



UNIVERSIDAD NACIONAL DEL ALTIPLANO
FACULTAD DE INGENIERÍA ESTADÍSTICA E INFORMÁTICA
ESCUELA PROFESIONAL DE INGENIERÍA ESTADÍSTICA E
INFORMÁTICA



MODELO LOGIT PARA IDENTIFICAR FACTORES ASOCIADOS
EN LA PREVALENCIA DE ANEMIA EN MENORES DE 3 A 59
MESES DEL CENTRO DE SALUD I-3 VALLECITO PUNO-2021

TESIS

PRESENTADA POR:

Bach. ROSA ANGELICA PAYEHUANCA MARCA

PARA OPTAR EL TÍTULO PROFESIONAL DE:
INGENIERO ESTADÍSTICO E INFORMÁTICO

PUNO – PERÚ

2023



Reporte de similitud

NOMBRE DEL TRABAJO

MODELO LOGIT PARA IDENTIFICAR FACTORES ASOCIADOS EN LA PREVALENCIA DE ANEMIA EN MENORES DE 3 A 59 MESES

AUTOR

ROSA ANGELICA PAYEHUANCA MARCA

RECuento de palabras

38581 Words

RECuento de caracteres

155265 Characters

RECuento de páginas

120 Pages

Tamaño del archivo

2.2MB

Fecha de entrega

Jul 20, 2023 4:40 PM GMT-5

Fecha del informe

Jul 20, 2023 4:42 PM GMT-5

● 12% de similitud general

El total combinado de todas las coincidencias, incluidas las fuentes superpuestas, para cada base de datos

- 10% Base de datos de Internet
- Base de datos de Crossref
- 6% Base de datos de trabajos entregados
- 3% Base de datos de publicaciones
- Base de datos de contenido publicado de Crossref

● Excluir del Reporte de Similitud

- Material bibliográfico
- Bloques de texto excluidos manualmente
- Coincidencia baja (menos de 12 palabras)



Firmado digitalmente por TITO LIPA
Jose Pantilio FAU 20145496170 hard
Motivo: Soy el autor del documento
Fecha: 20.07.2023 20:32:56 -05:00



Firmado digitalmente por CARPIO
VARGAS EDGAR ELOY
Motivo: Soy el autor del documento
Fecha: 20.07.2023 16:45:48 -05:00

Resumen



DEDICATORIA

*A DIOS por permitirme llegar a este punto,
por iluminarme en cada momento de mi vida, por
concederme la serenidad para aceptar las cosas y
darme el valor para seguir adelante.*

*A mi Madre Margarita Q.E.P.D quien
desde donde esta me ha estado cuidando y guiando
cada día de mi vida.*

*A mi Papa Adolfo quien me brindo su apoyo
incondicional, económico y moral en todo momento,
por el gran amor y confianza que me dio cada día,
por el respaldo que siempre me ha demostrado a lo
largo de mis estudios.*

*A mis hermanos a quienes amo con toda mi
vida Américo, Edy, Richard y Mónica, quienes en
todos estos años fueron y serán mi motor y motivo de
seguir adelante, agradecerles por estar siempre
conmigo en las buenas y malas por la confianza que
me tienen.*

R. ANGELICA PAYEHUANCA M.



AGRADECIMIENTOS

A la Universidad Nacional del Altiplano Puno, por ser mi alma mater estudiantil y hacer de mí una gran profesional.

A mis maestros de la Facultad de Ingeniería Estadística e Informática, por su entrega a la docencia, el haberme brindado sus conocimientos y enseñado a amar mi profesión

A mi compañero y gran amigo Wilber por su apoyo incondicional, durante el desarrollo de este estudio.

A mi asesor y jurados de tesis, por sus orientaciones durante el desarrollo hasta la finalización de este trabajo de investigación.

R. ANGELICA PAYEHUANCA M.



ÍNDICE GENERAL

DEDICATORIA

AGRADECIMIENTOS

ÍNDICE GENERAL

ÍNDICE DE TABLAS

ÍNDICE DE FIGURAS

INDICE DE ACRONIMOS

RESUMEN 13

ABSTRACT..... 14

CAPÍTULO I

INTRODUCCIÓN

1.1. PLANTEAMIENTO DEL PROBLEMA..... 16

1.1.1. Formulación del problema 18

1.2. JUSTIFICACIÓN DEL PROBLEMA 18

1.3. OBJETIVOS DE LA INVESTIGACIÓN..... 20

1.3.1. Objetivo general..... 20

1.3.2. Objetivos específicos..... 20

1.4. HIPÓTESIS DE LA INVESTIGACIÓN 20

CAPÍTULO II

REVISION DE LITERATURA

2.1. ANTECEDENTES DE LA INVESTIGACIÓN..... 21

2.1.1. Antecedentes internacionales 21



2.1.2. Antecedentes nacionales	23
2.1.3. Antecedentes regionales	25
2.2. MARCO TEÓRICO	27
2.2.1. Regresión logística	27
2.2.2. El modelo de regresión logística	28
2.2.3. Tipos de regresión logística.....	29
2.2.4. Modalidades	29
2.2.5. Modelo de regresión logística múltiple.....	30
2.2.6. Estimación de los parámetros.....	30
2.2.7. Pruebas de significancia	32
2.2.8. Estadístico wald.....	33
2.2.9. Pruebas de bondad de ajuste.....	33
2.2.10. Matriz de confusión	34
2.2.11. Curva ROC.....	35
2.2.12. Machine learning.....	36
2.2.13. Variables dummy	36
2.2.14. Funcion de verosimilitud	37
2.2.15. La curva logistica	37
2.2.16. Regresion logit en machine learning.....	37
2.2.17. Métodos de selección de variables.....	38
2.2.18. Anemia	38
2.2.19. causas de la anemia infantil	39
2.2.20. Clasificación de anemia	39



2.2.21. Estado nutricional	39
2.2.22. Evaluación del estado nutricional	40
2.2.23. Indicadores antropométricos	41
2.2.24. Factores de riesgo asociados	42
2.2.25. Factores biológicos del niño	43
2.2.26. Factores sociodemográficos	43
2.2.27. Factores gineco-obstétricos	45
2.2.28. Modelo aplicativo	46
2.2.29. Interpretación de los resultados.....	48
2.2.30. Marco conceptual	50

CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. LUGAR DE ESTUDIO	53
3.2. TIPO Y DISEÑO DE INVESTIGACIÓN	53
3.3. POBLACIÓN Y MUESTRA.....	54
3.4. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS	55
3.5. MÉTODO DE ANÁLISIS DE DATOS.....	57

CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

4.1. RESULTADOS.....	59
4.2. DISCUSIÓN	86
V. CONCLUSIONES.....	89
VI. RECOMENDACIONES	90



VII. REFERENCIAS BIBLIOGRÁFICAS.....	91
ANEXOS.....	96

Área: Estadística

Línea: Modelos Predictivos Uni y Multivariantes

FECHA DE SUSTENTACION: 24 de julio de 2023



ÍNDICE DE TABLAS

Tabla 1.	Resumen de modelo 00 con 15 variables significativas para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	67
Tabla 2.	Coefficientes de modelo 00 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	68
Tabla 3.	Matriz de confusión con el modelo 00 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	69
Tabla 4.	Métricas de evaluación con el modelo 00 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	70
Tabla 5.	Resumen de modelo 01 con 13 variables significativas para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	70
Tabla 6.	Coefficientes de modelo 01 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	71
Tabla 7.	Matriz de confusión con el modelo 01 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	72
Tabla 8.	Métricas de evaluación con el modelo 01 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	73
Tabla 9.	Resumen de modelo 02 con 10 variables significativas para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	74
Tabla 10.	Coefficientes de modelo 02 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	74
Tabla 11.	Matriz de confusión con el modelo 02 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	75



Tabla 12.	Métricas de evaluación con el modelo 02 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	76
Tabla 13.	Resumen de modelo 03 método de frodwrad selection para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	77
Tabla 14.	Coeficientes de modelo 03 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	77
Tabla 15.	Matriz de confusión con el modelo 03 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	78
Tabla 16.	Métricas de evaluación con el modelo 03 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	79
Tabla 17.	Resumen de modelo 04 método de Backward selection para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	80
Tabla 18.	Coeficientes de modelo 04 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021	80
Tabla 19.	Matriz de confusión con el modelo 04 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	81
Tabla 20.	Métricas de evaluación con el modelo 04 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	82
Tabla 21.	Resumen de los resultados de métricas de evaluación de los modelos planteados.	83
Tabla 21.	Resumen de los coeficientes significativos según los factores estudiados en el modelo optimo.....	84



ÍNDICE DE FIGURAS

Figura 1.	Modelo aplicativo basado en la teoría de Data Science.....	46
Figura 2.	Nivel de Anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.....	59
Figura 3.	Nivel de Anemia según el Sexo en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.....	60
Figura 4.	Nivel de Anemia en menores de 3 a 59 meses según el Estado Civil de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	61
Figura 5.	Nivel de Anemia en menores de 3 a 59 meses según el Nivel de Instrucción de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	62
Figura 6.	Nivel de Anemia en menores de 3 a 59 meses según el Número de hijos de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	63
Figura 7.	Nivel de Anemia en menores de 3 a 59 meses según Ingreso mensual de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	64
Figura 8.	Nivel de Anemia en menores de 3 a 59 meses según Anemia durante el embarazo de las madres atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	65
Figura 9.	Correlaciones según el nivel de hemoglobina y a las variables numéricas en los menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	66
Figura 10.	Curva Roc y área bajo la curva del modelo 02 con diferentes umbrales numéricas en los menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.....	82
Figura 11.	Métricas de evaluación de los diferentes modelos planteados según métrica.....	83



INDICE DE ACRONIMOS

C.S. Centro de Salud

dl: decilitro

ENDES: Encuesta Demográfica y de Salud Familiar

Gr/dl: Gramos por decilitro

Hb: Hemoglobina

HIS: Hospital Information System

INEI: Instituto Nacional de Estadística e informática

LM: Lactancia Materna

LME: Lactancia Materna Exclusiva

LA: Lactancia Artificial NSE: Nivel socio económico

Mg: miligramos

MINSA Ministerio de Salud

OMS: Organización Mundial de la Salud

OPS: Organización Panamericana De La Salud

SIEN Sistema Integral De Estado Nutricional



RESUMEN

la regresión logística es una técnica versátil en el aprendizaje automático. El presente estudio tuvo como objetivo: determinar el modelo óptimo para identificar los factores asociados en la prevalencia de anemia. El tipo de investigación fue cuantitativa de alcance descriptivo correlacional, el diseño de investigación es no experimental de corte transversal. Para lo cual se aplicó el modelo logit, la población fue el total de niños menores de 3 a 59 meses que acuden al centro de salud I-3 Vallecito Puno, durante el año 2021, el tamaño de muestra fue la información recolectada del 100% resultando un tamaño de muestra de 410 menores de 3 a 59 meses. La información de la prevalencia de anemia fue facilitada mediante el Sistema (SIEN) His-Minsa de la red de Salud Puno, asimismo se realizó un cuestionario para la recolección de datos de la madre. Se obtuvo el modelo óptimo de $Ln(odds) = [-0.778(Peso) - 2.0645(Sexo F) - 0.868(N CCred) + 1.60(Sachet No) + 1.548(Casada) + 5.56(Ane Madre gest)]$. Entre los factores identificados que influyen en el modelo fueron : factores biológicas el peso de niño(sig=0.000) y la categoría de sexo(sig=0.012), de los factores de SIEN CRED el número de consultas de crecimiento(sig=0.000) y consumo de sachet(sig=0.021), seguido de factores sociodemográficos la categoría de condición civil casada(sig=0.034) finalmente los factores gineco obstétricos de la madre influyeron cuando la madre tuvo anemia en la etapa gestacional(sig=0.000), Las mejores métricas de evaluación fueron sensibilidad de 0.984615, especificidad -1 de 0.1896, precisión del modelo fue de 0.8533 y exactitud del modelo fue de 0.9024, y teniendo una curva ROC de AUC de 97.5199%. concluyendo que se identificó un modelo óptimo con las mejores métricas de validación.

Palabras clave: modelo logit, métricas de precisión, factores asociados, anemia.



