

---

# Predicción de la inseguridad alimentaria leve en Bolivia a partir de características sociodemográficas y del hogar

Minería de Datos II  
01(1):1-17  
©Josias Sirpa 2025  
SAGE

Josias Milan Sirpa Pinto\*<sup>1</sup>

## Abstract

Este estudio aplica técnicas de clasificación supervisada para predecir la inseguridad alimentaria leve en hogares bolivianos, utilizando datos de la Encuesta de Hogares 2023. Se construyó una variable binaria que identifica la presencia de inseguridad alimentaria leve. A partir de variables sociodemográficas y de condiciones de vivienda del jefe/a de hogar, se entrenaron y compararon distintos modelos: logit, probit, Naive Bayes, KNN, árboles de decisión (CART y C5.0) y redes neuronales. Los resultados muestran que el modelo logit obtuvo el mejor desempeño en términos de exactitud, con un 63.9%, seguido por redes neuronales (63.0%). Estos modelos permiten identificar perfiles de riesgo y aportar evidencia para el diseño de políticas públicas orientadas a combatir la inseguridad alimentaria en el país. .

## Keywords

Inseguridad Alimentaria; Bolivia; Predicción; Logit;

---

<sup>1</sup> Universidad Católica Boliviana San Pablo sede "La Paz"

## Corresponding author:

Josias Sirpa, Carrera de Economía e Inteligencia de Negocios, UCB, La Paz, Bolivia

Email: [josias.sirpa@ucb.edu.bo](mailto:josias.sirpa@ucb.edu.bo)

## Introducción

La inseguridad alimentaria es una problemática persistente que afecta a millones de personas en todo el mundo, especialmente en países en desarrollo. En Bolivia, a pesar de la ampliación de la cobertura de programas sociales, un número importante de hogares continúa enfrentando dificultades para acceder de forma regular y suficiente a alimentos nutritivos y adecuados. Este fenómeno, además de constituir una vulneración de derechos fundamentales, tiene implicancias directas en la salud, el desarrollo infantil, la productividad y el bienestar general de la población.

Ante esta situación, resulta necesario contar con herramientas analíticas que permitan identificar de manera temprana los hogares en riesgo de sufrir inseguridad alimentaria, que muchas veces pasa desapercibida en los programas de asistencia. El uso de métodos de minería de datos y aprendizaje automático brinda una oportunidad valiosa para construir modelos predictivos que, a partir de características observables de los hogares y sus miembros (específicamente jefes/as de hogar), permitan anticipar escenarios de vulnerabilidad alimentaria.

## Objetivos

### *Objetivo General*

Desarrollar un modelo de clasificación que permita predecir la presencia de inseguridad alimentaria en hogares bolivianos, a partir de variables sociodemográficas de los jefes/as de hogar y características de la vivienda, utilizando la Encuesta de Hogares 2023 como base de análisis.

### *Objetivos Específicos*

- Identificar y seleccionar las variables independientes relevantes que influyen en la presencia de inseguridad alimentaria en los hogares.
- Construir y entrenar distintos modelos de clasificación supervisada (logit, probit, Naive Bayes, KNN, CART, C5.0 y redes neuronales) para evaluar su desempeño predictivo.
- Comparar el desempeño de los modelos mediante métricas como la exactitud (accuracy) y la matriz de confusión, seleccionando el modelo con mejor capacidad de clasificación.
- Diseñar una función que permita aplicar el modelo final a nuevos casos, facilitando la predicción de inseguridad alimentaria leve con base en las características ingresadas.

## Motivación

El presente trabajo surge como una propuesta para aplicar técnicas de clasificación al análisis de la inseguridad alimentaria en Bolivia, con base en la Encuesta de Hogares 2023. La motivación principal es desarrollar un modelo que no solo clasifique con precisión, sino que también sea interpretable y aplicable en contextos de política pública.

Si los organismos estatales, ONG o instancias locales pudieran contar con modelos predictivos eficientes, sería posible focalizar con mayor efectividad los esfuerzos de prevención y mitigación de la inseguridad alimentaria, optimizando los recursos disponibles.

