

UNIVERSIDAD PERUANA UNIÓN
FACULTAD DE INGENIERIA Y ARQUITECTURA
Escuela Profesional de Ingeniería de Sistemas



**Modelo de clasificación basado en entropía difusa para la
identificación de factores que contribuyen a la prevalencia de
anemia en niños peruanos**

Tesis para obtener el Título Profesional de Ingeniero de Sistemas

Autor:

Jhon Vilcarana Tintaya
Jorge Luis Quispe Sumire
Hu'bert Eli Falcon Lanchi

Asesor:

Doctor Juan Jesus Soria Quijaite

Lima, mayo de 2026

DECLARACIÓN JURADA DE ORIGINALIDAD DE TESIS

Yo Juan Jesus Soria Quijaite docente de la Facultad de Ingeniería y Arquitectura Escuela Profesional de Ingeniería de Sistemas , de la Universidad Peruana Unión.

DECLARO:

Que la presente investigación titulada: “**Modelo de Clasificación Basado en Entropía Difusa para la Identificación de Factores que Contribuyen a la Prevalencia de Anemia en Niños Peruanos**” del (los) autor (autores) (Nombres y apellidos de los autores) tiene un índice de similitud de 3 % verificable en el informe del programa Turnitin, y fue realizada en la Universidad Peruana Unión bajo mi dirección.

En tal sentido asumo la responsabilidad que corresponde ante cualquier falsedad u omisión de los documentos como de la información aportada, firmo la presente declaración en la ciudad de Lima, a los 12 días del mes de junio del año 2026



Juan Jesus Soria Quijaite

ACTA DE SUSTENTACIÓN DE TESIS

En Lima, Naña, Villa Unión, a. 20... día(s) del mes de mayo... del año 2026, siendo las 08:30 horas, se

reunieron los miembros del jurado en la Universidad Peruana Unión Campus Lima, bajo la dirección del (de la) presidente(a):

Mg. Immer Elias Cuellar Rodriguez, el (la) secretario(a): Mg. Lennin Henry

Centurión Julca y los demás miembros: Mg. Nemias Saboya Rios

Mg. Fernando Manuel Asin Gomez y el (la) asesor(a) Dr. Juan Jesus Soria

Guigaito con el propósito de administrar el acto académico de sustentación de la tesis titulado:

"Modelo de Clasificación Basado en Entropía Difusa para la Identificación de Factores que Contribuyen a la Prevalencia de Anemia en Niños Peruanos"

del(los) bachiller(es): a) Jhon Vilcarana Tintaya

b) Jorge Luis Guispe Sumire

c) Hu'bert Eli Falcon Lanchi

conducente a la obtención del título profesional de:

Ingeniero de Sistemas

El Presidente inició el acto académico de sustentación invitando al (a la) / a (los) (las) candidato(a)s hacer uso del tiempo determinado para su exposición. Concluida la exposición, el Presidente invitó a los demás miembros del jurado a efectuar las preguntas y aclaraciones pertinentes, las cuales fueron absueltas por al (a la) / a (los) (las) candidato(a)s. Luego, se produjo un receso para las deliberaciones y la emisión del dictamen del jurado.

Posteriormente, el jurado procedió a dejar constancia escrita sobre la evaluación en la presente acta, con el dictamen siguiente:

Bachiller (a): Jhon Vilcarana Tintaya

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	17.2	B+	Muy bueno	Sobresaliente

Bachiller (b): Jorge Luis Guispe Sumire

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	15.6	B	Bueno	Muy bueno

Bachiller (c): Hu'bert Eli Falcon Lanchi

CALIFICACIÓN	ESCALAS			Mérito
	Vigesimal	Literal	Cualitativa	
Aprobado	16.1	B	Bueno	Muy bueno

(*) Ver parte posterior

Finalmente, el Presidente del jurado invitó al (a la) / a (los) (las) candidato(a)s a ponerse de pie, para recibir la evaluación final y concluir el acto académico de sustentación procediéndose a registrar las firmas respectivas.

Presidente/a

Secretario/a

Asesor/a

Miembro

Miembro

Bachiller (a)

Bachiller (b)

Bachiller (c)

Esta sustentación fue realizada de manera virtual u online sincrónica según conforme al Reglamento General de Grados y Títulos.

Modelo de clasificación basado
en entropía difusa para la
identificación de factores que
contribuyen a la prevalencia de
anemia en niños peruanos

Fuzzy Entropy Based
Classification Model for the
Identification of Factors
Contributing to the Prevalence of
Anemia in Peruvian Children

Resumen

La anemia infantil constituye uno de los principales problemas de salud pública en los países en desarrollo, donde las deficiencias nutricionales afectan gravemente el crecimiento y desarrollo de millones de niños. Según la Organización Mundial de la Salud (OMS), alrededor del 40 % de los niños de entre 6 y 59 meses sufren esta condición en todo el mundo. La identificación temprana es fundamental para diseñar intervenciones eficaces; Sin embargo, los métodos tradicionales y los modelos convencionales de aprendizaje automático presentan limitaciones en la gestión de la complejidad e incertidumbre de los datos clínicos. En este estudio, se propone un modelo de clasificación que combina lógica difusa, entropía difusa y un clasificador de Extreme Gradient Boosting (XGBoost) con ajuste de umbral. El objetivo es captar tanto la incertidumbre como las relaciones no lineales presentes en los historiales médicos. Se evaluaron siete configuraciones: cuatro arquitecturas modernas de línea base (Random Forest, XGBoost, LightGBM y CatBoost), dos modelos mejorados con entropía difusa (Random Forest y XGBoost) y el modelo híbrido propuesto. Los resultados indican que el modelo XGBoost con entropía difusa y ajuste de umbral logró el mejor rendimiento, logrando una Precisión de 0,904, una puntuación F de 0,865, un ROC-AUC de 0,963 y un PR-AUC de 0,938, superando todas las configuraciones de referencia. Estos hallazgos demuestran que la integración de representaciones difusas de la incertidumbre con técnicas avanzadas de clasificación ofrece una estrategia robusta y flexible para apoyar la toma de decisiones en salud pública y fortalecer la detección de la prevalencia de la anemia infantil.

Abstract

Childhood anemia constitutes one of the main public health problems in developing countries, where nutritional deficiencies seriously affect the growth and development of millions of children. According to the World Health Organization (WHO), about 40% of children aged 6 to 59 months suffer from this condition worldwide. Early identification is critical to designing effective interventions; however, traditional methods and conventional machine learning models have limitations in managing the complexity and uncertainty of clinical data. In this study, a classification model is proposed that combines fuzzy logic, fuzzy entropy, and an Extreme Gradient Boosting (XGBoost) classifier with threshold adjustment. The goal is to capture both the uncertainty and non-linear relationships present in medical records. Seven configurations were evaluated: four modern baseline architectures (Random Forest, XGBoost, LightGBM, and CatBoost), two models enhanced with fuzzy entropy (Random Forest and XGBoost), and the proposed hybrid model. The results indicate that the XGBoost model with fuzzy entropy and threshold adjustment achieved the best performance, achieving an Accuracy of 0.904, an F_1 -score of 0.865, a ROC-AUC of 0.963, and a PR-AUC of 0.938, surpassing all reference configurations. These findings demonstrate that the integration of fuzzy representations of uncertainty with advanced classification techniques offers a robust and flexible strategy to support public health decision-making and strengthen the detection of childhood anemia prevalence.