ABSTRACT

Logistic regression is a versatile technique in machine learning. The present study had as objective: to determine the optimal model to identify the associated factors in the prevalence of anemia. The type of research was quantitative with a descriptive correlational scope, the research design is non-experimental and cross-sectional. For which the logit model was applied, the population was the total number of children under 3 to 59 months of age who attend the I-3 Vallecito Puno health center, during the year 2021, the sample size was the information collected from 100% resulting in a sample size of 410 children from 3 to 59 months. The information on the prevalence of anemia was provided through the His-Minsa System (SIEN) of the Puno Health network, a questionnaire was also carried out to collect data from the mother. The optimal model of $\text{Ln}(\text{odds}) = [-0.778(\text{Weight}) - 2.0645(\text{Sex F}) - 0.868(\text{NCCred}) + 1.60(\text{Sachet No}) + 1.548(\text{Married}) + 5.56(\text{Ane Mother gest})]$ was obtained. Among the factors identified that influence the model were: biological factors the weight of the child ($\text{sig}=0.000$) and the gender category ($\text{sig}=0.012$), of the factors of SIEN CRED the number of growth consultations ($\text{sig}=0.000$) and consumption of sachets ($\text{sig}=0.021$), followed by sociodemographic factors, the category of civil status married ($\text{sig}=0.034$), finally the gynecological and obstetric factors of the mother influenced when the mother had anemia in the gestational stage ($\text{sig}=0.000$), the best evaluation metrics were sensitivity of 0.984615, specificity -1 of 0.1896, model precision was 0.8533 and model accuracy was 0.9024, and having an AUC ROC curve of 97.5199%. concluding that an optimal model with the best validation metrics was identified.

Keywords: logit model, precision metrics, associated factors, anemia.



CAPÍTULO I

INTRODUCCIÓN

Actualmente la anemia es uno de los problemas de salud más relevante en el mundo, a pesar de conocer su etiología y como enfrentarla es uno de los problemas menos controlado especialmente en la población de niños menores de 5 años.

La anemia es la disminución del número de los eritrocitos que es insuficiente para satisfacer las necesidades del organismo. El valor por debajo del punto de corte del cual se considera que una persona padece de anemia varían en función de la edad, el sexo, la altitud sobre el nivel del mar y estado fisiológico en la que vive la persona. Su valor limite es de 11g/dl para niños de 6 meses a 5 años de edad (Organización Mundial de la Salud, 2011).

Para poder detectar la anemia en niños consiste en medir el nivel de hemoglobina (glóbulos rojos), siendo esta una sustancia que incluye al hierro la cual permite llevar el oxígeno a las células del cuerpo las cuales forman parte de la reparación celular. El hierro se encuentra también en enzimas y en neurotransmisores de allí que su deficiencia tenga consecuencias negativas en el desarrollo cognitivo (Ministerio de Salud, 2017).

En América Latina la anemia en niños menores de 5 años es un problema de salud pública moderado o severo en casi todos los países, donde Brasil ocupa el primer lugar de prevalencia de anemia en niños menores de 5 años (55%), seguido de Bolivia (52%) y como tercer lugar se encuentra Perú con un (50%) (OMS, citado en Álvarez Jiménez, 2019).



En el Perú la anemia constituye un problema de salud pública grave, dada la elevada prevalencia en niños menores de 3 años de edad (42.2%), fue mayor en el área rural (49.0%) y en el área urbana (39.6%) en el primer semestre del 2019.

Según región natural en el primer semestre 2019 la prevalencia de la anemia fue mayor en la región de la sierra (49.4%), selva (44.4%) y costa (37.6%) de niñas y niños menores de 3 años de edad.

En la región de Puno, a diferencia del 2018 el porcentaje de anemia se incrementó en 2.2%. Puno presenta la mayor prevalencia de anemia con 69.9%, que afecta a la mayor parte de menores de 3 a 59 meses de edad, lo que equivale a que siete de cada diez niños padecen de esta enfermedad (Instituto Nacional de Estadística e Informática, 2021)

1.1. PLANTEAMIENTO DEL PROBLEMA

Desde 1958 la Organización Mundial de Salud (OMS) se interesa por los problemas de anemia y viene patrocinando investigaciones en varios países. Esos estudios y otros recientes han revelado una prevalencia muy alta en anemia en personas de todos los grupos de edad, especialmente entre embarazadas y niños menores de cinco años (Organización Mundial de la Salud, 1975).

La anemia infantil afecta al 43.6% de los niños y niñas de 3 a 59 meses de edad, siendo más prevalente entre los niños de 6 a 18 meses, sector en el que 6 de cada 10 niños presenta anemia. La desnutrición infantil se ha reducido en los últimos años, sin embargo, aún afecta al 13.1% de menores de 5 años en el 2016; en las áreas rurales llega al 26.5% y 7.9% en las urbanas. En el último decenio nuestro país ha mostrado singulares avances en la reducción de la desnutrición crónica infantil, aún persisten inequidades a nivel regional y en zonas de pobreza, lo que exige la continuidad de las acciones destinadas a disminuir estas brechas (Ministerio de Salud, 2017).



Según el Instituto Nacional de Estadística e Informática (2021) durante el año 2019 los mayores niveles de anemia en niñas y niños de 3 a 59 meses de edad se registraron en la Sierra (48,8%), seguido de la Selva (44,6%), resto Costa (37,5%) y Lima Metropolitana (30,4%). Según el área de residencia, el porcentaje de las niñas y niños con anemia es mayor en los que residen en el área rural (49,0%) que en el área urbana (36,7%); así lo informó el Instituto Nacional de Estadística e Informática (INEI), según los resultados de la Encuesta Demográfica y de Salud Familiar.

Asimismo, se registró la mayor prevalencia de anemia en Puno (69,9%), Cusco (57,4%), Huancavelica (54,2%), Ucayali (53,7%), Loreto (53,0%), Junín (52,6%), Madre de Dios (51,4%) y Pasco (50,2%). Por el contrario, se observó menor prevalencia de anemia en Cajamarca (28,7%), Provincia de Lima (29,8%), Tacna (32,7%), Moquegua (33,2%), Arequipa (33,9%) y La Libertad (34,2%).

Al 40,0% de niñas y niños de 6 a 36 meses de edad residentes en Perú, se les detectó anemia en el año 2020; por área de residencia, la anemia fue más frecuente entre niñas y niños residentes en el área rural (48,4%) que en el área urbana (36,7%)

Por todo ello es importante conocer los factores influyentes en la prevalencia de anemia en niños de 3 a 59 meses de edad, porque en la actualidad los altos porcentajes de anemia se ha dado en niños menores de 5 años que vienen siendo un problema de salud.

El presente trabajo de investigación tiene como intención desarrollar diferentes modelos de regresión logística con enfoque al aprendizaje automático, determinando las mejores métricas de validación.

Mediante este trabajo de investigación se podrá analizar de la mejor manera cuales son las causas que contribuyen a la prevalencia de anemia en menores de 3 a 59 meses de edad de esta manera evitar dicha deficiencia.



La Organización Mundial de la Salud, grupos de expertos mundiales y el Colegio Médico del Perú recomiendan priorizar políticas públicas y sanitarias para reducir la prevalencia de anemia.

La realización de esta investigación es factible y viable por las siguientes razones: Existencia de datos, ya que la disponibilidad de datos relacionados con la prevalencia de anemia y sus factores asociados en la ciudad de Puno, permitirá llevar a cabo el análisis necesario; acceso a la población de estudio por que se cuenta con el acceso necesario a la población de estudio para la recolección de datos y la aplicación de encuestas; conocimientos y habilidades y colaboración institucional.

La prevalencia de anemia en la ciudad de Puno cuenta con un alto índice de anemia siendo un problema de salud pública de gran importancia y urgencia. Por lo tanto, la solución propuesta, a través de la identificación de factores asociados utilizando un modelo Logit, es de alta prioridad. La comprensión de los determinantes de la anemia en esta región permitirá desarrollar intervenciones efectivas y enfocadas en los factores clave, lo que a su vez contribuirá a la mejora de la salud de la población y la reducción de esta.

1.1.1. Formulación del problema

¿Cuál es el modelo logit óptimo para determinar los factores asociados en la prevalencia de anemia en menores de 3 a 59 meses del Centro de Salud I-3 Vallecito-Puno 2021?

1.2. JUSTIFICACIÓN DEL PROBLEMA

La anemia es un problema de salud pública de gran relevancia a nivel mundial, y el daño que produce es irreversible y persiste aún después de haberla corregido. Puede afectar la actividad cerebral y la capacidad cognitiva, incluyendo el desarrollo del



lenguaje; también daña el desarrollo motor y la coordinación. Se ha comprobado que niños anémicos han bajado hasta 5 puntos su nivel de coeficiente intelectual en comparación con niños no anémicos. De ahí, la importancia de su prevención y detección temprana.

Dada la prevalencia a lo largo de los años en el Departamento de Puno, con altos índices de anemia constituye un desafío significativo para la salud y el bienestar de la población. Es necesario comprender los factores asociados a esta alta prevalencia de anemia, ya que esto permitirá implementar estrategias efectivas de prevención y control.

El propósito de este trabajo de investigación será la aplicación de un modelo logit que permitirá analizar diversas variables y así determinar su probabilidad de tener anemia, lo que contribuirá a la identificación de factores de riesgo y la formulación de estrategias de intervención más efectivas. Esto proporcionará un enfoque metodológico sólido y basado en evidencia para comprender las causas subyacentes de la anemia. Esta investigación beneficiará directamente a la población ya que proporcionará información precisa y actualizada sobre los factores asociados a esta condición. Los resultados obtenidos serán de utilidad de la salud autoridades sanitarias y los responsables de la toma de decisiones, permitiéndole diseñar políticas y programas de salud específicos y orientados a reducir la prevalencia de anemia en la ciudad, así como también a los estudiantes y público general, interesados en el tema de la implementación de modelos logit para diferentes estudios.



1.3. OBJETIVOS DE LA INVESTIGACIÓN

1.3.1. Objetivo general

Determinar el modelo logit óptimo que permite identificar los factores asociados en la prevalencia de anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.

1.3.2. Objetivos específicos

- Describir los factores asociados en la prevalencia de anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.
- Identificar los factores asociados en la prevalencia de anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.
- Establecer el modelo logit con mejores métricas de validación para la prevalencia de anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.

1.4. HIPÓTESIS DE LA INVESTIGACIÓN

Existe un modelo logit óptimo que permite identificar los factores asociados en la prevalencia de anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito Puno -2021.



CAPÍTULO II

REVISION DE LITERATURA

2.1. ANTECEDENTES DE LA INVESTIGACIÓN

2.1.1. Antecedentes internacionales

En Ecuador, Salazar y Guambuete (2023) en su investigación “modelo matemático para establecer los factores asociados a las calificaciones escolares utilizando regresión logística en los estudiantes del primer ciclo del Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar” obtuvieron como principales resultados, que los cofactores del nivel de formación de la madre, cantidad de miembros en el hogar mayor a nueve y estado civil unión libre, son los que incrementan la posibilidad de reprobar el ciclo académico. Para concluir, se determina que el modelo predice el 84.9% de los casos posibles para lo cual, considera las variables más significativas mediante el uso de los Odd Ratios como en el caso del cofactor estado civil unión libre.

En Ecuador, Covri y Maldonado (2023) realizaron su estudio sobre la información proporcionada por la Cooperativa de Ahorro y Crédito (COAC) Jardín Azuayo para determinar la puntuación crediticia de sus socios y obtener beneficios en préstamos. Se aplicó un modelo Logit binario utilizando doce variables para predecir el cumplimiento de pago de préstamos por parte de los socios. Los resultados mostraron una alta capacidad predictiva del modelo, cercana al 90% de las observaciones. Se encontró que los socios que solicitaban créditos con tasas de interés más bajas, residían en zonas urbanas, refinanciaban préstamos, realizaban pagos semestrales y solicitaban microcréditos de acumulación ampliada tenían mayores probabilidades de incumplir en los pagos.



Ventura et al.(2022) en su trabajo de investigación se propuso un modelo de regresión logística binaria para predecir la deserción estudiantil de posgrado en la Universidad Técnica de Manabí. Se encontró que variables como estado civil, situación laboral, edad y salario fueron altamente significativas en la predicción. El modelo fue evaluado analíticamente y obtuvo resultados significativos en pruebas como Chi cuadrado, R^2 de Nagelkerke y Hosmer-Lemeshow. Además, se realizó una clasificación de resultados con una muestra piloto, alcanzando un 98,6% de respuestas correctas. Finalmente, se evaluó el modelo con una muestra de validación, logrando un 99,06% de aciertos.