## Marco Teórico

La seguridad alimentaria existe cuando todas las personas, en todo momento, tienen acceso físico, social y económico a alimentos suficientes, inocuos y nutritivos, que satisfacen sus necesidades alimentarias y preferencias, para llevar una vida activa y saludable (FAO (2011)). Este concepto, ampliamente aceptado en el ámbito internacional, se basa en cuatro dimensiones fundamentales que permiten su evaluación integral:

- **Disponibilidad:** La disponibilidad física se refiere a la oferta de alimentos, tanto a nivel nacional como local. Esta dimensión depende de factores como la producción agrícola interna, los niveles de reservas, las importaciones y exportaciones, y el comercio neto. Es la dimensión más asociada a la oferta de alimentos y a la capacidad del sistema alimentario de abastecer a la población.
- **Acceso:** El acceso económico y físico a los alimentos implica que las personas no solo deben poder encontrar alimentos en el mercado o en sus entornos, sino también tener los medios para adquirirlos. Esto incluye el ingreso del hogar, la estabilidad de precios, la infraestructura física (caminos, transporte) y los mecanismos sociales que facilitan el acceso equitativo. Una adecuada disponibilidad no garantiza la seguridad alimentaria si las personas no tienen capacidad adquisitiva.
- **Utilización:** La utilización biológica hace referencia al uso adecuado de los alimentos por parte del organismo. Esta dimensión abarca aspectos como la preparación, diversidad y calidad de la dieta, las prácticas de higiene y salud, y la absorción eficiente de nutrientes. La seguridad alimentaria implica, por tanto, que los alimentos no solo estén disponibles y sean accesibles, sino también que su consumo conlleve beneficios nutricionales efectivos para los individuos.
- **Estabilidad:** La estabilidad implica que las otras tres dimensiones (disponibilidad, acceso y utilización) se mantengan a lo largo del tiempo. Una persona puede tener acceso adecuado a los alimentos en un momento dado, pero aún así estar en situación de inseguridad alimentaria si ese acceso no está garantizado de manera sostenida. Factores como el desempleo, los conflictos, las crisis económicas, las variaciones estacionales o los desastres naturales pueden alterar esta estabilidad.

Estas dimensiones no deben analizarse de forma aislada, ya que solo una combinación armónica entre ellas permite alcanzar una seguridad alimentaria sostenible. Este enfoque ha sido clave para el diseño de políticas públicas y estrategias integrales de combate al hambre, especialmente en contextos vulnerables como el boliviano.

## *Medición basada en experiencia del hogar*

La **Organización de las Naciones Unidas para la Alimentación y la Agricultura (2024)** utiliza escalas validadas como la ELCSA (Escala Latinoamericana y Caribeña de Seguridad Alimentaria) y el FIES (Food Insecurity Experience Scale). Estas herramientas capturan directamente la vivencia de los hogares (preocupación, reducción de calidad, cantidad, hambre), distinguiendo niveles de severidad: leve, moderado y severo. En Bolivia la Encuesta Hogares de **Instituto Nacional de Estadística (2023)** utiliza la escala FIES, por lo cual a continuación desarrollaremos la misma.

### *Food Insecurity Experience Scale (FIES)*

La Escala de Experiencia de Inseguridad Alimentaria (FIES) es una herramienta desarrollada por la FAO en el marco del proyecto Voices of the Hungry, con el objetivo de medir la inseguridad alimentaria desde la vivencia directa de las personas. A diferencia de otros enfoques tradicionales, como la Prevalencia de la Subalimentación, que evalúan el fenómeno de forma indirecta a partir de la disponibilidad de alimentos o los niveles de ingreso, la FIES captura experiencias concretas vinculadas a dificultades en el acceso a alimentos por falta de recursos. Este enfoque centrado en la experiencia permite aproximarse al componente subjetivo y cotidiano de la inseguridad alimentaria, reflejando con mayor precisión los impactos reales sobre los hogares.

