



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

Facultad de  
**MEDICINA**

Inteligencia artificial en salud pública: desarrollo de un modelo predictivo para estimar el porcentaje de anemia en comunidades peruanas

Artificial intelligence in public health: development of a predictive model to estimate the percentage of anemia in peruvian communities

PROYECTO DE INVESTIGACIÓN PARA OPTAR POR EL  
TÍTULO DE SEGUNDA ESPECIALIDAD PROFESIONAL EN  
ADMINISTRACIÓN Y GESTIÓN EN SALUD

AUTOR

DIEGO DANILO GAYOSO LIVIAC

ASESOR

FERNANDO ENRIQUE DURAND CONCHA

LIMA – PERÚ

2025


# RESULTADO DEL INFORME DE SIMILITUD

ev.turnitin.com/app/carta/es/?u=1151562268&o=2665214763&ro=103&lang=es&ks=1

turnitin

1 de 251: DIEGO DANILO GAYOSO LIVIAC  
Inteligencia artificial en salud pública: desarrollo de u...

Similitud 9% Marcas de alerta



UNIVERSIDAD PERUANA  
**CAYETANO HEREDIA**

Facultad de  
**MEDICINA**

Inteligencia artificial en salud pública: desarrollo de un modelo predictivo para estimar el porcentaje de anemia en comunidades peruanas

Artificial intelligence in public health: development of a predictive model to estimate the percentage of anemia in peruvian communities

**PROYECTO DE INVESTIGACIÓN PARA OPTAR POR EL TÍTULO DE SEGUNDA ESPECIALIDAD PROFESIONAL EN ADMINISTRACIÓN Y GESTIÓN EN SALUD**

**AUTOR**  
DIEGO DANILO GAYOSO LIVIAC

**ASESOR**  
FERNANDO ENRIQUE DURAND CONCHA

LIMA – PERÚ  
2025

Página 1 de 13 2879 palabras 145%

Informe estándar  
Informe en inglés no disponible [Más información](#)

**9% Similitud estándar** [Filtros](#)

Fuentes  
Mostrar las fuentes solapadas

- 1 Trabajos del estudiante   
Universidad Internacional de la Ri... 2%  
4 bloques de texto 52 palabras que coinciden
- 2 Internet   
www.coursehero.com <1%  
3 bloques de texto 25 palabras que coinciden
- 3 Internet   
cybertesis.unmsm.edu.pe <1%  
1 bloques de bloques 18 palabras que coinciden
- 4 Internet   
www.insp.mx <1%  
1 bloques de bloques 17 palabras que coinciden
- 5 Trabajos del estudiante

Mostrar escritorio

## 2. RESUMEN

**Introducción:** La anemia es un problema mundial de salud pública, que afecta especialmente a poblaciones vulnerables como niños, embarazadas y adultos mayores. En el Perú, la prevalencia de anemia en niños menores de 3 años ronda entre el 40% y 50%, con mayor porcentaje en zonas rurales, situación que ha persistido en el tiempo a pesar de los esfuerzos por reducirla. **Objetivo del estudio:** Se propone desarrollar un modelo predictivo basado en técnicas de Inteligencia Artificial o *Machine Learning* para estimar el porcentaje de anemia en comunidades peruanas, aprovechando datos abiertos de salud pública. **Diseño del estudio:** Se propone un estudio observacional retrospectivo de tipo longitudinal, utilizando datos secundarios recopilados por el Ministerio de Salud. **Población y muestra:** La población de estudio comprende las comunidades a nivel nacional (departamentos, provincias y distritos del Perú), sin enfocarse en un departamento específico. No se realizará muestreo, dado que se utilizarán todos los datos disponibles en las fuentes abiertas del MINSA sobre anemia. **Procedimientos y técnicas:** Se recopilarán y depurarán los registros de anemia y otras variables teóricamente relacionadas. Luego se entrenará un modelo de *machine learning* utilizando algoritmos supervisados (*Random Forest*) para predecir la prevalencia de anemia en cada comunidad en función de los determinantes de salud identificados. **Análisis estadístico:** Se evaluará el desempeño del modelo mediante métricas como la exactitud (*accuracy*), recuperación (*recall*), precisión (*precision*) y el área bajo la curva (AUC). Se espera que el modelo logre identificar patrones complejos entre los determinantes de la salud y la anemia.