Cevallos Melo & Vélez Gonzáles (2018) demostraron en su trabajo de investigación Prevalencia de anemia en niños menores de 5 años atendidos en el área de pediatría del hospital general Guasmo Sur, periodo enero-junio 2017 donde Fueron analizados 309 registros hospitalarios obtenidos del área de censo y archivo del mencionado hospital, que del total de los sujetos estudiados 124 pacientes tienen un promedio de edad de 3 años, el cual corresponde (40.12%) y el diagnóstico más común presentado en la población estudiada, de acuerdo con el diagnostico categorizado por el cie 10, fue la anemia por deficiencia de hierro, no especificada que conllevó un 56.63%, teniendo en relación que existen otros como son el nivel socioeconómico que se expresó con un 33.33% relacionado con personas de nivel medio típico 55.25% de la población estudiada profesa practicar la religión cristiana. En relación al IMC (índice de masa corporal), se observó que el 40.12% de la población demostró obtener un IMC normal.

Várela (2019) determinó que la prevalencia de anemia en niños y niñas de 6 a 48 meses que concurren a dos Centros de Atención a la Primera Infancia (CAIF) de la ciudad de Salto fue de 22,5% en toda la muestra, siendo levemente mayor en el CAIF2 (27,4%) que en el CAIF1(18,4%). La frecuencia de anemia fue significativamente menor



en los niños que tuvieron lactancia materna exclusiva hasta los 6 meses de edad, mientras que las restantes variables analizadas no presentaron asociación con la anemia, concluyendo que la disminución observada de la frecuencia de anemia comparada con lo reportado en 2011 sugiere que las políticas de fortificación con hierro de los alimentos y de administración de hierro profiláctico a niños entre 6 y 24 meses ha sido positiva. Adicionalmente, la menor frecuencia de anemia observada en el CAIF1 indica la importancia de la acción coordinada entre los servicios de salud y la sociedad.

2.1.2. Antecedentes nacionales

Castillo (2021), determinó los factores de riesgo relacionados con el trabajo infantil utilizando el modelo de regresión logística múltiple. Mediante esta técnica paramétrica, se identificaron 12 variables que influyen en el riesgo de que un menor esté involucrado en algún tipo de trabajo ($X^2=46.88$; $p_valor=0.0000$). Estas variables incluyen el sexo, el área de residencia, la región rural, la macro región, la capacidad de leer y escribir, la lengua materna, el último nivel de estudios asistido en 2020, la asistencia actual a un centro educativo, el tipo de centro educativo, la relación de parentesco con el jefe del hogar, el nivel de pobreza y el número de miembros en el hogar (variable agrupada).

Yupari-Azabache(2022) analizó los factores de riesgo de mortalidad en pacientes hospitalizados con COVID-19 en el distrito de Trujillo, Perú. Se utilizó un enfoque descriptivo y correlacional, con datos obtenidos del sistema del Ministerio de Salud. La muestra incluyó a 64 pacientes entre marzo y mayo de 2020. Se encontró que la mayoría de los fallecidos eran hombres (85,71%) y tenían una ocupación de jubilados (28,57%), con una edad promedio de 64,67 años. Los síntomas más comunes en los pacientes fallecidos fueron dificultad respiratoria, fiebre, malestar general y tos. Se identificaron signos como disnea y auscultación pulmonar anormal, así como comorbilidades como



enfermedad cardiovascular y diabetes. Un modelo de regresión logística se utilizó para predecir la mortalidad, y factores como edad, sexo, tos, dificultad respiratoria y diabetes fueron identificados como factores de riesgo. El modelo demostró una capacidad considerable para explicar la variación y clasificar correctamente el 90,6% de los casos.

Charapaqui Jacinto (2019) halló que los factores sociodemográficos y materno infantil: área de residencia, grado de instrucción de la madre, sexo del niño, edad del niño y los controles prenatales son factores de riesgo asociados a la anemia en niños de 6 a 36 meses en Perú durante el 2017. La prevalencia de anemia es mayor en niños que se encuentran en el rango de 12 a 23 meses, (59.6%), y en aquellos que residen en área urbana, (66,7%). Tener una madre adolescente incrementa la probabilidad para el niño de tener anemia (OR=1,4) y la anemia en los niños es mayor cuando la madre no ha recibido educación (57.3%). Asimismo, la anemia es más frecuente en el sexo masculino. Tuvo como muestra total a 4911 niños de 6 a 36 meses de edad reportados en la ENDES el año 2017.

Paredes Flores (2019) encontró que los factores biológicos, alimentarios y culturales se relacionan significativamente a la presencia de anemia en niños de 6 a 23 meses de edad, del puesto de salud Intiorko – Tacna, año 2014, con un nivel de significancia $p < 0,05$. Factor biológico: sexo ($p=0,034$), parasitosis ($p=0,048$). Factor alimentario: frecuencia de consumo de alimentos dos veces al día ($p=0,011$), inadecuada cantidad de consumo de alimentos de origen animal ricos en hierro ($p=0,002$), el no consumo de pescado ($p=0,000$), huevo ($p=0,003$), menestra ($p=0,001$) y sangrecita en la semana ($p=0,046$) y bajo consumo de fruta ($p=0,003$), insuficiente suplementación con hierro ($p=0,001$). Factor cultural: grado de instrucción ($p=0,011$) y el nivel de conocimiento regular y bajo de la madre ($p=0,01$



Navarro y oliva (2020) demostró que, para identificar sus factores asociados en la presencia de anemia en menores de 5 años en el Perú, incluyó un estudio de 21 covariables analizados mediante la técnica multivariante de regresión logística binaria y obtuvo como resultados mostraron que la prevalencia en el Perú es de 36,4%, en el área rural esta prevalencia es de 43,2% y en el área urbana es de 33,7%. Demostrando que el modelo clasifica correctamente los casos de presencia de anemia en menores de cinco años a nivel nacional en el Perú en un 69,5% conformado por 14 covariables: área de residencia(OR=1,144), región natural1(OR=1,420), región natural2(OR=1,278), grado de instrucción de la madre(OR=0,884), edad de la madre(OR=0,702), lengua materna de la madre(OR=1,181), índice de riqueza del hogar(OR=0,834), sexo del menor(OR=0,824), edad en meses(OR=0,313), presencia de diarrea en el menor reciente(OR=1,181), visitas prenatales durante el embarazo(OR=0,852), consumo de suplemento de hierro durante el embarazo(OR=0,859), presencia de anemia durante el embarazo(OR=1,380), recibió suplemento de hierro en los últimos 12 meses(OR=1,156), consumo de medicamentos para lombrices intestinales en los últimos 6 meses(OR=0,824).

2.1.3. Antecedentes regionales

Roque (2018) tuvo como objetivo determinar los modelos de regresión logística multinomial para la calidad de fibra de alpaca Huacaya en el distrito de Corani, se consideraron tres categorías: Baby, Fleece y mediana. Se tomaron muestras de fibra de alpaca de la raza Huacaya y se analizaron mediante el OFDA 2000 por el Proyecto Especial Camélidos Sudamericanos. La muestra consistió en 1526 registros de las comunidades de Quelcaya y Chimboya. Los resultados revelaron modelos específicos para cada categoría de calidad de fibra, siendo el Factor de Confort la variable más determinante en la clasificación. El modelo mostró una tasa de clasificación del 88.4%,



lo que indica una buena capacidad de predicción. En conclusión, a medida que el Factor de Confort aumenta, la probabilidad de que la calidad de fibra sea Baby también aumenta.

Mayhua(2020) en su estudio realizado en el distrito de Corani, se determinaron los modelos de regresión logística multinomial para la calidad de fibra de alpaca Huacaya. Se analizaron características como el Factor de Confort, Índice de Curvatura y Longitud de Mecha en 1526 registros de fibra de alpaca. Los resultados mostraron que el Factor de Confort fue la variable más relevante para clasificar la calidad de fibra en Baby, Fleece y mediana. El modelo logró una tasa de clasificación del 88.4% y demostró una buena capacidad predictiva.

Cutipa Huarsaya (2019) determinó que los factores de riesgo materno perinatales asociados a anemia ferropénica en niños menores de 36 meses de los establecimientos de Salud Simón Bolívar I-3 y 4 de noviembre I-3 2016, de la ciudad de Puno son la edad de la madre, anemia de la madre y peso del niño, según la Chi cuadrada existe relación entre factores de riesgo materno perinatal la que se asocia de manera significativa con la anemia ferropénica en niños menores de 36 años. El 50% de los niños en ambos establecimientos presentan anemia leve, el 33,3% anemia moderada y el 16,7% anemia severa, dentro de los factores de riesgo maternales, el 29,6% tienen la edad entre 17 – 35 años, el 27,8% tienen edad gestacional 37-40 semanas y son a término, el 27,8% con periodo intergenésico menor a 2 años, el 31,5% son primípara y el 35,2% presentaron anemia leve en el embarazo. La población fue de 54 niños con anemia ferropénica en ambos establecimientos de salud. Se usó la ficha de revisión documentaria Historia Clínica de los niños. Como factores de riesgo perinatal el 40,7% tienen bajo peso, el 27,1% nacieron a término y el 31,5% reciben lactancia materna exclusiva.

Díaz Cayro (2019) indicó en su trabajo de investigación factores asociados a la presencia de anemia nutricional en niños de 1 a 5 años, Micro red Cono Sur diciembre



2015 marzo 2016, que el 63.91% (108) niños(as) presentan anemia moderada, por otro lado, el 21.30% (36) niños (as) evidencian anemia leve y el 14.79% (25) niños(as) presentan anemia severa. Respecto, a la presencia de anemia nutricional según la edad del niño(a), el 46.75% (79) se encuentran en el grupo de 25 a 48 meses de edad, según el género, el 62.72% (106) son de género femenino, según el espacio intergenésico de la madre, el 53.26% (90) son de espacio intergenésico corto (menor a 2 años), según la edad de la madre, el 76.33% (129) se encuentran en el grupo de 18 a 35 años, según la práctica de lactancia materna, el 62.72% (106) recibieron lactancia materna no exclusiva, según el ingreso económico familiar, el 69.82% (118) tienen un ingreso económico menor a 850 soles. La muestra estuvo conformada por 169 niños(as).

Canaza(2021) En su trabajo considero como objetivo determinar un modelo predictivo de riesgo asociado en la anemia en niños de la MICRORED yauri de la provincia espinar, llegando a la conclusión donde su clasificación de verdaderos negativos (no anémicos pronosticado como no anémicos) fue del 93.8% mientras que su tasa más relevante fue clasificar correctamente a los niños anémicos (verdadero positivos) fue de un 96.2% quedando demostrado gran capacidad predictiva de su modelado.

2.2. MARCO TEÓRICO

2.2.1. Regresión logística

La regresión logística es una de las técnicas estadístico-inferenciales más empleadas en la producción científica contemporánea. Surge en la década del 60 con la aparición de un trabajo sobre el riesgo de padecer una enfermedad coronaria debido a Cornfield, Gordon y Smith en 1961, que contenía la primera aplicación práctica trascendente. El algoritmo para la obtención de los estimadores de máxima verosimilitud de Walker-Duncan en 1967 fue un paso crucial para poder confeccionar programas de



acceso universal, los cuales confluían con el comienzo de la era de las computadoras personales. Tal escenario fue especialmente propicio para el desarrollo de un recurso tan versátil como la RL (Silva & Barroso Utra, 2000).

La regresión logística mezcla dos tradiciones del análisis estadístico: el análisis de tablas de contingencia con el tratamiento de modelos log-lineales, y el análisis de regresión por mínimos cuadrados ordinarios. En ambos casos nos encontramos con limitaciones que la regresión logística resuelve: en el primer caso los modelos de dependencia no podían utilizar variables continuas y en el segundo las variables categóricas no siempre funcionan como buenos predictores (López & Fachelli, 2015).

2.2.2. El modelo de regresión logística

La regresión logística evalúa la influencia de las variables independientes sobre la variable dependiente o de respuesta. Tendremos, por tanto, una variable dependiente Y , que puede ser dicotómica o politómica y una o más variables independientes X , que pueden ser de cualquier naturaleza, cualitativas o cuantitativas. Si la variable Y es dicotómica, podrá tomar el valor "0" si el hecho no ocurre y "1" si el hecho ocurre. Este proceso es denominado binomial ya que solo tiene dos posibles resultados, siendo la probabilidad de cada uno de ellos constante en una serie de repeticiones (Alderete, 2006).