La escala consiste en un módulo de ocho preguntas que indagan sobre experiencias relacionadas con preocupaciones por el acceso a alimentos, reducción en la calidad o cantidad de la dieta, e incluso episodios de hambre. Estas preguntas tienen respuestas dicotómicas (sí/no) y se administran con un período de 12 meses, sin embargo, esto puede variar dependiendo del diseño de la encuesta. En Bolivia, la Encuesta de Hogares 2023 aplica la versión para hogares, en la cual una persona responde en representación del grupo familiar.

Las ocho preguntas son las siguientes:

1. Se haya preocupado por no tener suficientes alimentos para comer
2. No haya podido comer alimentos sanos o nutritivos
3. Haya comido poca variedad de alimentos
4. Haya tenido que saltarse una comida
5. Haya comido menos de lo que pensaba que debía comer
6. Su hogar se haya quedado sin alimentos
7. Haya sentido hambre pero no comió
8. Haya dejado de comer durante todo un día

La FIES se construye como una escala estadística basada en la teoría de respuesta al ítem (TRI), empleando modelos probabilísticos que permiten clasificar a los hogares en distintos niveles de severidad de la inseguridad alimentaria (leve, moderada o severa). Esta metodología, similar a la utilizada en pruebas psicométricas, asegura la comparabilidad de los resultados a nivel internacional, lo que convierte a la FIES en uno de los dos indicadores oficiales para monitorear el Objetivo de Desarrollo Sostenible 2.1

**Organización de las Naciones Unidas (2025)**, que busca poner fin al hambre y garantizar el acceso a una alimentación adecuada.

Entre sus principales fortalezas se destaca su capacidad para producir estimaciones comparables entre países y subpoblaciones, así como su aplicabilidad en diversos contextos socioculturales. Además, su integración en encuestas nacionales permite identificar grupos vulnerables, monitorear políticas públicas y explorar factores de riesgo asociados a la inseguridad alimentaria. No obstante, su análisis requiere cierto nivel de especialización técnica y el uso de software específico, lo cual puede representar un desafío para algunas instituciones. La FAO, sin embargo, ofrece herramientas estadísticas, módulos en varios idiomas y asistencia técnica para facilitar su implementación.

### *Contexto boliviano*

Bolivia enfrenta una situación compleja en materia de seguridad alimentaria, determinada por factores estructurales como la pobreza, la desigualdad y la alta vulnerabilidad territorial. Según **Mallea and Maria (2014)**, el 37% de la población boliviana vive en extrema pobreza, una condición que repercute directamente en el acceso y la calidad de los alimentos. Esta problemática se refleja también en indicadores preocupantes como la desnutrición crónica infantil, que alcanza al 32% de la población menor de cinco años, y en una alarmante incidencia de anemia que afecta a ocho de cada diez niños entre los 6 y 23 meses de edad.

El índice de desarrollo humano del país se encuentra en un nivel medio-bajo (0,729), y aproximadamente el 53% de los municipios presentan un grado alto o muy alto de vulnerabilidad a la inseguridad alimentaria. Las regiones rurales, especialmente el altiplano, evidencian las peores condiciones de ingesta calórica y proteica, exacerbadas por factores como el aislamiento geográfico, el bajo acceso a servicios básicos y la escasa capacidad de respuesta ante eventos climáticos adversos (**Mallea and Maria (2014)**).

La población indígena, que representa alrededor del 65% del total nacional, es particularmente afectada por estas condiciones, debido a marcadas desigualdades en el acceso a recursos, tecnologías y oportunidades. Además, la migración campo-ciudad ha generado nuevas formas de inseguridad alimentaria en áreas periurbanas, donde se incrementa la demanda de servicios sin que existan mecanismos de protección adecuados.

En respuesta a esta realidad, el Estado boliviano ha incorporado instrumentos como la FIES en encuestas oficiales, lo que permite contar con una base empírica más sólida para evaluar y monitorear la inseguridad alimentaria. Sin embargo, como advierte la literatura, la información nacional sigue siendo escasa, dispersa y metodológicamente limitada, lo que dificulta una planificación eficaz (**Mallea and Maria (2014)**).

