

# Implementación de un modelo de Machine Learning para la clasificación de la desnutrición e inseguridad alimentaria en población infantil de los hogares del departamento de Antioquia, Colombia.

Huberth Rolando Hincapié Arango    Yeisson Alejandro Gutiérrez Cano  
Ingeniero Electrónico                      Profesional Gesis  
Universidad de Antioquia                  Universidad de Antioquia  
Medellín, Colombia                      Medellín, Colombia  
huberth.hincapie@udea.edu.co    yeisson.gutierrez@udea.edu.co

Daniel Escobar Saltaren                      Maria Bernarda Salazar Sánchez  
MSc Ingeniería                              PhD Ingeniería  
Universidad de Antioquia                  Universidad de Antioquia  
Medellín, Colombia                      Medellín, Colombia  
daniel.escobars@udea.edu.co    bernarda.salazar@udea.edu.co

**Resumen:** ~~This electronic document is a “live” template and already defines the components of your paper [title, text, heads, etc.] in its style sheet. \*CRITICAL: Do Not Use Symbols, Special Characters, Footnotes, or Math in Paper Title or Abstract. (Abstract)~~

**Palabras clave:** desnutrición, inseguridad alimentaria, análisis predictivo, políticas públicas.

## I. INTRODUCCIÓN

Alrededor del 38.8% de la población mundial enfrenta el hambre crónica y la inseguridad alimentaria moderada o grave [1]. En América Latina la desnutrición y las prevalencias de talla baja han sido un problema reconocido y analizado mediante estudios detallados [2], [3], además, en Colombia dicha problemática se refleja en las proporciones de mortalidad por desnutrición en población de 0 a 4 años (17.5% en 2019), que aunque han disminuido, estas presentan un incremento en adultos de 85 y más años (34.5% en 2019) [4].

Múltiples autores han comentado sobre los efectos de la desnutrición infantil y el impacto en la mortalidad [4], problemática social que cobra la vida de más de 6,800 niños al día; además, su asociación con problemas en el desarrollo físico y cognitivo, afectando el aprendizaje, el pensamiento y el habla [5], la reducción en la capacidad intelectual, su influencia sobre patrones de comportamiento durante la adultez [6], y la afectación sobre la productividad económica de los individuos y las sociedades futuras [7].

En Antioquia son las desigualdades sociales y económicas las que explican las altas prevalencias de inseguridad alimentaria en los hogares, el habitar áreas rurales (75.8%) y vivir en condiciones socioeconómicas deterioradas los hacen más vulnerables [8]. La capital del departamento para la década de 2011 a 2021 presentó una leve disminución del 3% en la desnutrición infantil, siendo este un problema observado a nivel local, regional, nacional e internacional [9].

Diferentes organizaciones e informes internacionales han estudiado la inseguridad alimentaria, el hambre y la desnutrición; la Organización de las Naciones Unidas para la Alimentación y la Agricultura -FAO- asegura que entre 691 y

783 millones de personas padecieron hambre en 2022 [10], el Fondo de las Naciones Unidas para Infancia -UNICEF- identifica como la desnutrición crónica en México afecta a 1 de cada 8 niños y niñas en la primera infancia, y además, plantea retos y soluciones por medio de políticas de salud pública integrales [11].

La Organización Mundial de la Salud -OMS- plantea que una buena nutrición influye positivamente en la salud de los lactantes, así como en embarazos y partos más seguros [12]. Según esta organización, los niños desnutridos son más vulnerables ante la enfermedad y la muerte, 149 millones de niños menores de 5 años en todo el mundo presentaban retraso del crecimiento en 2022 y casi la mitad de las defunciones en esta población, tienen que ver con la desnutrición [13]. Por otra parte la FAO y otras instituciones evidencian un aumento en la insuficiencia de consumo habitual de alimentos y la inseguridad alimentaria en el mundo desde 2014, proponiendo la ejecución de acciones urgentes [14].

La problemática de la desnutrición y la inseguridad alimentaria ha sido abordada desde diversos contextos: escolares [3], territoriales [15], y considerando factores y determinantes sociodemográficos [7], [16], así como programas sociales [17]. Además, se han utilizado técnicas de análisis estadístico y predictivo que van desde un modelo univariante [6], hasta modelos de regresión logística y lineal [2], [3], [7], [9], [17]–[20], árboles de decisión [5], [21], Análisis de Correspondencias Múltiples [18] y técnicas Random Forest y AdaBoost [22]. De las diferentes técnica implementadas, se destaca la regresión logística logrando hasta 95% en su coeficiente de determinación  $R^2$ , y Random Forest con una sensibilidad del 90.03%; otros modelos como el AdaBoost alcanzaron una sensibilidad hasta del 69.76% y los árboles de clasificación un poder predictivo del 59.8%.

A pesar de los diferentes estudios efectuados en el contexto de la desnutrición infantil y la inseguridad alimentaria a nivel regional y nacional, aun es necesario lograr un modelo que fortalezca el proceso de toma de decisiones y posibilite la formulación y aplicación de acciones enmarcadas en las políticas públicas, con mayor probabilidad de obtener

resultados efectivos e impactar positivamente los indicadores poblacionales. ¿Es posible entonces, accediendo a información recopilada por la Secretaría Seccional de Salud y Protección Social de Antioquia sobre aproximadamente 100 indicadores a nivel municipal para todo el departamento, y posteriormente con el desarrollo de técnicas de Machine Learning implementar un modelo para la clasificación de la incidencia de desnutrición en población infantil y la prevalencia de la inseguridad alimentaria en los hogares del territorio?