**Palabras clave:** Inteligencia Artificial, Informática en Salud Pública, Anemia.

### **3. INTRODUCCIÓN**

La anemia es un problema de salud pública en el Perú y el mundo, especialmente en los niños, las embarazadas y los adultos mayores. A nivel nacional, cerca de un tercio de los niños peruanos presentan esta enfermedad (1). En lactantes y niños pequeños entre los 6 y 35 meses de edad, la prevalencia es aún más alta, siendo alrededor del 40% en zonas urbanas y poco más del 50% en zonas rurales (2). Esta prevalencia se ha mantenido relativamente constante a pesar de las mejoras socioeconómicas de las últimas décadas y las estrategias de salud pública implementadas, lo que sugiere que el problema de la anemia sigue siendo un desafío importante en el país.

Las efectos a largo plazo de la anemia son importantes, y existe un consenso acerca del gran impacto que esta enfermedad tiene sobre el desarrollo infantil (3). Existen estudios que señalan que los niños con anemia en primeros años de vida presentan retrasos en el desarrollo cognitivo y el desarrollo motor, tienen menor rendimiento escolar, y una mayor tasa de incidencia de problemas de conducta en las siguientes etapas de vida (4). La anemia también afecta el sistema inmunológico y la salud en general, aumentando la susceptibilidad a infecciones, afectando negativamente la calidad de vida (5–8).

En años recientes, con el auge de los modelos de lenguaje generativos, la inteligencia artificial (IA) y el aprendizaje automático se han posicionado como soluciones innovadoras en el sector salud (9). Estas tecnologías permiten desarrollar modelos predictivos capaces de predecir la aparición o el riesgo de enfermedades, pueden ayudar en la planificación e implementación de intervenciones preventivas,

y hasta pueden facilitar la interpretación de pruebas diagnósticas (10–13). A modo de ejemplo, una revisión señala que la IA ha permitido desarrollar modelos predictivos enfocados en enfermedades transmisibles, ayudando en la toma de decisiones oportunas y mejorando los resultados en la salud de la población (14). Algo que caracteriza a estos modelos es que poseen la capacidad de analizar grandes volúmenes de datos de diferentes tipos, pudiendo identificar patrones y factores de riesgo que son difíciles de identificar usando métodos tradicionales. Así, la IA en el ámbito de la salud pública se podría utilizar, por ejemplo, para identificar de manera oportuna brotes epidémicos, categorizar el riesgo de enfermedades en pacientes y optimizar el gasto o distribución de los recursos disponibles.

En América Latina, ya se han realizado estudios que aplican técnicas de aprendizaje automático para abordar problemas nutricionales y de salud pública como la anemia. Por ejemplo, un estudio reciente analizó la Encuesta Demográfica y de Salud Familiar (ENDES) 2015–2019, empleando algoritmos de *machine learning* para identificar factores asociados a anemia infantil (15). A partir del ENDES, los investigadores probaron diversos modelos de clasificación, alcanzando una sensibilidad y especificidad cercanas al 64–66 % en los modelos entrenados más exitosos.

Este tipo de análisis permite explorar la influencia de los diferentes determinantes de la salud sobre la anemia en contextos locales. Por ejemplo, en el estudio antes mencionado, el modelo pudo identificar como principales factores influyentes: la edad del niño, la altitud del hogar (m.s.n.m.), la altura materna, y la cantidad y momento de visitas prenatales (15). Estos resultados coinciden con la literatura internacional, que menciona la relación de los diferentes factores individuales,

familiares y comunitarios con la anemia infantil (16–18). De este estudio se puede concluir que los datos abiertos de salud pueden ser aprovechados mediante técnicas de análisis o de aprendizaje automático para lograr comprender y predecir problemas que pueden resultar complejos (15).