En el modelo clásico de regresión lineal la variable dependiente es cuantitativa, condición que se extiende a las variables independientes si bien podemos utilizar variables cualitativas con una codificación dummy. En el caso de la regresión logística se trata de predecir una variable cualitativa o categórica, con la ventaja, frente al modelo de regresión clásico, de no tener que establecer la serie de condiciones de aplicación que dificultan su utilización y sus posibilidades, en particular, en el contexto de estudios por encuesta.



A diferencia de la regresión lineal pues, con la regresión logística el objetivo es explicar o pronosticar la pertenencia a un grupo, a partir de una variable dependiente categórica o cualitativa, en función de una o más variables independientes que pueden ser tanto cuantitativas como cualitativas. Se trata por tanto de identificar qué características o factores diferencian los grupos definidos por la variable dependiente, de forma similar a como lo hace el análisis discriminante, pero con la ventaja de poder considerar cualquier nivel de medición de las variables independientes.

Al considerar al análisis de regresión logística como técnica destinada al análisis de una relación de dependencia, nos referiremos fundamentalmente a ella como una técnica predictiva y no tanto como técnica destinada a establecer relaciones de causalidad, si bien implícitamente se razone la causalidad. Esto es, cuando diferenciamos a las variables independientes de la dependiente establecemos un modelo explicativo donde se fijan los factores que tienden a favorecer un efecto, a aumentar probabilidades de un comportamiento, que se dé un valor de la variable de dependiente (López & Fachelli, 2015).

2.2.3. Tipos de regresión logística

Vamos a clasificar la regresión logística en dos tipos, fundamentalmente, Regresión logística simple, si en el modelo hay una sola variable independiente y Regresión logística múltiple, si en el modelo hay varias variables independientes (Álvarez Cáceres, 1995).

2.2.4. Modalidades

El análisis de regresión logística tiene dos modalidades: la regresión logística binaria cuando se pretende explicar una característica o suceso dicotómico (estar desempleado o no, abstenerse en las elecciones o no), y la regresión logística multinomial en el caso más general de querer explicar una variable cualitativa politómica. Para ello se

requiere convertir la variable en diversas variables dicotómicas ficticias, es decir, creando tantas variables dicotómicas (dummy) como categorías tenga la variable menos una, la que actuará de categoría de referencia. En este segundo caso se diferencia la situación en que la variable categórica es politómica nominal (la elección de una marca de un producto o la filiación política) o politómica ordinal (el nivel salarial o el grado de acuerdo sobre una cuestión). Por otro lado, desde el punto de vista de las variables independientes, éstas pueden ser cualitativas, tanto dicotómicas como politómicas, o cuantitativas, y se puede considerar tanto el efecto individual de cada una como el efecto de la interacción (López & Fachelli, 2015).

2.2.5. Modelo de regresión logística múltiple

Es una generalización del modelo simple, relaciona la probabilidad de que ocurra un determinado suceso independiente denotado por el vector $X' = (x_1, \dots, x_k)$ con probabilidad condicional $P(Y=1/X)$ en función de k variables independientes que pueden ser cuantitativas, cualitativas o combinadas según sea el tipo de diseño de estudio.

El modelo logístico múltiple es:

$$\log\left(\frac{p_1}{1-p_1}\right) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$$

O también:

$$p_1 = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}{1 + e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}}$$

2.2.6. Estimación de los parámetros

Sea una muestra de n observaciones independientes definido por $(x_{i1}, x_{i2}, \dots, x_{ik}, y_i)$, $i=1, \dots, n$ y como en el caso univariante se elige el vector $\beta' = (\beta_0, \dots, \beta_k)$, el método más usado es el de máxima verosimilitud definido en (3.3.2.2) con

k+1 ecuaciones de verosimilitud que se obtienen derivando el log de la función de verosimilitud respecto a k+1 coeficientes.

Las ecuaciones de verosimilitud son:

$$\sum_{i=1}^n [y_1 - p_1] = 0$$

$$\sum_{i=1}^n [y_1 - p_1] = 0, j = 1, \dots, k; i = 1, \dots, n$$

Sea $\hat{\beta}$ el estimador de máxima verosimilitud para el sistema de ecuaciones de tal modo $\hat{\rho}_1$ es el modelo logístico múltiple de la ecuación (18). El método de estimación de varianzas y covarianzas de los coeficientes estimados, es a través del método de máxima verosimilitud con proceso iterativos, consiste en obtener la matriz de la segunda derivada parcial de la función de verosimilitud estas derivadas parciales tienen la siguiente forma general:

$$\frac{\delta^2 L(\beta)}{\delta \beta_j^2} = - \sum_{i=1}^n \frac{x_{ij}^2 e^{\beta_0 + \beta_1 x_1}}{(1 + e^{\beta_0 + \beta_1 x_1})^2} = - \sum_{i=1}^n X_j^2 p_i (1 - p_i) \dots \dots \dots (19)$$

$$\frac{\delta^2 L(\beta)}{\delta \beta_j \delta \beta_1} = - \sum_{i=1}^n \frac{X_{ij} X_{i1} e^{\beta_0 + \beta_1 x_1}}{(1 + e^{\beta_0 + \beta_1 x_1})^2} = - \sum_{i=1}^n X_{ij} X_{i1} p_i (1 - p_i) \dots \dots \dots (20)$$

Sea una matriz (k+1) x (k+1) que contiene los términos negativos en las ecuaciones (19) y (20) denotado por I (B), llamado matriz de información con varianzas y covarianzas de los coeficientes estimados por la inversa de la matriz de la siguiente forma $\Sigma(B) = I^{-1}(B)$, cuyos elementos de la diagonal son $\sigma^2(\beta_j)$ que es el j-ésimo elemento de la diagonal, la varianza de $\hat{\beta}_j$ y $\sigma(\hat{\beta}_j, \hat{\beta}_1)$ son las covarianzas de $\hat{\beta}_j$ y $\hat{\beta}_1$ para estimar la matriz de información del modelo estimado es $\hat{I}(\hat{\beta}) = X'VX$, donde $X_{nx(k+1)}$



es la matriz de datos de los sujetos y $V_{n \times n}$ es una matriz diagonal cuyo elemento general es $\hat{P}_i(1 - \hat{P}_i)$.

Las matrices son:

$$X = \begin{bmatrix} 1 & X_{1i} & \dots & X_{1k} \\ \dots & \dots & \dots & \dots \\ 1 & X_{ni} & \dots & X_{nk} \end{bmatrix}_{n \times (k+1)} \dots \dots \dots (21)$$

$$V = \begin{bmatrix} \hat{P}(X_1)(1 - \hat{P}(X_1)) & 0 & \dots & 0 \\ 0 \dots & \hat{P}(X_2)(1 - \hat{P}(X_2)) \dots & \dots & 0 \dots \\ 0 & 0 & \dots & \hat{P}(X_n)(1 - \hat{P}(X_n)) \end{bmatrix}_{n \times n} \dots \dots \dots (22)$$

2.2.7. Pruebas de significancia

Como en el caso univariante se prueba la significancia de las variables independientes del modelo mediante la prueba de verosimilitud con la significancia de los $k+1$ parámetros en la ecuación (14) bajo la hipótesis para determinar si las variables independientes influyen significativamente en la probabilidad del suceso del modelo relacionado a la variable del resultado del siguiente modo:

$$H_0: \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1: \text{para algún } \beta_i \neq 0$$

Estadístico de prueba:

$$D \sim \chi^2 \text{ con } k \text{ grados de libertad}$$

Decisión: si $D > \chi^2_{\alpha, k}$ rechazamos H_0 , entonces al menos uno de los coeficientes es diferente de cero y la variable correspondiente influye en la probabilidad del suceso estudiado.



Una vez encontrado el mejor conjunto de variables explicativas que predicen la variable Y, luego se debe evaluar mediante Wald cada coeficiente para determinar cuál o cuáles ingresan al modelo.

2.2.8. Estadístico wald

Evalúa la significancia de los coeficientes se define como el vector matriz de los coeficientes estimados del siguiente modo según la hipótesis:

$$H_0: \beta_i = 0$$

$$H_1: \text{algún } \beta_i \neq 0$$

Estadístico de prueba:

$$W = \hat{\beta}' [\hat{I}(\hat{\beta})]^{-1} \hat{\beta} = \hat{\beta}' (X' V X) \hat{\beta} \sim X_{\alpha, k+1}^2 \dots \dots \dots (24)$$

Donde:

$X_{n \times (k+1)}$ y $V_{n \times n}$, son las matrices de las ecuaciones (21) y (22).

Decisión: si $W > X_{\alpha, k}^2$ rechazamos H_0 con un nivel de significancia fijado, concluimos que la variable independiente influye en la probabilidad del suceso.

2.2.9. Pruebas de bondad de ajuste

Para evaluar la bondad de ajuste del modelo se utiliza la prueba de Hosmer-lemeshow, consiste en calcular para cada observación del conjunto de datos las probabilidades de la variable dependiente que predice el modelo, se agrupa en aproximadamente 10 grupos iguales a partir de las probabilidades esperadas y se compara con las frecuencias observadas mediante una prueba X^2 con j-2 grados de libertad, donde j es el número de grupos formados como se explicó en el modelo simple. El modelo se ajusta bien si no hay evidencias para rechazar la hipótesis nula.

2.2.10. Matriz de confusión

Esta matriz es muy utilizada para validar modelos de machine learning, redes neuronales entre otros en el campo de data science, compara los valores reales con los valores predichos y contabiliza aquellos predichos correctamente como los que no se predijeron bien. Cada fila en una matriz de confusión representa una clase real, mientras que cada columna representa una clase predicha (Géron, 2019).

El valor en la posición (0,0) de la matriz se denomina Verdaderos Negativos (NV), mientras el que se encuentra en la posición (0,1) se hace llamar Falso Positivo (FP). El valor en la posición (1,0) se denomina Falso Negativo (FN) y el que está en la posición (1,1) Verdadero Positivo (VP). Un clasificador perfecto solo tendría verdaderos positivos y verdaderos negativos, por lo que esta matriz de confusión solo tendría valores distintos de cero en su diagonal (de arriba izquierda, a abajo derecha).

Verdaderos negativos	Falsos Positivos
Falso negativo	Verdaderos Positivos

Precisión

Una interesante manera de ver el funcionamiento del modelo es observando la precisión de las predicciones positivas, esto es llamado precisión del clasificador.

$$Precision = \frac{VP}{VP + FP}$$

Exactitud

Contiene la relación entre el número de predicciones correctas y todas las predicciones. Es decir, toma la precisión de los datos válidos y la compara con el número

total de datos devueltos (ya sea que sea exacto o no). Al igual que con la medición anterior, cuanto mayor sea la precisión, con mayor precisión se podrá interpretar la matriz.

$$\text{Exactitud}(\text{accuracy}) = \frac{VP + VN}{VP + FP + VN + FN}$$

Sensibilidad: determina la capacidad del algoritmo para detectar casos positivos. Es decir, determina el éxito que ha mostrado al considerar casos que representan un rasgo particular definido por el programador como positivo.

$$\text{Sensibilidad} = \frac{VP}{VP + FN}$$

Especificidad: Esta métrica está muy relacionada con la primera, ya que expresa la tasa de éxito del algoritmo con respecto a los casos negativos. Compara el número de casos negativos detectados con el número total de casos negativos.

$$\text{Especificidad} = \frac{VN}{FP + VN}$$

Como se puede ver de lo anterior, se puede concluir que la exactitud, precisión, su sensibilidad y especificidad permiten una interpretación de la matriz de interferencia como funcional o no funcional. Solo así se puede evaluar una mayor o menor mejora en el rendimiento de este algoritmo de aprendizaje supervisado.

2.2.11. Curva ROC

La curva de Características de Funcionamiento del Receptor es otra herramienta común utilizada con clasificadores binarios. La curva ROC traza la Tasa de Verdaderos Positivos, otro nombre para la sensibilidad, contra la Tasa de Falsos Positivos (TFP). Los TFP es la proporción de instancias negativas que se clasifican incorrectamente como



positivas, es igual a $1 -$ la Tasa de Verdaderos Negativos (TFN), el cual es la proporción de instancias negativas que se clasifican como negativas. El TFN también es llamada especificidad. Por lo tanto, la curva ROC traza la sensibilidad versus $1 -$ especificidad (Géron, 2019).

