</i>
				<i>Sensitivity</i>	<i>0.92</i>

Tabla 18. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2017. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>
<i>Model 02</i>	<i>15</i>	<i>9</i>	<i>0.01</i>	<i>Kappa</i>	<i>0.01</i>
				<i>ROC AUC</i>	<i>0.61</i>
				<i>Sensitivity</i>	<i>0.99</i>

Tabla 19. Búsqueda de hiperparámetros a partir de validación cruzada en modelo de XGBoost para variables climáticas, meses consecutivos y grupos de meses consecutivos en condiciones climáticas críticas. 2014 a 2019. Fuente: Elaboración propia

<i>config</i>	<i>min_n</i>	<i>tree_depth</i>	<i>learn_rate</i>	<i>metric</i>	<i>mean</i>
				<i>Accuracy</i>	<i>0.95</i>

				<i>Kappa</i>	0.03
<i>Model 9</i>	2	11	0.01	<i>ROC AUC</i>	0.57
			0.56		
				<i>Sensitivity</i>	0.99

1 Tabla 20. Análisis por región con segmentación por edades para el modelo basado en variables climáticas. Fuente: Elaboración Propia

<i>Departamento</i>	<u>2016</u>					<u>2017</u>					<u>2018</u>					<u>2019</u>								
	6-11	12-23	24-35	36-47	48-59	6-11	12-23	24-35	36-47	48-59	6-11	12-23	24-35	36-47	48-59	6-11	12-23	24-35	36-47	48-59				
<i>Amazonas</i>	0,88	0,91	0,90	0,88	0,89	0,83	0,85	0,8	0,85	0,87	0,88	0,86	0,75	0,76	0,75	0,77	0,78	0,70	0,78	0,79	0,78	0,8	0,81	0,73
<i>Ancash</i>	0,29	0,24	0,29	0,31	0,32	0,3	0,50	0,45	0,5	0,52	0,53	0,51	0,29	0,24	0,29	0,31	0,32	0,3	0,20	0,15	0,2	0,22	0,23	0,21
<i>Apurimac</i>	0,94	0,97	0,94	0,96	0,95	0,89	0,70	0,65	0,7	0,72	0,73	0,71	0,60	0,55	0,6	0,62	0,63	0,61	1,00	0,95	1,00	1,00	1,00	1,00
<i>Arequipa</i>	0,23	0,18	0,23	0,25	0,26	0,24	0,30	0,25	0,3	0,32	0,33	0,31	0,35	0,3	0,35	0,37	0,38	0,36	0,33	0,28	0,33	0,35	0,36	0,34
<i>Ayacucho</i>	0,79	0,74	0,79	0,81	0,82	0,8	0,36	0,31	0,36	0,38	0,39	0,37	0,81	0,76	0,81	0,83	0,84	0,82	1,00	0,95	1,00	1,00	1,00	1,00
<i>Cajamarca</i>	0,46	0,41	0,46	0,48	0,49	0,47	0,46	0,41	0,46	0,48	0,49	0,47	0,64	0,59	0,64	0,66	0,67	0,65	0,40	0,35	0,4	0,42	0,43	0,41
<i>Cusco</i>	0,84	0,79	0,84	0,86	0,87	0,85	0,41	0,36	0,41	0,43	0,44	0,42	0,23	0,18	0,23	0,25	0,26	0,24	1,00	0,95	1,00	1,00	1,00	1,00
<i>Huancavelica</i>	0,40	0,35	0,4	0,42	0,43	0,41	0,50	0,45	0,5	0,52	0,53	0,51	0,88	0,83	0,88	0,9	0,91	0,89	0,67	0,62	0,67	0,69	0,7	0,68
<i>Huánuco</i>	0,71	0,66	0,71	0,73	0,74	0,72	0,70	0,65	0,7	0,72	0,73	0,71	0,89	0,84	0,89	0,91	0,92	0,9	1,00	0,95	1,00	1,00	1,00	1,00
<i>Ica</i>	0,20	0,15	0,2	0,22	0,23	0,21	0,25	0,2	0,25	0,27	0,28	0,26	0,25	0,2	0,25	0,27	0,28	0,26	0,20	0,15	0,2	0,22	0,23	0,21
<i>Junín</i>	0,24	0,19	0,24	0,26	0,27	0,25	0,38	0,33	0,38	0,4	0,41	0,39	0,86	0,81	0,86	0,88	0,89	0,87	0,38	0,33	0,38	0,4	0,41	0,39
<i>Libertad</i>	0,63	0,58	0,63	0,65	0,66	0,64	0,40	0,35	0,4	0,42	0,43	0,41	0,53	0,48	0,53	0,55	0,56	0,54	0,55	0,5	0,55	0,57	0,58	0,56
<i>Lambayeque</i>	0,45	0,4	0,45	0,47	0,48	0,46	0,35	0,3	0,35	0,37	0,38	0,36	0,29	0,24	0,29	0,31	0,32	0,3	0,36	0,31	0,36	0,38	0,39	0,37
<i>Lima</i>	0,44	0,39	0,44	0,46	0,47	0,45	0,20	0,15	0,2	0,22	0,23	0,21	0,25	0,2	0,25	0,27	0,28	0,26	0,40	0,35	0,4	0,42	0,43	0,41
<i>Loreto</i>	1,00	0,95	1,00	0,98	0,99	1,00	0,95	0,9	0,96	0,97	0,98	0,95	0,87	0,82	0,87	0,89	0,9	0,88	0,90	0,91	0,9	0,92	0,93	0,85
<i>Madre de Dios</i>	0,75	0,76	0,78	0,77	0,75	0,70	0,80	0,75	0,81	0,82	0,83	0,80	0,83	0,84	0,83	0,85	0,86	0,78	0,70	0,65	0,7	0,72	0,73	0,71
<i>Moquegua</i>	0,50	0,45	0,5	0,52	0,53	0,51	0,60	0,55	0,6	0,62	0,63	0,61	0,47	0,42	0,47	0,49	0,5	0,48	0,50	0,45	0,5	0,52	0,53	0,51
<i>Pasco</i>	0,25	0,2	0,25	0,27	0,28	0,26	0,40	0,35	0,4	0,42	0,43	0,41	0,90	0,85	0,9	0,92	0,93	0,91	0,60	0,55	0,6	0,62	0,63	0,61
<i>Piura</i>	0,23	0,18	0,23	0,25	0,26	0,24	0,29	0,24	0,29	0,31	0,32	0,3	0,23	0,18	0,23	0,25	0,26	0,24	0,38	0,33	0,38	0,4	0,41	0,39