El presente estudio cobra especial relevancia al proponer una herramienta predictiva basada en minería de datos, que permite anticipar situaciones de inseguridad alimentaria a partir de características observables del hogar. Esto no solo complementa los esfuerzos existentes, sino que también contribuye al diseño de políticas públicas más focalizadas,

orientadas a reducir la vulnerabilidad y mejorar la calidad de vida de las poblaciones más afectadas.

## Descripción del dataset

La presente investigación utiliza los microdatos de la Encuesta de Hogares 2023 (EH2023), elaborada por el Instituto Nacional de Estadística de Bolivia (INE). Esta encuesta tiene representatividad nacional, departamental y por área geográfica (urbana y rural), y recoge información detallada sobre las condiciones sociodemográficas, económicas, educativas y habitacionales de los hogares bolivianos.

En particular, el módulo de seguridad alimentaria de la EH2023 incluye preguntas asociadas a la Escala de Experiencia de Inseguridad Alimentaria (FIES), desarrollada por la FAO. La escala evalúa la inseguridad alimentaria en tres niveles de severidad: leve, moderada y severa. Para este estudio, se optó por analizar la variable s07a.01, que corresponde a la primera pregunta del módulo FIES: “¿Usted o algún miembro de su hogar estuvo preocupado porque no tenían suficientes alimentos?”. Esta pregunta mide la inseguridad alimentaria leve, centrada en la dimensión perceptiva de la preocupación, y fue seleccionada por su alta cobertura y baja incidencia de valores perdidos, a diferencia de las preguntas asociadas a las categorías moderada y severa, que presentaban menor variabilidad y mayor sesgo en su distribución.

### INSEGURIDAD ALIMENTARIA SEGÚN LA FIES. ¿QUÉ SIGNIFICA?



**Figura 1.** Escala FIES.

En la figura 1 podemos ver como la FIES clasifica las respuestas de las ocho preguntas de la encuesta, para medir una inseguridad leve, moderada y grave.

La variable dependiente fue transformada en binaria con los siguientes criterios:

- Se codificó como 1 si el hogar respondió afirmativamente (“sí”) a la pregunta de preocupación por la disponibilidad de alimentos.
- Se codificó como 0 si la respuesta fue negativa (“no”) o “no sabe/no responde”. Estas últimas se agruparon con las respuestas negativas por su baja frecuencia y para mantener una dicotomía estadísticamente robusta.

Se seleccionaron variables explicativas provenientes de distintos módulos del cuestionario de hogares:

- Características del jefe/a de hogar (eh23p): sexo, edad, autoidentificación indígena, es decir, si pertenece a algún pueblo indígena (p\_indig) y años de estudio (aestudio), finalmente la base de datos se convirtió en eh23p.f
- Condiciones de la vivienda (eh23v): número total de personas en el hogar (totper), ingreso total del hogar (yhog), tipo de vivienda, es decir, si es propia, rentada, etc. (s06a\_02), material del piso (s06a\_06), disponibilidad de un cuarto de cocina (s06a\_14) y acceso a internet (s06a\_19), la base de datos final es eh23v.f
- Disponibilidad de bienes durables (eh23e): tenencia de refrigerador (refrigerador), transformada a variable binaria (1 = posee, 0 = no posee), inicialmente también se consideró tenencia de cocina en una variable binaria, sin embargo, al realizar los modelos logit demostró no ser significativo por lo cual se eliminó. Además se alargó la base de datos, ya que originalmente estaba clasificada por ítems (del 1 al 17), se filtró refrigerador y cocina (ítems 2 y 4) y se transformó a variables binarias, la base de datos final es eh23ef.

Para asegurar la correcta integración de los módulos, se utilizó el identificador único de hogar (folio). La base de datos fue depurada para eliminar observaciones con valores perdidos (sin embargo no existían). Posteriormente, se realizó una partición aleatoria de la muestra en conjuntos de entrenamiento (70%) y prueba (30%), manteniendo la distribución de la variable dependiente para garantizar un entrenamiento balanceado.