Para ello en la sección de materiales y métodos se presentan la fuente de datos y su preparación, variables, métricas y desempeños de los modelos trabajados, posteriormente, algunos resultados y conclusiones del trabajo **investigativo**.

## II. MATERIALES Y MÉTODOS

### A. Fuente de datos

La principal fuente de datos se originó en el Servicio de Información y Estadística de la Secretaría Seccional de Salud y Protección Social de Antioquia [23], el cual pone a disposición con acceso libre y seguro datos con resultados de los Indicadores de Desnutrición y Seguridad Alimentaria, además de otras características Poblacionales [24] y de Estadísticas Vitales [25].

Se contó con un conjunto de datos de 135,386 instancias, con la desagregación para los 125 municipios del departamento de Antioquia y la medición de 98 indicadores enmarcados en los Objetivos de Desarrollo Sostenibles -ODS-, identificando el código y nombre del Municipio, Número y nombre del indicador con su resultado, numerador y denominador, además del año, fuente, unidad de medida, y también, la identificación de cada meta para cada ODS, para el periodo comprendido inicialmente entre el año 2007 y el 2022.

En general, la confiabilidad de esta información se puede estimar como alta, ya que son datos reportados por parte de la administración a entidades de control del orden nacional, lo cual otorga capacidad de crítica y auditoria. Adicionalmente, esta información asume diferentes escenarios de uso y rendición de cuentas públicas, además contiene indicadores que están gobernados por diferentes leyes dentro del margen de política pública, al igual que otras leyes de orden presupuestal [23]–[25].

### B. Preparación de datos

El proceso de preparación de la base de datos se llevó a cabo por medio del lenguaje de programación de alto nivel Python, mediante el uso del entorno de notebook Jupyter Google Colaboratory, iniciando con la selección de las variables de interés para la **investigación** y procediendo con actividades de transformación, codificación, medidas de correlación, imputación, depuración, definición de categorías y construcción de modelos de clasificación.

#### 1. Selección de variables

La fuente de datos contó inicialmente con 18 variables para la identificación de cada indicador poblacional. Las variables seleccionadas para proceder con el análisis propuesto son las que identifican el municipio, el año y la medición para cada indicador, y se listan en la siguiente tabla (Tabla 1).

TABLA 1. VARIABLES SELECCIONADAS

Variable	Descripción
Código entidad	Código municipio nomenclatura estandarizada DANE
Entidad	Nombre del municipio o entidad territorial
Año	Año al que corresponde la medición del indicador
Indicador	Nombre del indicador
Valor	Medición del indicador

#### 2. Transformación estructura de datos

Inicialmente la estructura de los datos identificaba en cada registro la medición y las características de cada indicador, con el objetivo de obtener la estructura de datos idónea para el entrenamiento de los modelos planteados, se realizó una transformación conocida como pivoteo, haciendo uso de la función `pivot_table()` de pandas, la cual reorganiza el dataframe basada en los atributos del nombre del indicador llevándolo a las columnas y su medición para el valor de cada registro. De esta forma se obtiene un conjunto de datos que en cada columna cuenta con las variables predictoras y las dependientes (Figura 1).

Datos originales					
Index	Código entidad	Entidad	Año	Indicador	Valor
0	5674	San Vicente Ferre	2019	Incidencia de la pobreza monetaria extrema	17.9614254
1	5004	Abriaquí	2019	Incidencia de la pobreza monetaria extrema	11.34554324
2	5674	San Vicente Ferre	2019	Porcentaje de mujeres de 15 a 19 años con embarazo subsiguiente	33.33333333
3	5004	Abriaquí	2019	Porcentaje de mujeres de 15 a 19 años con embarazo subsiguiente	18.51851852
4	5674	San Vicente Ferre	2019	Tasa de fecundidad específica en mujeres adolescentes de 15 a 19 años	76.92307692
5	5004	Abriaquí	2019	Tasa de fecundidad específica en mujeres adolescentes de 15 a 19 años	32.72727273

Datos pivoteados					
Código entidad	Entidad	Año	Incidencia de la pobreza monetaria extrema	Porcentaje de mujeres de 15 a 19 años con embarazo subsiguiente	Tasa de fecundidad específica en mujeres adolescentes de 15 a 19 años
5674	San Vicente Ferre	2019	17.9614254	33.33333333	76.92307692
5004	Abriaquí	2019	11.34554324	18.51851852	32.72727273

FIGURA 1. PIVOTEO DE LOS DATOS

El proceso de pivoteo permitió obtener finalmente un conjunto de datos con 1,753 instancias y 101 columnas.

#### 3. Codificación de variables

Mediante el uso de un diccionario en Python denominado `dictVars`, el cual contiene el nombre inicial de las variables predictoras representadas por cada uno de los indicadores disponibles, se realizó su codificación mapeando desde V0 hasta V95, además de las variables de salida O1 para Incidencia de desnutrición aguda en población menor de 5 años y O2 Prevalencia de inseguridad alimentaria en el hogar (moderada o severa), con el objetivo de optimizar la visualización de los nombres de las variables para los diferentes indicadores y su manejo y utilización dentro del proceso de análisis y construcción de modelos predictivos.