Considerando la magnitud del problema de la anemia en el Perú, la disponibilidad de datos abiertos y los avances en metodologías de inteligencia artificial, nace la siguiente pregunta de investigación: ¿ Es posible, mediante un modelo predictivo de Inteligencia Artificial entrenado con información proveniente de datos abiertos, estimar de forma precisa la prevalencia de anemia en comunidades peruanas? La respuesta a esta pregunta podría brindar un enfoque innovador para reforzar la toma de decisiones y acciones de salud pública en la lucha contra la anemia. Esta tecnología podría permitir identificar zonas de mayor riesgo que requieran intervenciones focalizadas, impactando directamente en la eficiencia del gasto de recursos en el sector público.

#### **4. OBJETIVOS**

**Objetivo general:** Desarrollar y validar un modelo predictivo basado en técnicas de inteligencia artificial que permita estimar la prevalencia de anemia infantil en las comunidades peruanas, a partir de datos abiertos de la Plataforma Nacional de Datos Abiertos del MINSA, el INEI, el RENIEC y otras fuentes públicas pertinentes, con el fin de identificar oportunamente las áreas de mayor riesgo y contribuir al diseño de intervenciones de salud pública más eficientes y efectivas.

**Objetivos específicos:**

- Identificar y describir las variables socioeconómicas, demográficas, de salud y nutrición relacionadas con la anemia infantil disponibles en las bases de datos abiertas del Perú para el periodo 2019-2023.
- Determinar la capacidad predictiva de un modelo de *machine learning* para estimar la prevalencia de anemia infantil a nivel comunitario, utilizando los datos abiertos identificados.
- Identificar las variables predictoras de mayor influencia en la estimación de la prevalencia de anemia infantil según el modelo desarrollado.

## **5. MATERIAL Y MÉTODO**

**a) Diseño del estudio:** Estudio transversal, retrospectivo y observacional, que analiza datos secundarios agregados a nivel comunitario para el periodo 2019-2023. Se trata de un estudio de análisis secundario de datos en el que se construirá y validará un modelo de inteligencia artificial. El diseño es transversal pues la estimación de la prevalencia de anemia se basará en datos de un periodo determinados. Adicionalmente, se trata un estudio retrospectivo de datos secundarios, dado que se utilizarán registros existentes recolectados previamente por el Ministerio de Salud. Es observacional porque no se intervendrá en las variables, sino que se analizarán tal como fueron reportadas, enfocándose en la construcción de un modelo predictivo.

**b) Población y ámbito de estudio:** La población de estudio comprenderá los registros agregados a nivel comunitario (distrito/provincia) correspondientes a niños menores de 3 años durante el periodo 2019-2023, obtenidos a partir de datos abiertos de la Plataforma Nacional de Datos Abiertos del MINSA, el INEI, el

RENIEC y otras fuentes públicas pertinentes. El estudio abarca toda la población descrita en el *dataset* “Anemia” de la Plataforma Nacional de Datos Abiertos (19), que recopila información de todos los registros en donde hubo una atención de anemia, tanto en zonas urbanas como rurales, en el periodo comprendido entre el 2019 al 2023. Se define como único criterio de selección el haber tenido 3 o menos años en algún momento del periodo de estudio.

**c) Muestra:** Al tratarse de un análisis de datos secundarios a escala nacional, no se realizará un muestreo probabilístico tradicional. Se empleará muestra censal de datos, es decir, se incluirán todos los registros y unidades disponibles pertinentes al problema. Se utilizarán los datos abiertos proporcionados por el MINSA disponibles en la Plataforma Nacional de Datos Abiertos mediante el *dataset* “Anemia”, específicamente los conjuntos de datos denominados “Anemia\_procedimientos”, “Anemia Atendidos”, “Anemia Tamizaje” y “Anemia Tratamientos”. Esto incluye los registros de atenciones por anemia en establecimientos de salud, los resultados de tamizaje de hemoglobina en niños, suplementación de micronutrientes y otra información de interés. La población de esta base de datos será organizada por cada comunidad (distrito/provincia) y se constituirá un punto de datos con una serie de variables asociadas. El único criterio de exclusión de la muestra es tener datos con información incompleta o de calidad insuficiente, los cuales serán depurados durante el proceso de limpieza de datos.