En resumen, se utilizó las siguientes variables independientes para el ejercicio (no se incluye variable dependiente s07a\_01):

**Table 1.** Variables utilizadas en el análisis, clasificadas según su origen

Variable	Pertenece.a	Detalles
folio	Vivienda	Identificador único del hogar
sexo	Jefe/a del hogar	Sexo del jefe/a del hogar
edad	Jefe/a del hogar	Edad del jefe/a del hogar (años cumplidos)
p_indig	Jefe/a del hogar	Pertenencia a algún pueblo indígena
aestudio	Jefe/a del hogar	Años de estudio del jefe/a del hogar
yhog	Vivienda	Ingreso mensual del hogar en bolivianos

Variable	Pertenece.a	Detalles
totper	Vivienda	Total de personas en el hogar
s06a_02	Vivienda	Tipo de tenencia de la vivienda
s06a_06	Vivienda	Material predominante del piso
s06a_14	Vivienda	Disponibilidad de cocina
s06a_19	Vivienda	Acceso a internet
refrigerador	Vivienda	Disponibilidad de refrigerador

En total, la base final de análisis incluyó 12.815 hogares, con un conjunto de predictores representativos de las condiciones estructurales, materiales y sociodemográficas que, desde la literatura especializada, se asocian a la probabilidad de experimentar inseguridad alimentaria leve.

## Metodología

La presente investigación aplica métodos de clasificación supervisada para predecir la presencia de inseguridad alimentaria leve en hogares bolivianos. Se parte de un enfoque de analítica predictiva, en el cual se busca estimar la probabilidad de que un hogar experimente inseguridad alimentaria leve a partir de un conjunto de variables explicativas observadas. Para ello, se entrenaron y compararon varios modelos de clasificación, utilizando una partición aleatoria de los datos en subconjuntos de entrenamiento (70%) y prueba (30%).

En términos formales, se busca estimar una función  $f$  tal que:

$$y = f(x)$$

donde  $y$  es una variable binaria con valores de 0 para “NO” y 1 para “SI”, y representa la variable dependiente (inseguridad alimentaria leve) y  $x$  es el vector de predictores. El conjunto de modelos utilizados incluye:

- Regresión logística (logit)
- Regresión probit
- Naive Bayes
- K-nearest neighbors (KNN)
- Árbol de clasificación CART
- Árbol C5.0
- Redes neuronales artificiales (ANN)

### Regresión logística y probit

Estos modelos estiman la probabilidad de que un hogar pertenezca a la categoría con inseguridad alimentaria:

$$P(x) = \mathbb{P}(Y = 1|X = x) = \mathbb{E}[Y|X = x]$$



**Modelo Probit:**

$$P(x) = \Phi(x'\beta)$$

**Modelo Logit:**

$$P(x) = \frac{1}{1 + \exp(-x'\beta)}$$

donde  $\phi$  es la función de distribución acumulada normal estándar. Ambos modelos fueron ajustados mediante `glm()` en R, utilizando la función `step()` para seleccionar variables significativas.

**Naive Bayes**

El clasificador Naive Bayes se basa en el teorema de Bayes, con el supuesto fuerte de independencia condicional entre predictores:

$$P(Y|X) = \frac{P(Y) \cdot P(X|Y)}{P(X)} \approx P(Y) \cdot P(X|Y)$$

Con independencia condicional:

$$P(Y|X) \propto P(Y) \cdot \prod_{i=1}^p P(X_i|Y)$$

Se utilizó la librería `naivebayes`, evaluando el rendimiento mediante la matriz de confusión sobre la base de prueba.

**K-Nearest Neighbors (KNN)**

El método KNN asigna una observación nueva a la clase más frecuente entre sus  $k$  vecinos más cercanos en el espacio de covariables. La distancia empleada fue la euclídeana:

Distancia Euclídeana:

$$d(x, x_i) = \sqrt{\sum_{j=1}^p (x_j - x_{ij})^2}$$

El valor de  $k$  se eligió como la raíz cuadrada del número de observaciones en la base de entrenamiento. Se normalizaron las variables predictoras para evitar sesgos por escala.