Una forma de comparar los clasificadores es medir el área bajo la curva (AUC). Un clasificador perfecto tendrá un AUC igual a 1, mientras que un clasificador puramente aleatorio tendrá un AUC igual 0.5.

2.2.12. Machine learning

Machine learning es una forma de la IA que permite a un sistema aprender de los datos en lugar de aprender mediante la programación explícita. Un modelo de machine learning es la salida de información que se genera cuando entrena su algoritmo de machine learning con datos. Si se proporciona el modelo predictivo con datos, recibirá un pronóstico basado en los datos que entrenaron al modelo (IBM, s/f)

Es importante precisar que la inteligencia artificial es una clasificación aún más amplia que incluye tanto técnicas para el análisis de datos estructurados como el aprendizaje automático y datos no estructurados como procesamiento de lenguaje natural, el aprendizaje automático es un término amplio que agrupa varias estrategias analíticas cuyo propósito es el desarrollo de algoritmos para extraer información de los datos ya sea para explicación, clasificación o predicción.

2.2.13. Variables dummy

Las variables ficticias, también llamadas variables dummy o indicadoras, son variables artificiales que se crean para representar atributos con múltiples niveles o categorías. Por ejemplo, si estamos examinando el impacto del nivel educativo en los ingresos, podemos crear variables ficticias para representar diferentes niveles de



educación, como "educación primaria", "educación secundaria" y "educación universitaria". El objetivo de usar variables ficticias es poder incluir atributos no numéricos en un análisis de regresión, donde todas las variables independientes deben ser numéricas y comparables en escala (Parra, 2022).

2.2.14. Funcion de verosimilitud

La función de verosimilitud es una medida utilizada en estadística para evaluar la adecuación de un modelo estadístico a los datos observados. Se calcula como la probabilidad de observar los datos dados los parámetros del modelo. La maximización de la función de verosimilitud permite encontrar los valores óptimos de los parámetros que mejor se ajustan a los datos. La función de verosimilitud es fundamental en la estimación de máxima verosimilitud, un método comúnmente utilizado para obtener estimaciones de los parámetros en diversos modelos estadísticos (Casella, G., & Berger, R. L.; 2002)

2.2.15. La curva logística

La curva logística es una función matemática que se utiliza para modelar y describir el crecimiento o la propagación de un fenómeno a lo largo del tiempo. Esta curva tiene forma de "S" y se caracteriza por tener un crecimiento inicial rápido, seguido de un período de crecimiento más lento a medida que se acerca a un límite superior. Además, la curva logística también puede describir el decrecimiento de un fenómeno.

2.2.16. Regresion logit en machine learning

La regresión logística es ampliamente utilizada en Machine Learning debido a su capacidad para manejar variables categóricas y continuas, y su interpretación intuitiva de las probabilidades. Es comúnmente aplicada en problemas de clasificación, como la predicción de enfermedades, la detección de fraudes y la evaluación del riesgo crediticio (Smith et al., 2019).



2.2.17. Métodos de selección de variables

Estos métodos involucran agregar o eliminar variables del modelo en función de su impacto en la calidad del ajuste y en los criterios de selección, como el criterio de información de Akaike (AIC) o el criterio de información bayesiano (BIC).

Uno de los métodos comúnmente utilizados es el enfoque de selección de variables paso a paso, que incluye los métodos hacia adelante, hacia atrás y paso a paso bidireccional. Estas técnicas imponen restricciones en los coeficientes de las variables, lo que permite la selección automática de variables al forzar ciertos coeficientes a cero (Smith et al., 2019).

2.2.18. Anemia

La anemia es la reducción de la concentración de la hemoglobina, una condición en la cual la sangre carece de suficientes glóbulos rojos considerados para una determinada talla, edad y sexo (Instituto Nacional de Estadística e Informática, 2018). El hierro es un componente esencial de la hemoglobina, junto con el ácido fólico y la vitamina B12, su función en la oxigenación del organismo es esencial (Vásquez Servín, 2017). Se reconoce también que la vitamina A, la vitamina B2, las vitaminas B6, B12 y el ácido fólico intervienen en la formación de los glóbulos rojos en la médula ósea. Las vitaminas A, C y Riboflavina favorecían la absorción del hierro a nivel intestinal, cumpliendo un rol movilizador del mineral a partir de las reservas; mientras que la vitamina C y E tienen una función antioxidante para la protección de los glóbulos rojos (Ministerio de Salud, 2017). En los niños la deficiencia de hierro es una de las causas más comunes para concebir anemia, la falta de hierro en la alimentación diaria constituye más de la mitad del número total de casos de anemia.

2.2.19. Causas de la anemia infantil

Según el Ministerio de Desarrollo e Inclusión Social (2018) las principales causas de la anemia son:

- Deficiencia de hierro en el embarazo.
- Nacimiento prematuro.
- Bajo peso al nacer.
- Insuficiente hierro en la dieta diaria, especialmente proteínas y hierro de origen animal.
- Limitada absorción de hierro debido a las infecciones (principalmente diarreicas y respiratorias). Estas infecciones provocan además la pérdida de micronutrientes como el hierro y el zinc.

2.2.20. Clasificación de anemia

En niños y niñas de 6 a 59 meses la anemia se clasifica de la siguiente manera:

Anemia	Nivel de hemoglobina
Severa	< 7,0 g/dl
Moderada	7,0 – 9,9 g/dl
Leve	10,0 – 10,9 g/dl

Fuente: Instituto Nacional de Estadística e Informática

2.2.21. Estado nutricional

El estado nutricional de los niños (as) es empleado en el ámbito internacional como parte de los indicadores con los cuales se verifica el desarrollo de los países.



Los efectos del estado nutricional en los primeros años se prolongan a lo largo de la vida, ya que incrementa el riesgo de padecer enfermedades crónicas (sobrepeso, obesidad, diabetes, enfermedades cardiovasculares, entre otras) y está asociado a menores logros educativos.

Un estado nutricional deficiente tiene efectos adversos sobre el proceso de aprendizaje y rendimiento escolar, incrementa tanto la morbilidad como la mortalidad en la temprana infancia. La alimentación inadecuada y la prevalencia e incidencia de enfermedades infecciosas son determinantes en el crecimiento de las niñas y niños.

Los patrones de alimentación influyen en la nutrición de los niños (as) y a su vez, condicionan su potencial de desarrollo cognitivo y crecimiento. La lactancia materna provee los nutrientes adecuados y sirve de inmunización para un gran número de enfermedades comunes en la infancia. El inicio de la alimentación complementaria temprana limita lo valioso de lactancia materna, más cuando se hace sin las medidas de higiene y esterilización adecuadas, pues se pone a los niños (as) en contacto con sustancias contaminadas en el ambiente. Los malos hábitos de alimentación complementaria y lactancia están relacionados con el grado de nutrición, riesgos de enfermedad y muerte (Instituto Nacional de Estadística e Informática, 2018).

2.2.22. Evaluación del estado nutricional

Para determinar el estado nutricional del niño se puede realizar mediante las medidas antropométricas que son las más usadas en los servicios de salud por su utilidad y bajo costo.

Las medidas antropométricas que generalmente se usan es el peso, talla, circunferencia braquial y pliegues cutáneos. Los índices más usados en niños menores de 5 años son el peso para la talla, la talla para la edad y el peso para la edad.

2.2.23. Indicadores antropométricos

Según Ravasco et al. (2010) los indicadores antropométricos se definen:

Peso para la talla (P/T): Es la relación que puede existir entre el peso obtenido de un individuo de una talla determinada y el valor de referencia para su misma talla y sexo. Es más específico para el diagnóstico de desnutrición actual en niños de 2 a 10 años.

Talla para la edad (T/E): Es la relación entre la talla de un individuo y la referencia para su misma edad y sexo. Se emplea para el diagnóstico de desnutrición crónica.

Peso para la edad (P/E): Es la relación entre el peso de un individuo a una edad determinada y la referencia para su misma edad y sexo. Se utiliza para diagnosticar y cuantificar desnutrición actual o aguda.

La clasificación de los indicadores según el Centro Nacional de Alimentación y Nutrición (2007) son:

PESO PARA LA TALLA (mide la desnutrición aguda)

Clasificación	Peso
Desnutrición severa	< -3 DE
Desnutrición	≥ -3 DE
Normal	-2 DE a 2 DE
Sobrepeso	≤ 3 DE
Obesidad	> 3 DE

Fuente: Tabla de valoración nutricional antropométrica: niños < 5 años del CENAN.

TALLA PARA LA EDAD (mide la desnutrición crónica)



Clasificación	Longitud o talla
Talla baja severa	< -3 DE
Talla baja	\geq -3 DE
Normal	-2 DE a 2 DE
Talla alta	> 2 DE

Fuente: Tabla de valoración nutricional antropométrica: niños < 5 años del CENAN.

PESO PARA LA EDAD (mide la desnutrición global o desnutrición aguda y crónica)

Clasificación	Peso
Desnutrición	< -2 DE
Normal	-2 DE a 2 DE
Sobrepeso	> 2 DE

Fuente: Tabla de valoración nutricional antropométrica: niños < 5 años del CENAN.

2.2.24. Factores de riesgo asociados

Según el Ministerio de Salud (2016) podemos observar los siguientes factores:

Relacionados a la persona

- Recién nacidos prematuros y/o con bajo peso al nacer.
- Niñas y niños pequeños para la edad gestacional.
- Corte precoz del cordón umbilical.
- Niñas y niños menores de 2 años.



- Alimentación complementaria deficiente en productos de origen animal ricos en hierro.
- Niñas y niños con infecciones recurrentes.
- Niñas y niños menores de 6 meses sin lactancia materna exclusiva.
- Hijos de madres con embarazo múltiple.
- Hijos de madres adolescentes.
- Hijos de madres con periodo intergenésico corto.
- Hijos de madre anémica.

Relacionados al medio ambiente

- Zonas con alta inseguridad alimentaria.
- Zonas endémicas con parasitosis.
- Zonas endémicas de malaria.
- Zonas con saneamiento ambiental deficiente.
- Población expuesta a contaminación con metales pesados (plomo, mercurio, etc.).
- Familias con limitado acceso a información nutricional.

2.2.25. Factores biológicos del niño

Se consideró el peso, talla, edad y sexo en semanas según la fecha de atención identificándolo por el número de DNI y número de historia clínica para ello se extrajo la información del sistema SIEN de la red de salud Puno.

2.2.26. Factores sociodemográficos

Los factores sociodemográficos es un aspecto influyente en la salud de los niños y niñas, en la salud reproductiva, entre otros. Dentro de los factores sociodemográficos



encontramos: la edad de la madre, la condición civil (soltera, conviviente, casada, divorciada y separada), el nivel de instrucción (sin instrucción, primaria, secundaria, superior no universitaria y superior universitaria), el ingreso promedio mensual (menos de 1000.00 soles, entre 1000.00 a 3000.00 soles, entre 3000.00 a 5000.00 soles y más de 5000.00 soles) y número de hijos (uno, dos, tres, más de tres).

Edad de la madre: La edad de la madre es un aspecto importante, ser madre implica la responsabilidad de soporte económico del hogar. Un embarazo adolescente conlleva a una unión conyugal temprana, dejar los estudios, entre otros, pero un embarazo a edad adecuada conlleva a la formación de una familia más estable. La edad a la cual las mujeres comienzan su vida reproductiva constituye uno de los factores demográficos determinantes en la fecundidad de una población (Instituto Nacional de Estadística e Informática, 2021).

Condición civil: El estado civil de la madre puede ser un factor influyente en el estado nutricional del niño. Una familia estable con la cooperación de la pareja ayuda a la madre en su etapa de gestación, después del parto, y en la alimentación del niño. El niño puede tener una buena alimentación, considerando una estabilidad económica y nivel educativo de los padres (Torres Garcia, 2020).

Nivel de instrucción: El nivel de educación de las madres influye de manera importante en sus actitudes y prácticas relacionadas con la salud y el comportamiento reproductivo, sus actitudes hacia el tamaño ideal de la familia, la práctica de la planificación familiar y la violencia intrafamiliar. Por otro lado, el nivel de educación está positivamente asociado con la situación socioeconómica (Rivera Palomino, 2019).