**d) Definición operacional de variables:**

1. **Variable dependiente** (*resultado a predecir*): Porcentaje de anemia en la comunidad, definido operativamente como la prevalencia de anemia en

niños de 6 meses y 0 días a 35 meses y 29 días en dicha comunidad. Esta prevalencia se calculará como el número de niños con diagnóstico de anemia, dividido entre el total de niños evaluados en esa comunidad, multiplicado por 100.

2. **Variables independientes** (*predictores candidatos*): Se considerarán múltiples variables potencialmente asociadas a la anemia, basadas en la literatura y en la disponibilidad en los conjuntos de datos abiertos. Es importante aclarar que cada variable será extraída de los datos abiertos sin modificar la interpretación de acuerdo a su metodología, pero, dependiendo del análisis previo de las variables, éstas podrían pasar por un proceso de transformación, estandarización o imputación para el entrenamiento del modelo predictivo. Las variables por considerar son las siguientes:

- **Indicadores socioeconómicos:** porcentaje de población en la comunidad en condición de pobreza o pobreza extrema (INEI); nivel educativo promedio de las madres o porcentaje de madres sin educación formal (HIS); porcentaje de hogares sin acceso a agua potable o saneamiento básico (ENAHO); índice de desarrollo humano de la localidad (SINIA).
- **Indicadores demográficos, de servicios de salud y nutrición:** porcentaje de niños con bajo peso al nacer en la comunidad (MINSA); tasa de desnutrición crónica infantil en la comunidad (MINSA); cobertura de suplementos de hierro o multimicronutrientes en niños (MINSA); cobertura de CRED en

menores de 3 años (MINSA); cobertura de atención prenatal con suplemento de hierro en gestantes (MINSA).

- **Variables geográficas/ambientales:** altitud promedio de la comunidad (RENIEC); ruralidad (RENIEC); ubicación por región natural (RENIEC).
- **Periodo temporal:** año o periodo del que provienen los datos de anemia, para capturar posibles tendencias en el tiempo o efectos coyunturales.

**e) Procedimientos y técnicas de recolección de datos:** Para cumplir el primer objetivo específico, se realizará la recopilación de datos secundarios desde las plataformas oficiales. Se accederá a la Plataforma Nacional de Datos Abiertos del Perú (19), en la sección de Salud, y se utilizará el conjunto de datos titulado "Anemia" y todos los subconjuntos relacionados con anemia. Este *dataset* contiene los registros de casos de anemia (códigos CIE-10 D50-D53, D64) atendidos en establecimientos de salud a nivel nacional, desagregados por edad, sexo, fecha y lugar de atención. Cabe mencionar que el *dataset* "Anemia" y sus complementarios comparten un identificador único por cada persona natural ("id\_persona", identificador anonimizado) y por cada unidad geográfica ("id\_ubigeo", identificador no anonimizado), permitiendo cruzar la información entre ellos. Adicionalmente, se recopilarán datos públicos del INEI, SINAI, RENIEC y otras fuentes para obtener las variables independientes necesarias; en el caso de estas fuentes, como se tratan de datos socio-demográficos, se utilizará el campo "id\_ubigeo" del conjunto de datos principal para cruzar las fuentes de información.

En las variables que sea posible, se aprovecharán conjuntos de datos integrados del Repositorio Único Nacional de Información en Salud (REUNIS) del MINSA, que provee indicadores de anemia por departamentos, provincias y distritos alimentados por datos de establecimientos de salud.

Una vez obtenidos los datos, se procederá a la integración, limpieza, transformación e imputación de los mismos. Esto implica unir las bases de datos de distintas fuentes mediante identificadores comunes de modo que cada unidad de análisis tenga asociada tanto su prevalencia de anemia como los posibles predictores. Se revisará la calidad de los datos, aplicando técnicas como el manejo de valores faltantes, corrección de inconsistencias y transformación de variables.