**Árboles de clasificación: CART y C5.0**

Los árboles de decisión construyen reglas de partición mediante medidas de impureza como la entropía o el índice Gini:

Entropía:

$$\text{Entropía}(t) = - \sum_{i=1}^k p_i \log_2(p_i)$$

Índice Gini:

$$\text{Gini}(t) = 1 - \sum_{i=1}^k p_i^2$$

El algoritmo CART se implementó mediante la librería rpart, mientras que C5.0 se ajustó con C50. Ambos modelos generan una estructura de árbol que facilita la interpretación de las reglas de decisión.

### *Redes neuronales artificiales (ANN)*

Las redes neuronales permiten modelar relaciones no lineales entre las variables explicativas y la respuesta. Una ANN se compone de capas de nodos (entrada, ocultas y salida) interconectadas por pesos ajustables. La salida de una red simple con una capa oculta se puede representar como:

Red neuronal con una capa oculta:

$$\hat{y}(x) = f \left( \sum_{j=1}^h w_j \cdot \sigma \left( \sum_{i=1}^p w_{ij} x_i + b_j \right) + b_o \right)$$

Función de activación sigmoide:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

donde  $\sigma$  es una función de activación (e.g. sigmoide), y los pesos  $w$  se ajustan mediante backpropagation y descenso del gradiente. Se empleó la librería nnet con una arquitectura de una capa oculta y hasta 20 nodos.

### *Evaluación del rendimiento*

El desempeño de cada modelo se evaluó mediante la matriz de confusión y la métrica de exactitud (accuracy) sobre la base de prueba. También se examinó la sensibilidad y especificidad para interpretar la capacidad de clasificación en función de verdaderos positivos y negativos.

### **Resultados, análisis y predicción**

Al inicio del ejercicio se consideraron un total de 17 variables (sin considerar folio y la variable dependiente) que se consideraban significativas, sin embargo, al momento de realizar el logit y gracias a la función step(), se termino descartando las siguientes variables:

- s06a\_12: Existencia de energia electrica en la vivienda.
- leer\_esc: Si el jefe/a de hogar sabia leer y escribir.

- s06a\_08a: Dias a la semana con agua en la vivienda.
- cocina: Variable binaria de tenencia de cocina en la vivienda
- s06a\_07: Proveniencia del agua de la vivienda (pozo, alcantarilla, etc.)
- s06a\_09: Tipo de baño en la vivienda

A continuación se muestran los resultados de los diferentes métodos y al final un resumen

### *Logit*

**Table 2.** Resultados del modelo Logit (Matriz de confusión)

Métrica	Valor
Accuracy	0.6398
Kappa	0.2736
Sensitivity	0.6437
Specificity	0.6344
Valor predictivo positivo (PPV)	0.7070
Valor predictivo negativo (NPV)	0.5651
Balanced Accuracy	0.6391
No Information Rate	0.5782
P-Value (Acc > NIR)	0.0000
Mcnemar's Test P-Value	0.0000

### *Probit*

**Table 3.** Resultados del modelo Probit (Matriz de confusión)

Métrica	Valor
Accuracy	0.6380
Kappa	0.2699
Sensibilidad	0.6421
Especificidad	0.6323
Valor predictivo positivo (PPV)	0.7055
Valor predictivo negativo (NPV)	0.5629
Balanced Accuracy	0.6372
No Information Rate	0.5784
P-Value (Acc > NIR)	0.0000
Mcnemar's Test P-Value	0.0000

### *Naive Bayes*

**Table 4.** Resultados del modelo Naive Bayes (Matriz de confusión)

Métrica	Valor
Accuracy	0.5961
Kappa	0.1735
Sensibilidad	0.5858
Especificidad	0.6218
Valor predictivo positivo (PPV)	0.7945
Valor predictivo negativo (NPV)	0.3756
Balanced Accuracy	0.6038
No Information Rate	0.7139
P-Value (Acc > NIR)	1.0000
Mcnemar's Test P-Value	0.0000

### KNN

**Table 5.** Resultados del modelo K-Nearest Neighbors (Matriz de confusión)

Métrica	Valor
Accuracy	0.6114
Kappa	0.2125
Sensibilidad	0.6102
Especificidad	0.6135
Valor predictivo positivo (PPV)	0.7248
Valor predictivo negativo (NPV)	0.4854
Balanced Accuracy	0.6118
No Information Rate	0.6252
P-Value (Acc > NIR)	0.9624
Mcnemar's Test P-Value	0.0000