Ingreso promedio mensual: Uno de los componentes claves de la calidad de vida de las personas y del desarrollo económico social de un país es el nivel educativo de su población (Torres Garcia, 2020).

Número de hijos: Un gran número de hijos en la familia implica que ambos padres trabajen fuera de casa, y esto trae consecuencia en el cuidado del niño, la falta de tiempo para cocinar y atender a los hijos, lo que puede provocar desnutrición en los hijos (Torres Garcia, 2020).

2.2.27. Factores gineco-obstétricos

Se consideraron el tipo de parto (normal, cesárea), anemia en etapa gestacional (con anemia, sin anemia), patología materna (pre eclampsia, enfermedad infecciosa, otros y ninguno).

Tipo de parto: El corte tardío del cordón umbilical favorece mayor transferencia de hierro al recién nacido y así reserva de hierro adicional (Ministerio de Salud, 2017).

Anemia en etapa gestacional: El estado nutricional de la madre es un factor importante en el peso del bebé al nacer y de la muerte perinatal e infantil, un adecuado estado nutricional tiene efecto directo sobre su propio bienestar y el de sus hijos, tanto al momento de nacer como en la vida adulta de ellos (Instituto Nacional de Estadística e Informática, 2021).

La suplementación a la madre con hierro y ácido fólico, control prenatal oportuno y alimentación balanceada favorece mejores reservas de hierro fetal (Ministerio de Salud, 2017).

Clasificación de anemia en mujeres gestantes

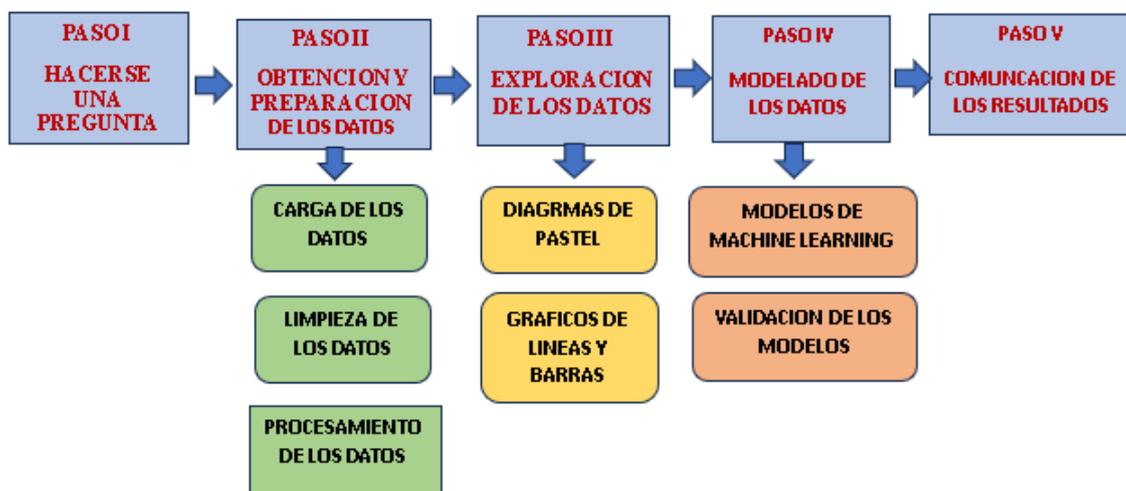
Mujeres gestantes y puérperas	Mujer gestante de 15 años a más	Mujer puérpera
Anemia severa	< 7,0	< 8,0
Anemia moderada	7,0 – 9,9	8,0 – 10,9
Anemia leve	10,0 – 10,9	11,0 – 11,9
Sin anemia	≥ 11,0	≥ 12,0

Fuente: Ministerio de Salud del Perú 2017.

Patología materna: Asiste a cualquier gestante con factores de riesgo materno, tanto desde la vertiente gestacional como desde el postparto. Esta en íntima conexión con la unidad de neonatología y se estructura sobre una serie de clínicas monográficas como la diabetes, hipertensión o infecciones de transmisión vertical (Vall d’Hebron Barcelona Hospital Campus, 2020).

2.2.28. Modelo aplicativo

Figura 1. Modelo aplicativo basado en la teoría de Data Science.



Elaboración: Propia.



PASO I. Pregúntate a ti mismo

Este paso presenta un problema a resolver, para lo cual se debe considerar el contexto del problema y los datos necesarios y adecuados, ya que los datos son el recurso más importante, necesario y fundamental para llevar a cabo a cabo un trabajo de previsión.

FASE II. Recolección y procesamiento de datos

En esta fase se buscan datos para pronósticos utilizando fuentes internas y externas. Esta investigación es importante porque, como se mencionó en el paso anterior, los datos son relevantes para el problema en el que desea trabajar. Tenga en cuenta que estos datos ayudarán a responder la pregunta. En segundo lugar, en este paso debe cargar los datos para los que necesita usar la biblioteca panda y el banco de trabajo Júpiter basado en el lenguaje de programación Python. Seguiremos con la preparación o limpieza de los datos, en esta parte por ejemplo se eliminarán los registros vacíos.

FASE III. Minería de datos El tercer paso es encontrar patrones en las relaciones entre los datos para que el analista tenga una idea clara de lo que los datos intentan mostrar. Cuando el analista utiliza gráficos específicos que ayudan a detectar patrones o tendencias.

PASO IV. Modelado de datos En esta fase, con los datos apropiados y procesados, se modela el comportamiento de los datos para futuras predicciones. En esta fase se utilizan modelos aleatorios de aprendizaje forestal y de regresión logística.

PASO V. Comunicar los resultados En esta etapa de obtener los resultados de los pronósticos a través de clasificadores, es necesario y fundamental presentar estos resultados de manera comprensible para las partes interesadas para lo cual esto se hace mediante gráficos son más adecuado para Presentación de los resultados.

python

Python en su versión para análisis de datos es casi muy completo porque ofrece un sistema de paquetes para análisis de datos y aprendizaje automático. Esto garantiza que se obtiene todo lo que pueda necesitar en el proceso de análisis de datos y, a veces, más. Gracias a su uso mínimo de memoria y su excelente administración de memoria, puede manejar grandes cantidades de datos en la memoria. Un recolector de basura en memoria generalmente ahorra días de carga, transformación, corte, corte, descarte o descarte de datos a través de varias iteraciones y manipulaciones de datos. Python puede integrar fácilmente diferentes herramientas y proporciona una base verdaderamente unificada para diferentes lenguajes, estrategias de datos y algoritmos de aprendizaje que se pueden integrar fácilmente y pueden ayudar a los científicos de datos a crear soluciones poderosas

2.2.29. Interpretación de los resultados

La interpretación de los resultados obtenidos se realiza a partir de la interpretación de los coeficientes del modelo. Para ello basta tener en cuenta que, si el modelo ajustado es bueno, entonces se dice que el modelo es significativo, pero además se debe analizar el grado de asociación estadística que existen en sus parámetros, a partir de la ecuación (4) se tiene:

$$\log \left[\frac{p(X_1, \dots, X_k; \beta)}{1 - p(X_1, \dots, X_k; \beta)} \right] = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$$

Donde el “odds ratio” que es el factor de riesgo está dado por la razón de esta expresión:

$$\frac{p(X_1, \dots, X_k; \beta)}{1 - p(X_1, \dots, X_k; \beta)} = e^{\beta_0 + \beta_1 X_1 + \dots + \beta_k X_k}$$

Entonces:

$$\frac{\frac{p(X_1 + 1, \dots, X_k; \beta)}{1 - p(X_1 + 1, \dots, X_k; \beta)}}{\frac{p(X_1, \dots, X_k; \beta)}{1 - p(X_1, \dots, X_k; \beta)}} = e^{\beta_1}$$

Por lo tanto, e^{β_1} es el factor de cambio en el “odds ratio” (OR) de riesgo si el valor de la variable X_1 cambia en una unidad. Así, se $\beta_1 > 0$ (ó $\beta_1 < 0$) el factor será mayor que 1 y $p(X_1, \dots, X_k; \beta)$ aumentará (disminuirá). Si $\beta_1 = 0$ la variable X_1 no ejerce ningún efecto sobre p_i .

β_0 es un ajuste de escala. Su mejor interpretación se obtiene calculando el valor de $p(X_1, \dots, X_k; \beta)$ en los valores medios de X_1, \dots, X_k y usar como variables explicativas sus valores estandarizados.

En regresión logística la media de asociación más empleada es el OR debido que el número e es la base de los logaritmos neperianos y elevados a un coeficiente de regresión logística del factor, si es mayor que 1 supone un aumento unitario, indica que el factor de riesgo es mayor.

Si el modelo de Regresión Logística es significativo y una de las variables independientes es dicotómica con valores de 0 y 1, el número e elevado al coeficiente de regresión logística es el OR, denominado factor de riesgo o protección que implica un aumento unitario de la variable independiente. En el caso de una variable cuantitativa, e elevado a β_1 es el número de veces que aumenta la probabilidad de padecer una enfermedad por cada unidad de aumento de la variable independiente, o dicha de otra manera, cuantas veces es más probable que padezca la enfermedad una persona que presenta síntomas relacionadas a ella (Flores Manrique, 2002).



2.2.30. Marco conceptual

Machine learning o aprendizaje automático.

Consiste en programas de codificación que ajustan automáticamente su funcionamiento en base a la comparación con la información contenida en los datos. El aprendizaje se realiza mediante un modelo parametrizado cuyas variables se ajustan automáticamente a diferentes criterios (Iguar & Seguí, 2017).

Aprendizaje no supervisado.

Es un tipo de aprendizaje automático que involucra algoritmos que aprenden de un conjunto de datos sin etiquetar que sirven como entrenamiento. Esta ciencia se utiliza para examinar datos según determinados modelos estadísticos, geométricos o criterios similares (Iguar & Seguí, 2017).

Aprendizaje supervisado.

Es un tipo de aprendizaje automático que involucra algoritmos que aprenden de un conjunto de datos etiquetados, utilizados como ejemplos de entrenamiento, para generalizar a un conjunto diferente de datos de salida (Iguar y Seguí, 2017).

Clasificación:

un método para predecir una respuesta categórica, lo que significa que la respuesta solo tiene un número finito de opciones. Ejemplos: grado de cáncer, preguntas correctas o incorrectas, predicción del año de nacimiento.



Ciencia de datos.

Es el arte de adquirir conocimiento a través de datos. Esto está relacionado con cómo se utilizan los datos depara generar conocimientos que se utilizan para tomar decisiones, predecir el futuro y/o comprender el pasado/presente (Ozdemir, 2016).

Júpiter.

Un entorno de desarrollo integrado, originalmente llamado IPython, que se limita a trabajar solo con Python. El libro de trabajo de IPython se separó del software de IPython y todos sus componentes heredados se importaron a Jupyter (Igual & Seguí, 2017).

Regresión logística

La regresión logística es una técnica de análisis multivariante, en la que la variable dependiente o variable respuesta es una variable dicotómica y la variable o variables independientes pueden ser cualitativas o cuantitativas. Si en el modelo la variable independiente es cualitativa con H categorías, habrá que generar H-1 variables DUMMY, a fin de que todas las posibilidades de la variable queden debidamente representadas en el modelo. Con los modelos de regresión logística podemos investigar factores causales de una determinada característica de la población y también estudiar qué factores modifican la probabilidad de un suceso determinado (Álvarez Cáceres, 1995).

Anemia

Tambien conocida como nivel de hemoglobina baja, es una afección en la cual careces de suficientes globulos rojos sanos para transportar un nivel adecuado de oxígeno a los tejidos del cuerpo. La anemia puede ser temporal o prolongada y puede variar de leve a severa (Mayo clinic, 2022).



Hemoglobina

Es una proteína al interior de los glóbulos rojos rica en hierro, cada proteína de la hemoglobina transporta cuatro moléculas de oxígeno y estas se distribuyen en todo el cuerpo mediante los glóbulos rojos, también transporta dióxido de carbono de vuelta a los pulmones.

Estado nutricional

El estado nutricional es la condición física de una persona, el resultante final del balance entre ingesta y requerimiento de nutrientes.



CAPÍTULO III

MATERIALES Y MÉTODOS

3.1. LUGAR DE ESTUDIO

El trabajo de investigación se llevó a cabo en el Centro de Salud I-3 Vallecito de la ciudad de Puno identificado con condigo RENIPRESS 3255 de la Microred Metropolitano de la red de salud Puno con ubicación geográfica de coordenadas (-15.82941167, -70.020545) perteneciente al distrito, provincia y departamento de Puno.