Para garantizar la transparencia del proceso, se documentarán todos los pasos de obtención y limpieza en un cuaderno de bitácora o script reproducible de la plataforma Google Colab ([colab.research.google.com](https://colab.research.google.com)). Una vez consolidada la base de datos final, se procederá con la construcción del modelo de inteligencia artificial.

**f) Aspectos éticos:** Este estudio no implica experimentación con seres humanos ni recolección directa de datos personales nuevos. Las bases de datos abiertas del MINSA son información pública agregada o anonimizada, por lo que no contienen identificadores personales que permitan comprometer la confidencialidad de individuos. Aun así, el investigador se compromete a manejar los datos con responsabilidad, almacenándolos de forma segura y únicamente con fines de investigación. Se obtendrá la aprobación del Comité de Ética de la Universidad Peruana Cayetano Heredia para el análisis de datos secundarios.

**g) Plan de análisis de datos:** Para cumplir los objetivos específicos, se seguirá la siguiente secuencia:

- **Análisis descriptivo y exploratorio:** Se calcularán medidas resumen de la prevalencia de anemia por comunidad (media, mediana, desviación estándar, mínimos y máximos). Se explorará la distribución geográfica: por ejemplo, se puede elaborar un mapa del Perú sombreando las provincias o distritos según su nivel de anemia, identificando visualmente clústeres de alta prevalencia. También se realizarán tablas comparando la prevalencia media de anemia según categorías de algunas variables (ej. promedio por quintil de pobreza, por región natural, por nivel de urbanización) para detectar relaciones iniciales. Correlaciones simples de Pearson o Spearman entre la tasa de anemia y los posibles predictores cuantitativos ayudarán a identificar qué factores muestran asociación lineal. En esta etapa, cualquier hallazgo es exploratorio y servirá para orientar la selección de variables e interpretación posterior.
- **Construcción del modelo predictivo:** Para el desarrollo del modelo predictivo, se consideran los siguientes pasos:
  - Se definirá el conjunto de datos en conjunto de entrenamiento y conjunto de prueba (*test*). Se asignará aleatoriamente alrededor del 80% de los puntos de datos para entrenar el modelo y reservar el 20% restante para probar el desempeño del modelo.
  - Luego se aplicarán algoritmos de árboles de decisión, en específico *Random Forest* (Bosque Aleatorio), el cual suele desempeñarse bien con puntos de datos con muchas variables, además que suele

mantener su poder predictivo incluso cuando no se cuente con datos completos.

- El modelo entrenado será ajustado y optimizado. Se utilizarán la técnica de búsqueda de hiperparámetros para optimizar parámetros relevantes del algoritmo. La selección del mejor modelo se basará en su desempeño promedio en el conjunto de validación.

- **Validación del modelo:** Una vez entrenados los modelos, se realizará la validación final del modelo utilizando el conjunto de prueba (el 20% de los datos reservados que no se usaron durante el entrenamiento). Sobre este conjunto de prueba se calcularán las siguientes métricas de desempeño:

- Exactitud (*accuracy*): Mide la proporción general de predicciones correctas sobre el total de casos evaluados. Se calcula como:  $(\text{Verdaderos Positivos} + \text{Verdaderos Negativos}) / (\text{Total de Casos})$ .
- Recuperación (*recall*): Evalúa la proporción de casos identificados como positivos por el modelo que realmente lo eran. Se calcula como:  $(\text{Verdaderos Positivos}) / (\text{Total Positivos})$ .
- Precisión (*precision*): Mide la proporción de casos positivos reales que el modelo fue capaz de identificar correctamente. Se calcula como:  $(\text{Verdaderos Positivos}) / (\text{Casos de Anemia})$ .
- Área bajo la curva (AUC): Representa la capacidad del modelo para distinguir entre las clases a través de diferentes umbrales de clasificación.