<i>Puno</i>	0,47	0,42	0,47	0,49	0,5	0,48	0,70	0,65	0,7	0,72	0,73	0,71	0,50	0,45	0,5	0,52	0,53	0,51	0,67	0,62	0,67	0,69	0,7	0,68
<i>San Martin</i>	0,90	0,91	0,93	0,92	0,90	0,85	0,88	0,89	0,88	0,9	0,91	0,83	0,87	0,88	0,87	0,89	0,9	0,82	0,87	0,88	0,87	0,89	0,9	0,82
<i>Tacna</i>	0,20	0,15	0,2	0,22	0,23	0,21	0,20	0,15	0,2	0,22	0,23	0,21	0,20	0,15	0,2	0,22	0,23	0,21	0,20	0,15	0,2	0,22	0,23	0,21
<i>Tumbes</i>	0,67	0,62	0,67	0,69	0,7	0,68	0,42	0,37	0,42	0,44	0,45	0,43	0,43	0,38	0,43	0,45	0,46	0,44	0,56	0,51	0,56	0,58	0,59	0,57
<i>Ucayali</i>	0,90	0,91	0,9	0,92	0,93	0,88	0,95	0,96	0,95	0,97	0,98	0,9	1,00	0,95	1,00	0,99	1,00	0,99	1,00	0,95	1,00	0,99	0,99	1,00

2

3 Tabla 21. Análisis por región con segmentación por edades para el modelo basado en variables sociodemográficas. Fuente: Elaboración
4 Propia