#### 4. Exclusión de variables predictoras con alta correlación

A partir del cálculo de la matriz de correlación basada en el Coeficiente de correlación de Person, haciendo uso de la función `corr()` de pandas, se hizo una selección de las variables con una correlación débil o nula, es decir, con un coeficiente de correlación menor o igual a 0.3, se excluyeron las variables para las cuales se obtuvo una correlación moderada o fuerte, es decir, un coeficiente de correlación mayor a 0.3 [26]. La figura 2 compara en un mapa de calor la correlación de las diferentes variables, donde el tono rojizo determina una mayor correlación entre las mismas (Figura 2).

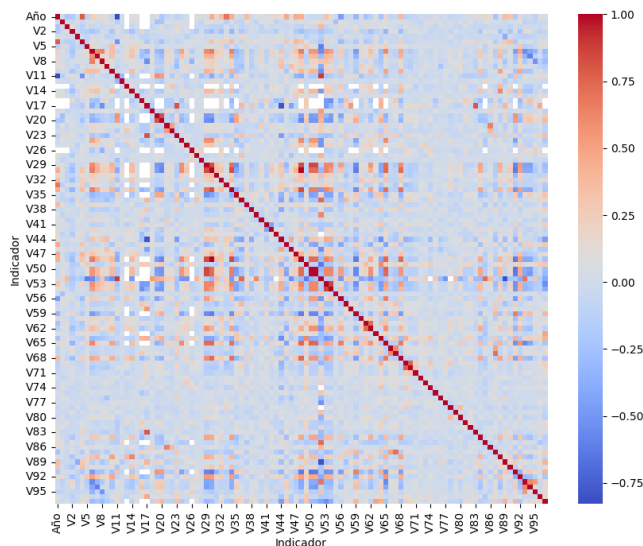


FIGURA 2. MATRIZ DE CORRELACIÓN PARA LAS VARIABLES PREDICTORAS

#### 5. Imputación de datos

Con la intención de optimizar el reemplazo de los valores faltantes en las mediciones de los indicadores poblacionales como variables predictoras (Figura 3), se eligió el método de imputación K-nearest Neighbors [27]. Haciendo uso de la clase `KNNImputer` de la biblioteca `scikit-learn` disponible en Python y utilizando un parámetro `n_neighbors` de 15, con el fin de captar mejor la estructura subyacente de los datos, se tomaron en cuenta los 15 vecinos más cercanos para la imputación de cada valor faltante, y un parámetro de ponderación de los vecinos `weights` "distance", donde los valores de los

vecinos más cercanos tienen un mayor impacto en el valor imputado que los valores de los vecinos más lejanos. Mediante el método `fit_transform()`, se aplicó el proceso de imputación a los datos sin incluir las variables dependientes con el fin de no perder los valores de la medición real de cada indicador de resultado [28], [29].

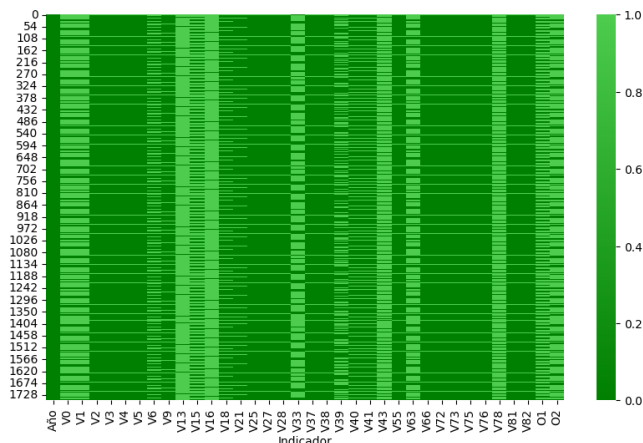


FIGURA 3. MAPEO DE DATOS FALTANTES PARA VARIABLES PREDICTORAS

#### 6. Depuración de registros con datos nulos en variables desenlace

Con el objetivo de depurar los registros con datos nulos en las variables dependientes, O1 Incidencia de desnutrición aguda en población menor de 5 años y O2 Prevalencia de inseguridad alimentaria en el hogar (moderada o severa), donde no se contó con una medición para el indicador, además, manteniendo la integridad de los datos desenlace, se usó el método `dropna` de pandas para excluir las instancias con valores NaN.

El proceso se llevó a cabo haciendo una separando del conjunto de datos por cada indicador, con el fin de excluir las instancias correspondientes manteniendo para cada desenlace la totalidad de mediciones contenidas en los datos, sin perder mediciones en el proceso de depuración.

#### 7. Categorización de variables dependientes

La medición de los indicadores desenlace O1 y O2, al ser inicialmente cuantitativa, no permite llevar a cabo el proceso de modelamiento para su clasificación, por ello se procedió mediante el uso de la función de pandas `qcut` a categorizar la incidencia de desnutrición aguda en población menor de 5 años y la prevalencia de inseguridad alimentaria en el hogar (moderada o severa), a partir de los cuartiles Q1, Q2 y Q3 en los niveles de Leve, Moderado y Alto la magnitud de cada indicador [30].