3.2. TIPO Y DISEÑO DE INVESTIGACIÓN

Tipo de investigación

La investigación corresponde al enfoque cuantitativo descriptivo correlacional, descriptivo porque se explica las características del fenómeno, es decir, como es y cómo se manifiesta, correlacional porque en esta investigación se analiza y estudia la relación de los factores en la prevalencia de anemia en menores de 3 a 59 meses del Centro de Salud I-3 Vallecito – Puno 2021.

Esta investigación tiene un enfoque aplicado, ya que utiliza Machine Learning en un proceso convencional para realizar un modelo logit para los factores influyentes de la anemia. Según Hernández et al. (2014), la investigación aplicada tiene como objetivo abordar de manera práctica problemas reales y solucionar deficiencias identificadas en un determinado ámbito de desarrollo.



Diseño de investigación

El diseño de investigación es no experimental de corte transversal, es no experimental porque no se manipula las variables deliberadamente y se observa el fenómeno en su ambiente natural para su análisis, es de corte transversal porque los datos se recolectan en un tiempo único.

El diseño no experimental es una metodología de investigación que se basa en la observación y recopilación de datos sin la manipulación de variables independientes. Se utiliza cuando no es posible realizar experimentos controlados. A través de técnicas como estudios de caso, encuestas y análisis de contenido, el diseño no experimental permite describir y comprender fenómenos en su entorno natural (Johnson & Smith, 2018).

3.3. POBLACIÓN Y MUESTRA

Población

La población de estudio estuvo comprendida por menores de 3 a 59 meses del Centro de Salud I-3 Vallecito – Puno que fueron atendidos durante todo el año 2021.

Según Boddy (2016, p. 4321), se define el término "población" como la agrupación de objetos o personas con el propósito de recopilar datos relevantes para una investigación. Esta población puede adoptar diferentes formas y tipos, desde registros hasta animales.

Muestra (muestreo)

El método muestral utilizado fue la totalidad de los niños atendidos durante el año 2021.



Unidad muestral

La unidad muestral para esta investigación son los menores de 3 a 59 meses del Centro de Salud I-3 Vallecito – Puno 2021.

3.4. TÉCNICAS E INSTRUMENTOS DE RECOLECCIÓN DE DATOS

Para determinar los factores asociados en la prevalencia de anemia en menores de 3 a 59 meses:

Técnica

La técnica usada en esta investigación fue la encuesta, técnica que nos permitió recopilar los datos de los menores de 3 a 59 meses del Centro de Salud I-3 Vallecito – Puno 2021.

El cuestionario es ampliamente reconocido en las ciencias sociales como un método tradicional y efectivo para recopilar y registrar datos, según lo señalado por García (2003, p.2). Su versatilidad le permite desempeñar un papel crucial tanto como herramienta de investigación como en la evaluación de personas, procesos y programas de formación. Además, Aigner (2009, p.6) describe el cuestionario como un formulario que contiene una serie de preguntas estandarizadas y estructuradas, las cuales se plantean de manera uniforme a todos los participantes.

Instrumento

Se utilizaron 02 instrumentos uno fue el cuestionario para la recolección de características de la madre y otro fue el análisis del registro del Sistema de Información del Estado Nutricional (SIEN) CRED de la Red de Salud Puno.

El instrumento usado en este trabajo fue el cuestionario, este instrumento se elaboró de acuerdo a las variables de estudio, cuenta con 12 ítems el cual nos permitió



recopilar la información de los factores asociados en la prevalencia de anemia en menores de 3 a 59 meses del Centro de Salud I-3 Vallecito – Puno 2021. (Anexo 03)

Procedimiento

Para la recopilación de datos sobre los factores se realizó el siguiente procedimiento:

- La encuesta fue realizado solo a las madres que asistieron al centro de salud I-3 Vallecito Puno.
- La encuesta se realizó de forma manual para cada una de las madres de menores de 3 a 59 meses.
- Se explicó el propósito de la encuesta y se dio a conocer la estructura del instrumento para su correcto llenado.
- Se estuvo al tanto sobre las dudas del llenado que pudieran tener las madres.
- Los datos fueron procesados después de haber culminado las encuestas.

Para determinar el nivel de anemia en menores de 3 a 59 meses:

Técnica

La técnica usada en esta investigación fue la observación directa de los datos del Sistema de Información del Estado Nutricional (SIEN) CRED del centro de salud, que contienen los datos de los menores de 3 a 59 meses del Centro de Salud I-3 Vallecito – Puno 2021.

Instrumento

El segundo instrumento utilizado fue el registro de datos (ver Anexo 04) facilitado por el Sistema de Información del Estado Nutricional (SIEN) del niño menor de cinco



años que acceden al Centro de Salud de la Red de Salud de Puno de todo el año 2021, donde la información fue brindada solo del año 2021 y del Centro de salud I-3 Vallecito.

3.5. MÉTODO DE ANÁLISIS DE DATOS

Los datos recopilados en esta investigación fueron procesados considerando la estadística descriptiva para luego realizar un análisis correlacional. Se ha realizado un sistema de seriación en el instrumento de recolección de datos para poder identificarlos fácilmente, se codificó las variables en números en el software Excel para luego ser procesados en el software estadístico. Con la regresión logística se demostró el modelo logit óptimo para factores asociados en la prevalencia de anemia en menores de 3 a 59 meses. Los resultados se mostraron en tablas estadísticas debidamente analizados e interpretados.

Operacionalización de variables

VARIABLE	DIMENSIONES	INDICADORES	CATEGORÍA	ESCALA	TIPO DE VARIABLE
VARIABLE INDEPENDIENTE FACTORES ASOCIADOS	Factores Biológicos	Sexo	Femenino Masculino	1 2	Discreta
		Edad	3-59 meses	Intervalo	Continua
		Peso	En kg	Intervalo	Continua
		Talla	En cm	Intervalo	Continua
	Factores De Aplicación SIEN	N_CCred	Numérico	Intervalo	
		Sachets	Si No	1 2	Ordinal
		Consejerías	Si No	1 2	Ordinal
		Sesiones	Si No	1 2	Ordinal
		Visitas	Si No	1 2	Ordinal
	Factores sociodemográficos de madre	Condición civil	Soltera Conviviente Casada Divorciada Viuda	1 2 3 4 5	Ordinal
		Nivel de instrucción	Sin instrucción Primaria Secundaria Superior No Univ. Superior Univer.	1 2 3 4	Ordinal
		Edad Madre	Numérico	De intervalo	Discreta
		Número de hijos	Uno Dos Tres Más de tres	1 2 3 4	Ordinal
		SIS	SI NO	1 2	Nominal
		Ingreso	Menos de 1000 De 1000 – 3000 De 3000 – 5000 Mas de 5000	1 2 3 4	Ordinal
		Factores gineco-obstétricos	Tipo de parto	Normal Cesárea	1 2
	Anemia en etapa gestacional		Con anemia Sin anemia	1 2	Nominal
	Patología materna		Pre eclampsia Enfermedad infecciosa Otra Ninguna	1 2 3 4	Nominal
	Semanas de embarazo		Numérico	De intervalo	Discreta
	VARIABLE DEPENDIENTE ANEMIA	Prevalencia de anemia	Nivel de anemia	Si No	1 2

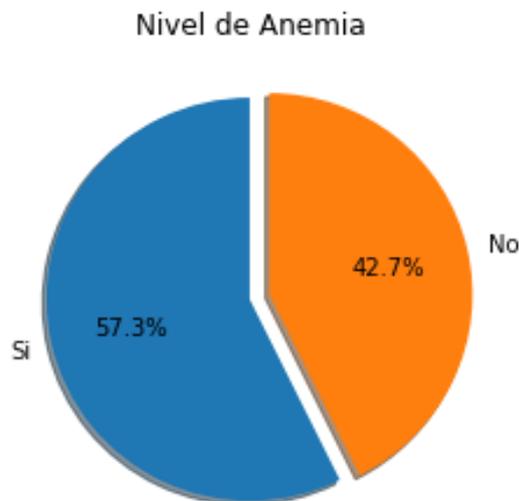
CAPÍTULO IV

RESULTADOS Y DISCUSIÓN

Luego de consolidar toda la información, con el apoyo de las librerías de Python se corroboró el tipo de datos, el cómo están distribuidos donde esta etapa está comprendida como limpieza de datos reacomodando y juntando la base de datos que se tenía separados; llegó el momento de la exploración y la interpretación de los datos.

4.1. RESULTADOS

Figura 2. Nivel de Anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.

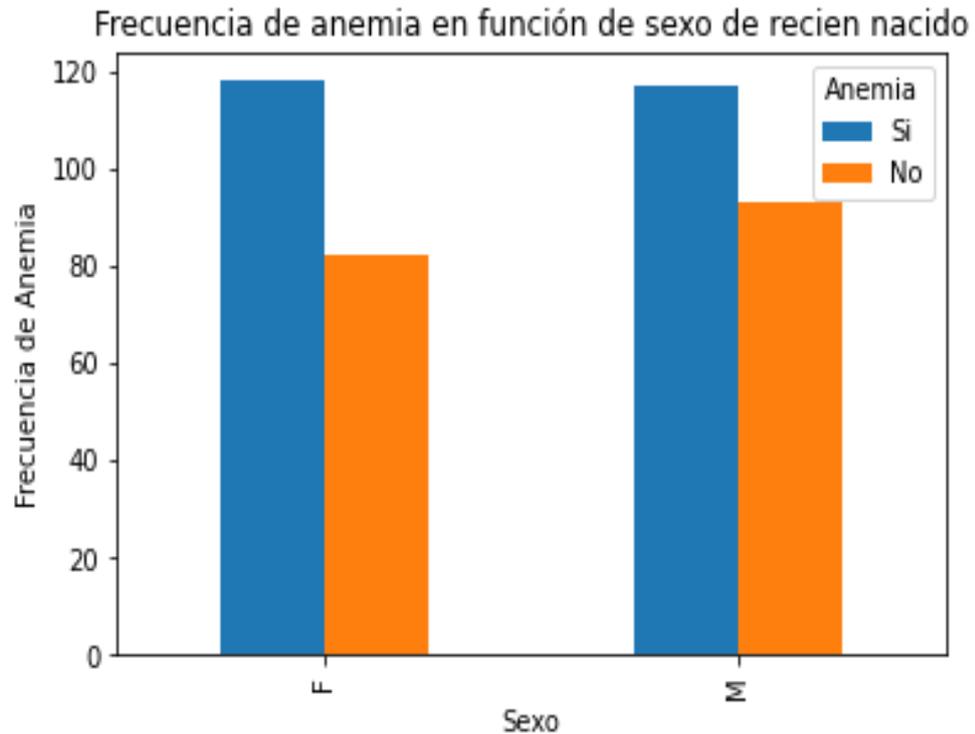


Fuente: Elaboración propia en base a los datos.

Del grafico anterior se evidencia que del 100% de los menores de 3 a 59 meses que asistieron en algún momento del año 2021 al centro de salud I-3 Vallecito-Puno, el 57.3% tuvieron clasificación como niños con anemia, es decir esos niños se

encontraban por debajo de 11g/dl mientras que el 42.7% no tuvieron anemia o no tuvieron un nivel de hemoglobina igual o mayor que 11g/dl.