### Arbol de decisión

**Table 6.** Comparación de desempeño: Árboles de decisión (CART vs. C5.0)

Métrica	CART	C5.0
Accuracy	0.6120	0.6161
Kappa	0.2208	0.2243
Sensibilidad	0.6286	0.6115
Especificidad	0.5926	0.6192
Valor predictivo positivo (PPV)	0.6423	0.5195
Valor predictivo negativo (NPV)	0.5783	0.7031
Balanced Accuracy	0.6106	0.6154
No Information Rate	0.5378	0.5977

Métrica	CART	C5.0
P-Value (Acc > NIR)	0.0000	0.0101
McNemar's Test P-Value	0.2656	0.0000

## Redes Neuronales

**Table 7.** Resultados del modelo Red Neuronal (nnet)

Métrica	Valor
Accuracy	0.6299
Kappa	0.2543
Sensibilidad	0.6372
Especificidad	0.6203
Valor predictivo positivo (PPV)	0.6897
Valor predictivo negativo (NPV)	0.5634
Balanced Accuracy	0.6287
No Information Rate	0.5698
P-Value (Acc > NIR)	0.0000
McNemar's Test P-Value	0.0000

**Table 8.** Comparación de Accuracy entre modelos de clasificación

Modelo	Accuracy
Logit	0.6398
Probit	0.6380
Naive Bayes	0.5961
KNN	0.6114
CART	0.6120
C5.0	0.6161
Red Neuronal	0.6299

Dado los resultados que se pueden ver en la tabla 8, se eligió el modelo logit para realizar la función predictiva, se elaboró una función `predecir_inseguridad()` para el modelo predictivo:

```
predecir_inseguridad <- function(sexo, edad, p_indig,
                                aestudio, yhog, totper,
                                s06a_02, s06a_06,
                                s06a_14, s06a_19, refrigerador) {
  nuevo_hogar <- data.frame(
    sexo = factor(sexo,
                  levels = levels(bdtrain$sexo)),
```

```

edad = edad,
p_indig = p_indig,
aestudio = aestudio,
yhog = yhog,
totper = totper,
s06a_02 = factor(s06a_02,
                  levels = levels(bdtrain$s06a_02)),
s06a_06 = factor(s06a_06,
                  levels = levels(bdtrain$s06a_06)),
s06a_14 = s06a_14,
s06a_19 = s06a_19,
refrigerador = refrigerador
)
if (any(is.na(nuevo_hogar))) {
  print("Error: alguno de los valores ingresados
        no coincide con los niveles esperados.")
  return(nuevo_hogar)
}
prob <- predict(m1, newdata = nuevo_hogar,
                type = "response")
resultado <- ifelse(prob > 0.5, "Sí sufre inseguridad
                             alimentaria leve",
                    "NO sufre inseguridad alimentaria leve")
list(probabilidad = round(prob, 3), resultado = resultado)
}

```

Para poder utilizar la función y predecir si existe inseguridad alimentaria en un hogar, dadas 11 variables, de las cuales 4 son del jefe/a de hogar y el resto son características de la vivienda, se debe llenar el formulario de la siguiente forma:

```

predecir_inseguridad(
  sexo = "1. Hombre",
  edad = 40,
  p_indig = 0,
  aestudio = 8,
  yhog = 2000,
  totper = 4,
  s06a_02 = "3. ¿Alquilada?",
  s06a_06 = "5. CEMENTO",
  s06a_14 = 1,
  s06a_19 = 1,
  refrigerador = 1
)

```

en la que las variables categoricas n6minales deben ser llenadas con alguna de las siguientes opciones:

- sexo: 1. Hombre, 2. Mujer
- s06a\_02: 1. ¿Propia y totalmente pagada?, 2. ¿Propia y la est1n pagando?, 3. ¿Alquilada? 4. ¿En contrato Mixto (alquiler y anticretico)?, 5. ¿En contrato anticretico?, 6. ¿Cedida por servicios?, 7. ¿Prestada por parientes o amigos? u 8. ¿Otra?(Especifique)
- s06a\_06: 1. TIERRA, 2. TABL3N DE MADERA, 3. MACHI-HEMBRE/PARQUET, 4. PISO FLOTANTE, 5. CEMENTO, 6. MOSAICO/BALDOSAS/CER1MICA, 7. LADRILLO u 8. OTRO (Especifique)

Una vez llenado correctamente y ejecutado la funci3n, se obtendra un resultado como el siguiente:

```
# Ejemplo de salida de la funci3n predecir_inseguridad
$probabilidad
      1
0.633

$resultado
                                     1
"SÍ sufre inseguridad alimentaria leve"
```

Podemos apreciar que el modelo esta funcionando correctamente, las características de la vivienda de ejemplo son un hogar con 4 integrantes y un ingreso promedio de Bs. 2000, lo que puede ya darnos un gran indicio de que este hogar sufre inseguridad alimentaria por falta de recursos. Adem1s podemos ver las otras variables como por ejemplo que la vivienda es alquilada, por lo tanto gran parte del ingreso del hogar se destina a alquiler, restando recursos disponibles para una buena alimentaci3n.

## Conclusiones y recomendaciones

Los resultados de este estudio muestran que es posible predecir la probabilidad de que un hogar boliviano experimente inseguridad alimentaria leve a partir de características sociodemográficas y materiales observables, utilizando modelos de clasificaci3n supervisada. El modelo logit, que obtuvo el mejor desempe1o entre las t1cnicas comparadas, alcanz3 una precisi3n de aproximadamente 63.9%, lo cual sugiere una capacidad predictiva moderada.

Si bien este nivel de exactitud no es el mejor (63.9%), especialmente considerando que se trata de un problema socialmente sensible, es importante destacar que los modelos desarrollados no pretenden reemplazar mecanismos formales de diagn3stico, sino ofrecer una herramienta complementaria para focalizar intervenciones. La inseguridad alimentaria leve, al estar basada en la experiencia subjetiva de preocupaci3n por la

alimentación, puede estar influida por factores no directamente observables en las encuestas. Esto podría explicar parte del margen de error del modelo.

A pesar de esta limitación, el enfoque predictivo basado en minería de datos tiene alto valor aplicado. Puede servir como un insumo útil para el diseño de políticas públicas focalizadas, especialmente en programas de transferencias, subsidios alimentarios o asistencia nutricional.

Se recomienda ampliar el conjunto de variables predictoras incluyendo información sobre empleo, gasto alimentario o acceso a redes de apoyo, aplicar otras técnicas como random forest o boosting, que podrían mejorar el rendimiento predictivo, utilizar métricas adicionales como el AUC-ROC para evaluar el balance entre sensibilidad y especificidad.

En conclusión, el presente trabajo demuestra que, incluso con exactitudes intermedias, los modelos de clasificación pueden aportar evidencia valiosa para combatir la inseguridad alimentaria desde un enfoque preventivo y focalizado, orientando recursos limitados hacia los hogares con mayor riesgo.



## References

- FAO (2011) Una introducción a los conceptos básicos de la seguridad alimentaria. URL <https://www.fao.org/4/a1936s/a1936s00.pdf>. Accedido el 10 de junio de 2025.
- Instituto Nacional de Estadística (2023) Encuesta de hogares 2023. <https://www.ine.gob.bo/index.php/estadisticas-sociales/vivienda-y-servicios-basicos/encuestas-de-hogares-vivienda/>. La Paz, Bolivia.
- Mallea T and Maria I (2014) Situación actual y prioridades básicas de la seguridad alimentaria nutricional en bolivia. *CienciAgro* 4(1).
- Organización de las Naciones Unidas (2025) Objetivos de desarrollo sostenible. <https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>. Consultado el 10 de junio de 2025.
- Organización de las Naciones Unidas para la Alimentación y la Agricultura (2024) Sitio oficial de la fao. <https://www.fao.org/home/es>. Consultado el 10 de junio de 2025.