El cálculo de estas métricas se realizará utilizando funciones de la biblioteca Scikit-learn en Python, específicamente con el paquete "sklearn.metrics"

- **Interpretación de resultados:** Se examinarán los predictores más influyentes en el modelo. Esto servirá para validar si el modelo concuerda con la evidencia epidemiológica existente y para generar conocimiento adicional. Se contrastarán estos resultados con la literatura.
- **Síntesis y respuesta a la pregunta de investigación:** Finalmente, se integrarán los hallazgos evaluando en qué medida el modelo logró el objetivo. Se discutirá si la precisión obtenida es suficiente para usos prácticos. También se identificarán limitaciones: posibles sesgos en los datos, variables importantes que no estuvieron disponibles, y la generalización del modelo solo a contextos similares a los datos usados.

Los pasos descritos se van a realizar utilizando idiomas de programación orientados a *machine learning*, específicamente Python con las bibliotecas pandas, NumPy, scikit-learn y TensorFlow. Se asegurará la reproducibilidad guardando los scripts y se adjuntarán fragmentos de código relevantes en anexos si es pertinente para la comprensión del método.

## 6. REFERENCIAS BIBLIOGRÁFICAS

1. Al-kassab-Córdova A, Mendez-Guerra C, Robles-Valcarcel P, Iberico-Bellomo L, Alva K, Herrera-Añazco P, et al. Inequalities in anemia among Peruvian children aged 6–59 months: A decomposition analysis. *Front Public Health*. 31 de marzo de 2023;11:1068083.
2. Encuesta Demográfica y de Salud Familiar - ENDES 2024 [Internet]. [citado 18 de marzo de 2025]. Disponible en: <https://www.gob.pe/institucion/inei/informes-publicaciones/3364616-encuesta-demografica-y-de-salud-familiar-endes-2024>
3. Colegio Médico del Perú. INFORME DEL SEMINARIO LA ANEMIA INFANTIL EN EL PERÚ: SITUACIÓN Y RETOS, UNA NUEVA PERSPECTIVA. LA ANEMIA INFANTIL EN EL PERÚ: SITUACIÓN Y RETOS, UNA NUEVA PERSPECTIVA [Internet]. 1 de noviembre de 2023

[citado 18 de marzo de 2025]; Disponible en:  
<https://repositorio.cmp.org.pe/handle/20.500.12971/89>

4. Grantham-McGregor S, Ani C. A review of studies on the effect of iron deficiency on cognitive development in children. *J Nutr.* febrero de 2001;131(2S-2):649S-666S; discussion 666S-668S.
5. Oppenheimer SJ. Iron and Its Relation to Immunity and Infectious Disease. *J Nutr.* 1 de febrero de 2001;131(2):616S-635S.
6. Kwak BO, Kim K, Kim SN, Lee R. Relationship between iron deficiency anemia and febrile seizures in children: A systematic review and meta-analysis. *Seizure.* 1 de noviembre de 2017;52:27–34.
7. Abuga KM, Nairz M, MacLennan CA, Atkinson SH. Severe anaemia, iron deficiency, and susceptibility to invasive bacterial infections. *Wellcome Open Res.* 2 de febrero de 2023;8:48.
8. Palmer AC, Bedsaul-Fryer JR, Stephensen CB. Interactions of Nutrition and Infection: The Role of Micronutrient Deficiencies in the Immune Response to Pathogens and Implications for Child Health. *Annu Rev Nutr.* 29 de agosto de 2024;44(Volume 44, 2024):99–124.
9. Shi J yuan, Yue S jin, Chen H shuang, Fang F yu, Wang X lian, Xue J jun, et al. Global output of clinical application research on artificial intelligence in the past decade: a scientometric study and science mapping. *Syst Rev.* 15 de marzo de 2025;14(1):62.
10. Perumandla R, III R. Population Health Case Management Automation: A Technical Architecture and Implementation Guide. *Int J Inf Technol Manag Inf Syst.* 13 de marzo de 2025;16:1094–107.
11. Raff D, Stewart K, Yang MC, Shang J, Cressman S, Tam R, et al. Improving Triage Accuracy in Prehospital Emergency Telemedicine: Scoping Review of Machine Learning–Enhanced Approaches. *Interact J Med Res.* 11 de septiembre de 2024;13:e56729.
12. Manna A, Dewan D, Sheet D. Structured hashing with deep learning for modality, organ, and disease content sensitive medical image retrieval. *Sci Rep.* 14 de marzo de 2025;15(1):8912.
13. Ozobu CO, Adikwu FE, Odujobi O, Onyeke FO, Nwulu EO. Advancing Occupational Safety with AI-Powered Monitoring Systems: A Conceptual Framework for Hazard Detection and Exposure Control. 2025;9(1).
14. Martínez TMB, Restrepo FP de la H. Inteligencia artificial en las enfermedades transmisibles: impacto en la toma de decisiones y la salud. *Rev Cuba Salud Pública [Internet].* 27 de agosto de 2022 [citado 17 de marzo de 2025];48. Disponible en:  
<https://revsaludpublica.sld.cu/index.php/spu/article/view/2669>