#### C. Construcción de modelos

Para experimentar y explorar de manera rápida y eficiente la construcción de los modelos de machine learning y poder comparar ágilmente entre diferentes técnicas y resultados, se usó la biblioteca de código abierto de Python `PyCaret` 3.0, permitiendo obtener métricas como la precisión (Accuracy),

el área bajo la curva ROC (AUC), la exhaustividad (Recall) o sensibilidad o tasa de verdaderos positivos, la precisión (Prec.), el puntaje F1 (F1-score), el coeficiente Kappa, y el coeficiente de correlación de Matthews (MCC) [31].

El uso de PyCaret permitió dentro de su `setup` visualizar la magnitud de los datos de entrenamiento y testeo, como también el número de pliegues de entrenamiento, además, el resultado de cada una de las métricas para los diferentes modelos de clasificación entrenados. Posterior se hizo una priorización de los 5 modelos más relevantes dentro de los resultados obtenidos para proceder a su despliegue con el objetivo de lograr la clasificación más óptima para ambos indicadores desenlace.

correr lo modelos con hiperparámetros ....

### III. RESULTADOS

Hacer un recorrido del notebook, figuras, resultados, con estilo, conectado con la metodología, los resultados para la limpieza, para la imputación en la figura xx

#### A. Fuente de datos

#### B. Variables seleccionadas

Se excluyeron 63 variables las cuales obtuvieron una correlación moderada o fuerte, y que para criterio obtuvieron un coeficiente de correlación de Pearson mayor a 0.3. Posterior a la depuración de dichas variables, se obtuvo una matriz de correlación que reflejó en un mapa de calor la ausencia de correlación entre las variables definitivas (Figura 4), las cuales representan un total de 33 indicadores poblacionales como posibles predictores de la incidencia de desnutrición aguda en población menor de 5 años y la prevalencia de inseguridad alimentaria en el hogar (moderada o severa).

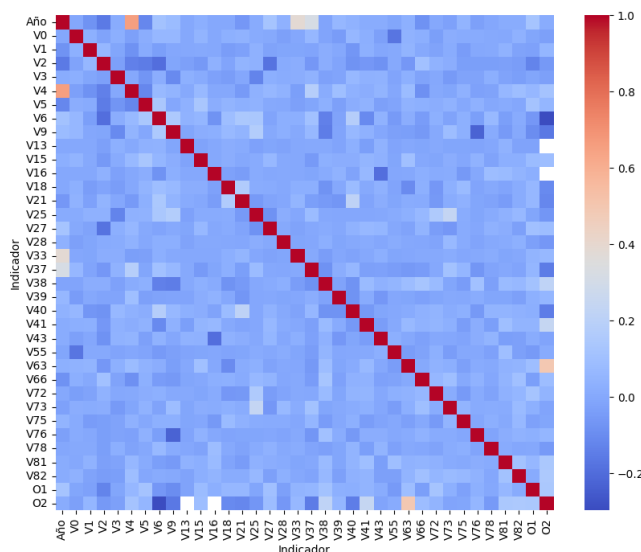


FIGURA 4. MATRIZ DE CORRELACIÓN PARA LAS VARIABLES PREDICTORAS

A continuación se listan los indicadores poblacionales definidos para la construcción de los modelos de machine learning como variables predictoras:

1. V0: Brecha de ingreso mensual promedio entre hombres y mujeres
2. V1: Brecha de pobreza extrema urbano-rural
3. V2: Brecha entre cobertura neta urbano-rural en educación básica y media
4. V3: Capacidad instalada de generación de energía eléctrica
5. V4: Cobertura de establecimientos financieros
6. V5: Cobertura de vacunación con BCG en nacidos vivos
7. V6: Cobertura de viviendas con agua potable
8. V9: Cobertura de viviendas con energía eléctrica
9. V13: Entidades territoriales con planes territoriales de cambio climático
10. V15: Hectáreas de pequeña y mediana propiedad rural, formalizadas
11. V16: Hogares urbanos con déficit cualitativo de vivienda
12. V18: Incidencia de intoxicaciones por sustancias químicas
13. V21: Incidencia de violencia intrafamiliar (Atención x 2 fuentes)
14. V25: Intensidad energética
15. V27: Miles de hectáreas de áreas protegidas
16. V28: Mortalidad departamental causada por eventos recurrentes
17. V33: Participación de mujeres en cargos de elección popular
18. V37: Porcentaje de menores de 1 año con tercera dosis de pentavalente
19. V38: Porcentaje de mujeres de 15 a 19 años con embarazo subsiguiente
20. V39: Porcentaje de mujeres en cargos directivos, según municipios de Antioquia
21. V40: Porcentaje de mujeres que han experimentado alguna violencia física por una persona diferente al esposo o compañero
22. V41: Porcentaje de mujeres que han sido forzadas físicamente por el esposo o compañero a tener relaciones o actos sexuales
23. V43: Porcentaje de municipios que cuentan con mecanismo intersectorial para el abordaje de las violencias de género
24. V55: Razón de mortalidad materna



25. V63: Tasa de desempleo
26. V66: Tasa de homicidio de mujeres (ocurrencia del hecho)
27. V72: Tasa de mortalidad por VIH/SIDA
28. V73: Tasa de mortalidad por accidentes de tránsito terrestre
29. V75: Tasa de mortalidad por lesiones auto infligidas intencionalmente (suicidio)
30. V76: Tasa de mortalidad por malaria
31. V78: Tasa de mortalidad por tumores malignos en menores de 18 años
32. V81: Tasa de mortalidad prematura por diabetes
33. V82: Tasa de mortalidad prematura por enfermedades cardiovasculares

### C. Imputación de datos

El proceso de imputación realizado sobre las 33 variables predictoras permitió obtener un conjunto de datos almacenado en nuevo dataframe llamado “df\_Imputed”, el cual conservo los índices y columnas del dataframe original. La siguiente figura evidencia la totalidad de datos para las variables predictoras posterior al proceso de imputación (Figura 5).