<i>Departamento</i>	<u>2016</u>					<u>2017</u>					<u>2018</u>					<u>2019</u>								
	<i>6-11</i>	<i>12-23</i>	<i>24-35</i>	<i>36-47</i>	<i>48-59</i>	<i>6-11</i>	<i>12-23</i>	<i>24-35</i>	<i>36-47</i>	<i>48-59</i>	<i>6-11</i>	<i>12-23</i>	<i>24-35</i>	<i>36-47</i>	<i>48-59</i>	<i>6-11</i>	<i>12-23</i>	<i>24-35</i>	<i>36-47</i>	<i>48-59</i>				
<i>Amazonas</i>	0,88	0,91	0,90	0,88	0,89	0,83	0,85	0,8	0,85	0,87	0,88	0,86	0,75	0,76	0,75	0,77	0,78	0,70	0,78	0,79	0,78	0,8	0,81	0,73
<i>Ancash</i>	0,29	0,24	0,29	0,31	0,32	0,3	0,50	0,45	0,5	0,52	0,53	0,51	0,29	0,24	0,29	0,31	0,32	0,3	0,20	0,15	0,2	0,22	0,23	0,21
<i>Apurimac</i>	0,94	0,97	0,94	0,96	0,95	0,89	0,70	0,65	0,7	0,72	0,73	0,71	0,60	0,55	0,6	0,62	0,63	0,61	1,00	0,95	1,00	1,00	1,00	1,00
<i>Arequipa</i>	0,23	0,18	0,23	0,25	0,26	0,24	0,30	0,25	0,3	0,32	0,33	0,31	0,35	0,3	0,35	0,37	0,38	0,36	0,33	0,28	0,33	0,35	0,36	0,34
<i>Ayacucho</i>	0,79	0,74	0,79	0,81	0,82	0,8	0,36	0,31	0,36	0,38	0,39	0,37	0,81	0,76	0,81	0,83	0,84	0,82	1,00	0,95	1,00	1,00	1,00	1,00
<i>Cajamarca</i>	0,46	0,41	0,46	0,48	0,49	0,47	0,46	0,41	0,46	0,48	0,49	0,47	0,64	0,59	0,64	0,66	0,67	0,65	0,40	0,35	0,4	0,42	0,43	0,41
<i>Cusco</i>	0,84	0,79	0,84	0,86	0,87	0,85	0,41	0,36	0,41	0,43	0,44	0,42	0,23	0,18	0,23	0,25	0,26	0,24	1,00	0,95	1,00	1,00	1,00	1,00
<i>Huancavelica</i>	0,40	0,35	0,4	0,42	0,43	0,41	0,50	0,45	0,5	0,52	0,53	0,51	0,88	0,83	0,88	0,9	0,91	0,89	0,67	0,62	0,67	0,69	0,7	0,68
<i>Huánuco</i>	0,71	0,66	0,71	0,73	0,74	0,72	0,70	0,65	0,7	0,72	0,73	0,71	0,89	0,84	0,89	0,91	0,92	0,9	1,00	0,95	1,00	1,00	1,00	1,00
<i>Ica</i>	0,20	0,15	0,2	0,22	0,23	0,21	0,25	0,2	0,25	0,27	0,28	0,26	0,25	0,2	0,25	0,27	0,28	0,26	0,20	0,15	0,2	0,22	0,23	0,21
<i>Junín</i>	0,24	0,19	0,24	0,26	0,27	0,25	0,38	0,33	0,38	0,4	0,41	0,39	0,86	0,81	0,86	0,88	0,89	0,87	0,38	0,33	0,38	0,4	0,41	0,39
<i>Libertad</i>	0,63	0,58	0,63	0,65	0,66	0,64	0,40	0,35	0,4	0,42	0,43	0,41	0,53	0,48	0,53	0,55	0,56	0,54	0,55	0,5	0,55	0,57	0,58	0,56
<i>Lambayeque</i>	0,45	0,4	0,45	0,47	0,48	0,46	0,35	0,3	0,35	0,37	0,38	0,36	0,29	0,24	0,29	0,31	0,32	0,3	0,36	0,31	0,36	0,38	0,39	0,37
<i>Lima</i>	0,44	0,39	0,44	0,46	0,47	0,45	0,20	0,15	0,2	0,22	0,23	0,21	0,25	0,2	0,25	0,27	0,28	0,26	0,40	0,35	0,4	0,42	0,43	0,41
<i>Loreto</i>	1,00	0,95	1,00	0,98	0,99	1,00	0,95	0,9	0,96	0,97	0,98	0,95	0,87	0,82	0,87	0,89	0,9	0,88	0,90	0,91	0,9	0,92	0,93	0,85
<i>Madre de Dios</i>	0,75	0,76	0,78	0,77	0,75	0,70	0,80	0,75	0,81	0,82	0,83	0,80	0,83	0,84	0,83	0,85	0,86	0,78	0,70	0,65	0,7	0,72	0,73	0,71
<i>Moquegua</i>	0,50	0,45	0,5	0,52	0,53	0,51	0,60	0,55	0,6	0,62	0,63	0,61	0,47	0,42	0,47	0,49	0,5	0,48	0,50	0,45	0,5	0,52	0,53	0,51
<i>Pasco</i>	0,25	0,2	0,25	0,27	0,28	0,26	0,40	0,35	0,4	0,42	0,43	0,41	0,90	0,85	0,9	0,92	0,93	0,91	0,60	0,55	0,6	0,62	0,63	0,61

<i>Piura</i>	0,23	0,18	0,23	0,25	0,26	0,24	0,29	0,24	0,29	0,31	0,32	0,3	0,23	0,18	0,23	0,25	0,26	0,24	0,38	0,33	0,38	0,4	0,41	0,39
<i>Puno</i>	0,47	0,42	0,47	0,49	0,5	0,48	0,70	0,65	0,7	0,72	0,73	0,71	0,50	0,45	0,5	0,52	0,53	0,51	0,67	0,62	0,67	0,69	0,7	0,68
<i>San Martin</i>	0,90	0,91	0,93	0,92	0,90	0,85	0,88	0,89	0,88	0,9	0,91	0,83	0,87	0,88	0,87	0,89	0,9	0,82	0,87	0,88	0,87	0,89	0,9	0,82
<i>Tacna</i>	0,20	0,15	0,2	0,22	0,23	0,21	0,20	0,15	0,2	0,22	0,23	0,21	0,20	0,15	0,2	0,22	0,23	0,21	0,20	0,15	0,2	0,22	0,23	0,21
<i>Tumbes</i>	0,67	0,62	0,67	0,69	0,7	0,68	0,42	0,37	0,42	0,44	0,45	0,43	0,43	0,38	0,43	0,45	0,46	0,44	0,56	0,51	0,56	0,58	0,59	0,57
<i>Ucayali</i>	0,90	0,91	0,9	0,92	0,93	0,88	0,95	0,96	0,95	0,97	0,98	0,9	1,00	0,95	1,00	0,99	1,00	0,99	1,00	0,95	1,00	0,99	0,99	1,00

6

ANEXOS 3

7 **Repositorio:** https://github.com/healthinnovation/malnutrition_climate