15. Panduro BC. Aplicación del algoritmo del bosque aleatorio a un modelo de clasificación de la anemia en niños peruanos. *Mediciego*. 19 de septiembre de 2022;28:e3471–e3471.
16. Zanjurne PS, Patil VV. Statistical Modelling for Determinants of Anaemia among Women at Reproductive Age. *Int J Renew Energy Exch*. 14 de septiembre de 2023;11(9):127–32.
17. The Intersection of a Child's Demographics and Household Socioeconomic Status in the Multimorbidity of Malaria, Anaemia, and Malnutrition among Children Aged 6-59 Months in Nigeria - PubMed [Internet]. [citado 18 de marzo de 2025]. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/38791859/>
18. Girma S, Abdureshid N, Ayele K, Dagne I, Mekonnen BA, Abate S, et al. Burden and determinants of anemia among lactating women in Ethiopia: evidence from demographic health survey. *Sci Rep*. 27 de junio de 2024;14(1):14816.
19. Plataforma Nacional de Datos Abiertos [Internet]. [citado 18 de marzo de 2025]. Disponible en: <https://www.datosabiertos.gob.pe/>

## **7. PRESUPUESTO Y CRONOGRAMA**

El presupuesto estimado para la planificación, ejecución y publicación del presente proyecto de investigación es de S/. 1,500.00 (ver Anexo N° 1).

El presente estudio se ha planificado para ejecutarse en un periodo de 8 meses desde la elaboración del presente informe (ver Anexo N° 2).

## **8. ANEXOS**

ANEXO N° 1 – PRESUPUESTO

ANEXO N° 2 - CRONOGRAMA

**ANEXO N° 1 – PRESUPUESTO**

<b>Rubro</b>	<b>Detalle</b>	<b>Monto (S/.)</b>
Personal investigador	Investigador principal (dedicación parcial 8 meses)	0.00 ( <i>Trabajo de tesis no remunerado</i> )
Equipo de cómputo	Uso de computadora personal y software (Python, R)	0.00 ( <i>Recursos existentes</i> )
Acceso a datos	Datos abiertos del MINSA (descarga en línea)	0.00
Difusión de resultados	Publicación de artículo científico y presentación	1,500.00
<b>Total previsto:</b>		<b>1,500.00</b>

**ANEXO N° 2 - CRONOGRAMA**

N.º	ETAPAS	2025							
		MAR	ABR	MAY	JUN	JUL	AGO	SEP	OCT
1	Elaboración del proyecto.								
2	Presentación del proyecto. Evaluación por Comité de Ética.								
3	Solicitud de datos a instituciones correspondientes.								
4	Procesamiento de datos.								
5	Análisis exploratorio de los datos								
6	Desarrollo y entrenamiento del modelo predictivo								
7	Validación y ajuste final del modelo								
8	Elaboración del informe.								
9	Presentación del informe.								
10	Sustentación.								