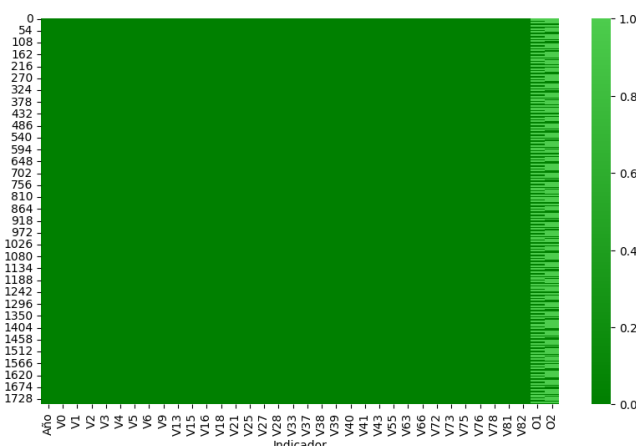


FIGURA 5. MAPEO DE DATOS IMPUTADOS PARA LAS VARIABLES PREDICTORA

### D. Depuración de registros con datos nulo en O1 y O2

La depuración de los registros con datos faltantes en los indicadores de la incidencia de desnutrición aguda en población menor de 5 años y la prevalencia de inseguridad alimentaria en el hogar (moderada o severa), permitió contar finalmente con mediciones reales para O1 en los años de 2017 a 2022, y para O2 en los años de 2013, 2017, 2019 y 2021.

Estos indicadores presentan una distribución con valores de medición a partir de cero y con datos extremos a los cuales no se les realizó manejo como atípicos, ya que corresponden a mediciones del indicador dentro de un contexto dinámico poblacional, y reflejan el comportamiento del evento en los territorios sujetos de dicha medición (Figura 6).

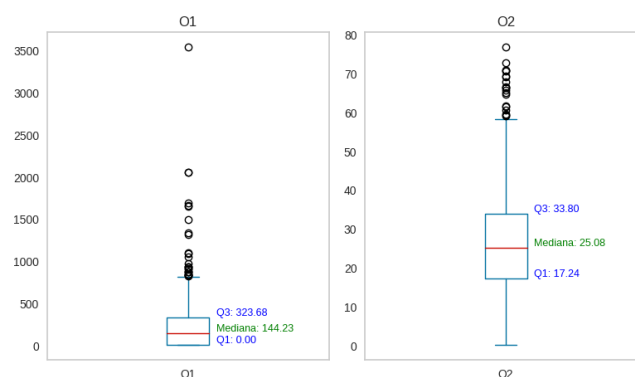


FIGURA 6. DISTRIBUCIÓN DE VARIABLES DESENLACE O1 Y O2

### E. Categorización de variables dependientes

La categorización de la incidencia de desnutrición aguda en población menor de 5 años permitió conocer el comportamiento de dicho indicador desde tres niveles de severidad, en un nivel leve con un rango que va desde 0 hasta 63.5 casos por 100 mil menores de 5 años, moderado hasta 240.5 casos por 100 mil menores de 5 años y alto hasta 3,533.6 casos por 100 mil menores de 5 años. En términos de frecuencia, el nivel alto presenta picos de mayor incidencia en la población de estudio (Figura 7).

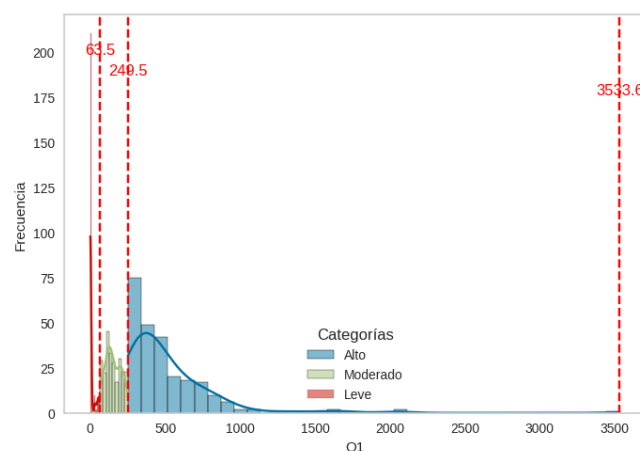


FIGURA 7. FRECUENCIA DE LA INCIDENCIA DE DESNUTRICIÓN AGUDA EN MENORES DE 5 AÑOS SEGÚN NIVEL DE SEVERIDAD

La categorización de la prevalencia de inseguridad alimentaria en el hogar (moderada o severa) permitió conocer el comportamiento de dicho indicador desde tres niveles de severidad, en un nivel leve con un rango que va desde 0 hasta 19,7%, moderado hasta 30.7% y alto hasta 76.6%. En términos de frecuencia, el nivel alto presenta picos de mayor prevalencia en la población de estudio (Figura 8).

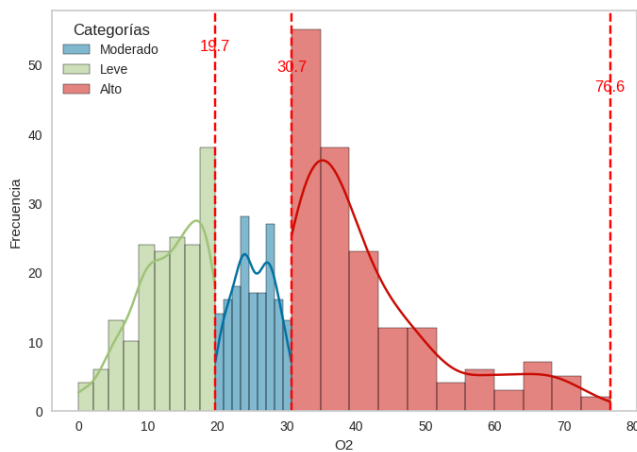


FIGURA 8. FRECUENCIA DE LA PREVALENCIA DE INSEGURIDAD ALIMENTARIA EN EL HOGAR SEGÚN NIVEL DE SEVERIDAD

### F. Construcción y métricas de los modelos

Mediante el uso de la clase `ClassificationExperiment()` contenida en PyCaret, fue posible con el método `setup`, establecer y visualizar la configuración del entorno de experimentación para cada indicador O1 y O2 (Tabla 2).

TABLA 2. CONFIGURACIÓN ENTORNO DE EXPERIMENTACIÓN VARIABLES DESENLAJE O1 Y O2

Descripción	Valores O1	Valores O2
Session id	123	135
Target	O1_bins	O2_bins
Target type	Multiclass	Multiclass
Target mapping	Alto: 0, Leve: 1, Moderado: 2	Alto: 0, Leve: 1, Moderado: 2
Original data shape	(750, 36)	(500, 36)
Transformed data shape	(750, 36)	(500, 36)
Transformed train set shape	(525, 36)	(350, 36)
Transformed test set shape	(225, 36)	(150, 36)
Numeric features	35	35
Preprocess	True	True
Imputation type	simple	simple
Numeric imputation	mean	mean
Categorical imputation	mode	mode
Fold Generator	StratifiedKfold	StratifiedKfold
Fold Number	10	10
CPU Jobs	-1	-1
Use GPU	False	False
Log Experiment	False	False
Experiment Name	clf-default-name	clf-default-name
USI	128a	3971

TABLA 3. MÉTRICAS DE DESEMPEÑO DE LOS MODELOS PRIORIZADOS PARA CADA VARIABLE DESENLAJE

Desenlace	Modelo		Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
Incidencia de desnutrición aguda en población menor de 5 años	dt	Decision Tree Classifier	0.9981	0.9986	0.9981	0.9982	0.9981	0.9972	0.9972	0.064
	ada	Ada Boost Classifier	0.9981	0	0.9981	0.9982	0.9981	0.9972	0.9972	0.449
	gbc	Gradient Boosting Classifier	0.9981	0	0.9981	0.9982	0.9981	0.9972	0.9972	0.722
	catboost	CatBoost Classifier	0.9981	1	0.9981	0.9982	0.9981	0.9972	0.9972	20.289
Prevalencia de inseguridad alimentaria en el hogar (moderada o severa)	dt	Decision Tree Classifier	0.9943	0.9957	0.9943	0.9947	0.9943	0.9914	0.9916	0.054
	ada	Ada Boost Classifier	0.9943	0	0.9943	0.9947	0.9943	0.9914	0.9916	0.205
	gbc	Gradient Boosting Classifier	0.9943	0	0.9943	0.9947	0.9943	0.9914	0.9916	0.791
	xgboost	Extreme Gradient Boosting	0.9943	0.9986	0.9943	0.9947	0.9943	0.9914	0.9917	0.157
	lightgbm	Light Gradient Boosting Machine	0.9943	0.997	0.9943	0.9947	0.9943	0.9914	0.9916	0.411
	catboost	CatBoost Classifier	0.9943	1	0.9943	0.9947	0.9943	0.9914	0.9916	18.281

## IV. CONCLUSIONES

Analizar el numerador y denominador sobre las mediciones no solo basado en la tasa calculada del evento

Que tiene mas peso,

Usabilidad de los modelos finales ...

## V. REFERENCES

[1] N. Unidas, “Informe de los Objetivos de Desarrollo Sostenible,” 2023. [Online]. Available: [https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023\\_Spanish.pdf?\\_gl=1\\*vtezeu\\*\\_ga\\*MjEyODc3Nzk3Mi4xNjk1MzQ5NzQ5\\*\\_ga\\_TK9BQL5X7Z\\*MTY5NTM0OTc0OS4xLjEuMTY5NTM1MDA4Ny4wLjAuMA](https://unstats.un.org/sdgs/report/2023/The-Sustainable-Development-Goals-Report-2023_Spanish.pdf?_gl=1*vtezeu*_ga*MjEyODc3Nzk3Mi4xNjk1MzQ5NzQ5*_ga_TK9BQL5X7Z*MTY5NTM0OTc0OS4xLjEuMTY5NTM1MDA4Ny4wLjAuMA)

[2] G. Estrada and L. Roldán, “Modelo estadístico para predecir la

- prevalencia de desnutrición crónica infantil en los departamentos de Guatemala,” *Rev. médica (Colegio Médicos y Cir. Guatemala)*, vol. 156, no. 2, pp. 61–66, 2017, doi: 10.36109/rmg.v156i2.57.
- [3] A. Ávila Curiel *et al.*, “Mala nutrición en población escolar mexicana: factores geográficos y escolares asociados,” *Glob. Health Promot.*, vol. 29, no. 2, pp. 126–135, 2021, doi: 10.1177/17579759211038381.
  - [4] L. C. Forero Ballesteros and A. Y. Forero Torres, “Tendencia temporal de la mortalidad por desnutrición en Colombia, 2005-2019,” *Rev. Panam. Salud Pública*, vol. 46, p. 1, 2022, doi: 10.26633/rpsp.2022.4.
  - [5] G. Congacha, “Comparación de modelos logísticos y árboles de decisión para identificar y predecir factores asociados a la desnutrición crónica infantil basados en la encuesta nacional de salud y nutrición – ENSANUT 2018-2019,” *Esc. Super. Politécnica Chimborazo*, pp. 1–67, 2020, [Online]. Available: <http://dspace.esPOCH.edu.ec/bitstream/123456789/14551/1/226T0063%281%29.pdf>
  - [6] C. Machaca, “Modelo Univariante Para La Predicción De La Desnutrición Crónica De Los Niños Menores De Cinco Años En El Hospital Regional Manuel Nuñez Butrón, Puno 2012 - 2016,” p. 98, 2017, [Online]. Available: [http://repositorio.unap.edu.pe/bitstream/handle/UNAP/9386/Machaca\\_Uriel\\_Clever\\_Uriel.pdf?sequence=3&isAllowed=y](http://repositorio.unap.edu.pe/bitstream/handle/UNAP/9386/Machaca_Uriel_Clever_Uriel.pdf?sequence=3&isAllowed=y)
  - [7] O. Reséndiz, C. Torres, A. Arredondo, Y. García, and E. Orozco, “Determinantes de la inseguridad alimentaria y su asociación con prácticas de alimentación y lactancia materna en el primer año de vida,” *Glob. Health Promot.*, vol. 29, no. 4, pp. 161–170, 2021, doi: 10.1177/17579759221076881.
  - [8] Gobernación de Antioquia and Universidad de Antioquia, “Perfil Alimentario y Nutricional de Antioquia 2019,” *Perf. Alimentario y Nutr. Antioquia 2019*, p. 70, 2019, [Online]. Available: [https://www.udea.edu.co/wps/wcm/connect/udea/ba8c4eac-c8ad-43aa-9a6f-7f9f143d28a0/resumen\\_ejecutivo\\_perfil\\_alimentario\\_FINAL.pdf?MOD=AJPERES&CVID=n3658bG](https://www.udea.edu.co/wps/wcm/connect/udea/ba8c4eac-c8ad-43aa-9a6f-7f9f143d28a0/resumen_ejecutivo_perfil_alimentario_FINAL.pdf?MOD=AJPERES&CVID=n3658bG)
  - [9] Á. A. Loaiza Duque, W. G. Moreno Puerta, and J. D. Rios Rios, “Proceso de analítica de datos aplicado a la desnutrición infantil en niños de 0 a 5 años en la ciudad de Medellín,” 2023, [Online]. Available: [http://siawebpr.itm.edu.co/bitstream/handle/20.500.12622/6011/AlvaroAndres\\_LoalizaDuque\\_WilliamGuillermo\\_MorenoPuerta\\_JuanDavid\\_RiosRios\\_2023.pdf?sequence=2&isAllowed=yes](http://siawebpr.itm.edu.co/bitstream/handle/20.500.12622/6011/AlvaroAndres_LoalizaDuque_WilliamGuillermo_MorenoPuerta_JuanDavid_RiosRios_2023.pdf?sequence=2&isAllowed=yes)
  - [10] Organización de las Naciones Unidas para la Alimentación y la Agricultura, “Hambre e inseguridad alimentaria,” 2023. [Online]. Available: <https://www.fao.org/hunger/es/#:~:text=Una>
  - [11] Fondo de las Naciones Unidas para la Infancia UNICEF, “Desnutrición Infantil,” México. [Online]. Available: <https://www.unicef.org/mexico/desnutricion-infantil>
  - [12] Organización Mundial de la Salud, “Nutrición,” 2024. [Online]. Available: <https://www.who.int/es/health-topics/nutrition>
  - [13] Organización Mundial de la Salud, “Malnutrición,” 2024. [Online]. Available: <https://www.who.int/es/news-room/fact-sheets/detail/malnutrition>
  - [14] Organización de las Naciones Unidas para la Alimentación y la Agricultura, Fondo Internacional de Desarrollo Agrícola, Fondo de las Naciones Unidas para la Infancia, Programa Mundial de Alimentos, and Organización Panamericana de la Salud, *Panorama Regional De Seguridad Alimentaria Y Nutricional*. 2023. [Online]. Available: <https://iris.paho.org/handle/10665.2/57048>
  - [15] H. F. Martínez, “Desnutrición a nivel municipal en Colombia,” *República Colomb. Nac. Planeación-Dirección Estud. Económicos-Archivos Econ.*, pp. 1–34, 2009, [Online]. Available: <https://www.dnp.gov.co/LinkClick.aspx?fileticket=cbZhW6NhfZE%3D&tabid=897>
  - [16] G. Hoyos, Y. Beatriz, A. Dra, M. Colmenares, and P. Clementina, “Determinantes socioeconómicos de la desnutrición crónica infantil en el Perú, año 2021,” 2023.
  - [17] A. Vilca Mamani, E. B. García Castro, L. Lipa Tudela, E. Calancho Mamani, and R. M. Cruz Huiza, “Impacto de los programas sociales alimentarios sobre la desnutrición infantil en la región de Puno,” *Comuni@cción Rev. Investig. en Comun. y Desarro.*, vol. 14, no. 3, pp. 220–234, 2023, doi: 10.33595/2226-1478.14.3.871.
  - [18] F. DE Mecánica, P. Por, and J. Diego Cruz Freire Darwin Vinicio Chimbo Chimbo, “Factores asociados en la desnutrición crónica en niños menores de 5 años en el Hospital Alfonso Villagómez, mediante la comparación de modelo de regresión y análisis de correspondencia en el período 2020-2022,” 2015.
  - [19] C. A. Minaya Gutierrez and A. Sánchez, “Crecimiento económico y desnutrición crónica infantil regional en el Perú: Una aplicación del modelo de probabilidad en el período 2000-2016,” *An. Científicos*, vol. 79, no. 2, p. 249, 2018, doi: 10.21704/ac.v79i2.1218.
  - [20] B. Muñoz Díaz *et al.*, “Nutritional status and associated factors in ambulatory elderly patients,” *Aten. Primaria*, vol. 52, no. 4, pp. 240–249, 2020, doi: 10.1016/j.aprim.2018.12.002.
  - [21] P. Flores and G. Congacha, “Factores Asociados a La Desnutrición Crónica Infantil En Ecuador. Estudio Basado En Modelos De Regresión Y Árboles De Clasificación,” *Perfiles*, vol. 1, no. 26, pp. 21–33, 2021, doi: 10.47187/perf.v1i26.132.
  - [22] Norabuena Figueroa Roger Pedro, *Random forest de los factores asociados a la desnutrición crónica en niños peruanos menores de 5 años 2020*. 2023. [Online]. Available: [https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/19783/Paucar\\_gb.pdf?sequence=1&isAllowed=y](https://cybertesis.unmsm.edu.pe/bitstream/handle/20.500.12672/19783/Paucar_gb.pdf?sequence=1&isAllowed=y)
  - [23] Secretaría Seccional de Salud de Antioquia, “Servicio de Información y Estadística,” Gobernación de Antioquia. [Online]. Available: <https://dssa.gov.co/inicio-estadisticas>
  - [24] Secretaría Seccional de Salud de Antioquia, “Agenda Antioquia 2040,” Gobernación de Antioquia. [Online]. Available: <https://agendaantioquia.com/biblioteca/>
  - [25] Secretaría Seccional de Salud de Antioquia, “Estadística Vitales,” Gobernación de Antioquia. [Online]. Available: <https://dssa.gov.co/estadisticas-vitales>
  - [26] J. D. Hernández-Lalinde *et al.*, “Sobre el uso adecuado del coeficiente de correlación de Pearson,” *Arch. Venez. Farmacol. y Ter.*, vol. 37, no. 5, pp. 586–601, 2019, [Online]. Available: [http://saber.ucv.ve/ojs/index.php/rev\\_aavft/article/view/16165](http://saber.ucv.ve/ojs/index.php/rev_aavft/article/view/16165)
  - [27] scikit-learn developers, “Scikit Learn, KNNImputer (k-Nearest Neighbors),” scikit-learn.org. [Online]. Available: <https://scikit-learn.org/stable/modules/generated/sklearn.impute.KNNImputer.html>
  - [28] N. Emiliano García D’Urso, A. Fuster Guilló, and J. Azorín López, “Análisis de la evolución morfológica del cuerpo y los datos clínicos de pacientes en tratamiento dietético nutricional,” 2021, [Online]. Available: [https://rua.ua.es/dspace/bitstream/10045/118325/1/Analisis\\_de\\_la\\_evolucion\\_morfologica\\_del\\_cuerpo\\_Garcia\\_DUrso\\_Nahuel\\_Emiliano.pdf](https://rua.ua.es/dspace/bitstream/10045/118325/1/Analisis_de_la_evolucion_morfologica_del_cuerpo_Garcia_DUrso_Nahuel_Emiliano.pdf)
  - [29] F. I. Araneda, Rodrigo Ignacio Navarro. Dra, Rosa L., “Algoritmo de aprendizaje automático para el estudio de la asociación entre enfermedad cardiovascular y depresión,” pp. 356–363, 2023, [Online]. Available: <http://repositorio.udel.cl/jspui/bitstream/11594/11348/1/Navarro>

- [30] Pandas.pydata.org, “Pandas.qcut,” OVHcloud. [Online]. Available: <https://pandas.pydata.org/docs/reference/api/pandas.qcut.html>
- [31] M. Ali, “PyCaret, low-code machine learning,” pycaret.gitbook.io. [Online]. Available: <https://pycaret.org/>