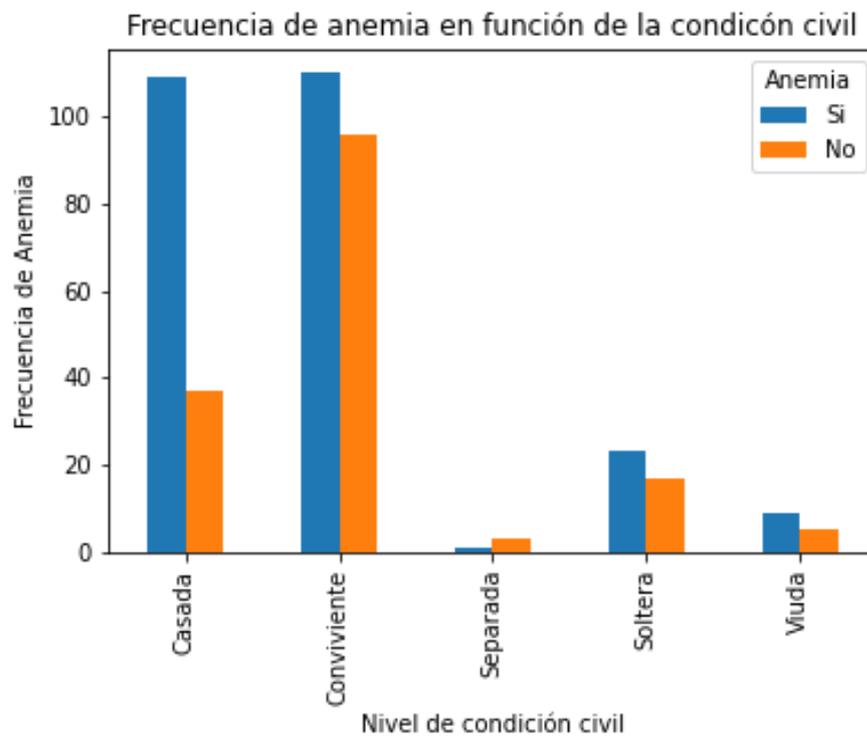
Figura 3. Nivel de Anemia según el Sexo en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.



Fuente: Elaboración propia en base a los datos.

De la totalidad de niños, se evidencia que tanto en el sexo masculino como en el sexo femenino el nivel de anemia fue en mayor frecuencia, respecto a los que no tuvieron anemia en menores de 3 a 59 meses del centro de salud I-3 Vallecito-Puno-2021.

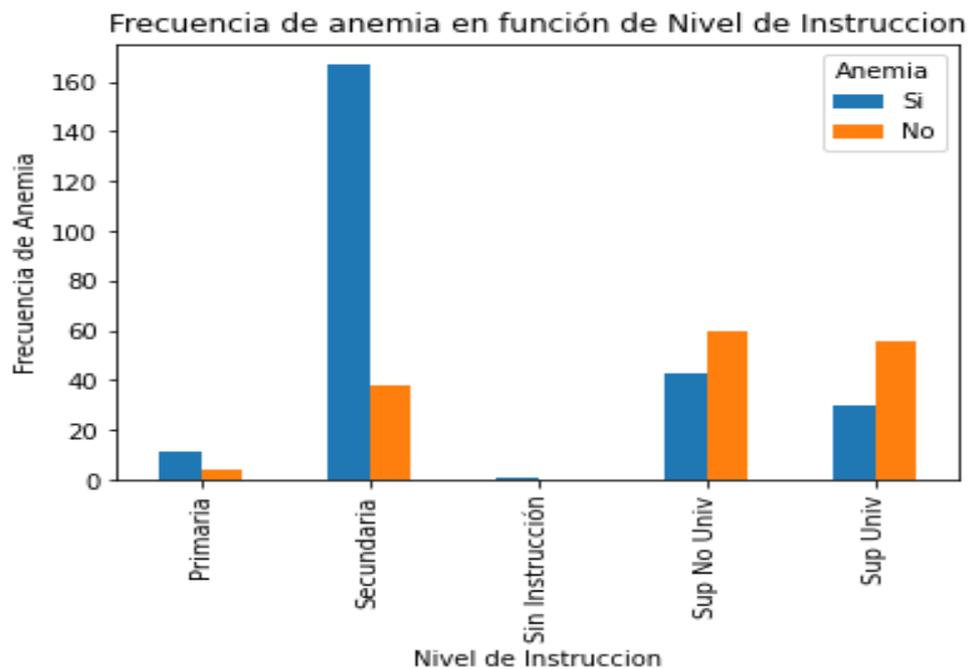
Figura 4. Nivel de Anemia en menores de 3 a 59 meses según el Estado Civil de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.



Fuente: Elaboración propia en base a los datos.

Del gráfico anterior se verifica que la mayor parte de las madres encuestadas fueron casadas o convivientes donde en las madres que están casadas la diferencia es significativa con respecto a sus hijos si tienen anemia contra los que no tienen, mientras que en las madres convivientes esta diferencia disminuye casi al punto de igualarse entre los niños con anemia y sin anemia. Seguidamente también se evidencia que en menor proporción estuvieron las madres separadas y la mayoría de ellas tuvieron hijos con anemia, finalmente entre las madres solteras y viudas no se evidencia diferencias de proporciones en cuanto a sus hijos con anemia o sin anemia.

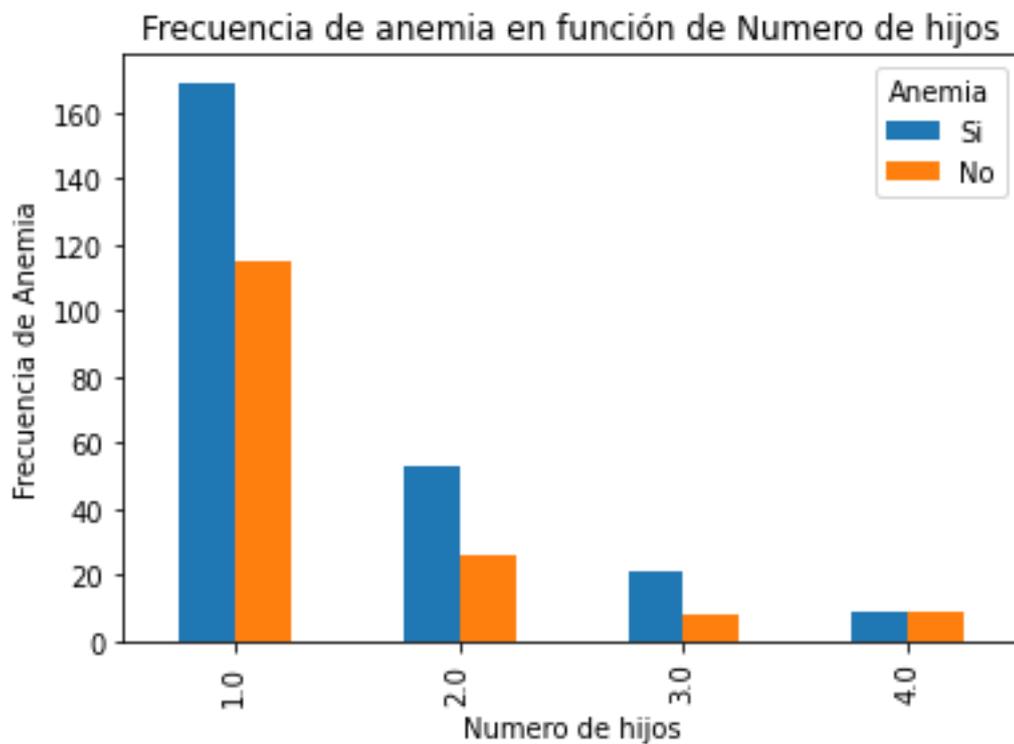
Figura 5. Nivel de Anemia en menores de 3 a 59 meses según el Nivel de Instrucción de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.



Fuente: Elaboración propia en base a los datos.

De las madres encuestadas, se identificó que la mayoría de ellas tuvieron educación secundaria donde una gran proporción de estas madres si tuvieron a sus hijos con anemia, mientras que en las madres con educación superior no universitaria y educación superior universitaria tuvieron a sus hijos sin anemia dándonos a entender que la educación de la madre si está influyendo en el nivel de anemia de los menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

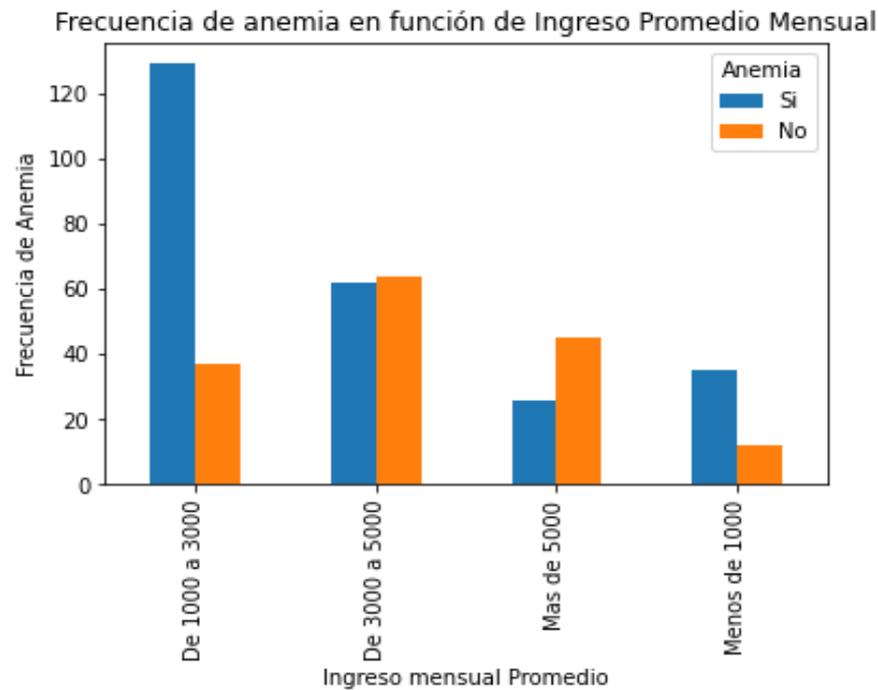
Figura 6. Nivel de Anemia en menores de 3 a 59 meses según el Número de hijos de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.



Fuente: Elaboración propia en base a los datos.

De la totalidad de las madres estudiadas en el 2021, según el grafico no hay una asociación entre la cantidad de hijos de la madre y el nivel de anemia, asimismo cerca de 160 madres atendidas tuvieron 01 hijo hasta esa fecha y que estuvieron con anemia, mientras que cerca de 120 niños no tuvieron anemia y fueron hijos únicos.

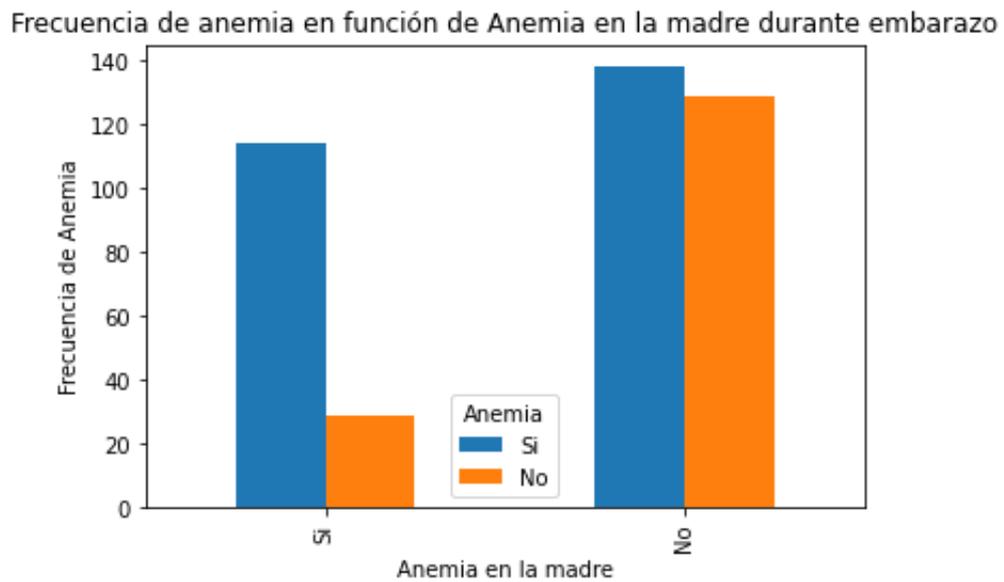
Figura 7. Nivel de Anemia en menores de 3 a 59 meses según Ingreso mensual de la madre atendidos en el centro de salud I-3 Vallecito-Puno-2021.



Fuente: Elaboración propia en base a los datos.

Del grafico anterior se muestra que más de 120 madres afirmaron tener un promedio total de ingreso mensual entre 1000 y 3000 soles y que sus hijos tuvieron anemia, mientras que los que tuvieron ingresos promedio mensual de 3000 a 5000 soles no se evidencio una diferencia significativa con respecto al nivel de anemia de sus hijos. Seguidamente en el grupo de madres que tuvieron ingresos de más de 5000 soles se evidencia que la mayoría de los menores analizados no tuvieron anemia, pero en los que perciben ingresos mensuales promedio de menos del 1000 sol existen que si tuvieron anemia en su mayoría.

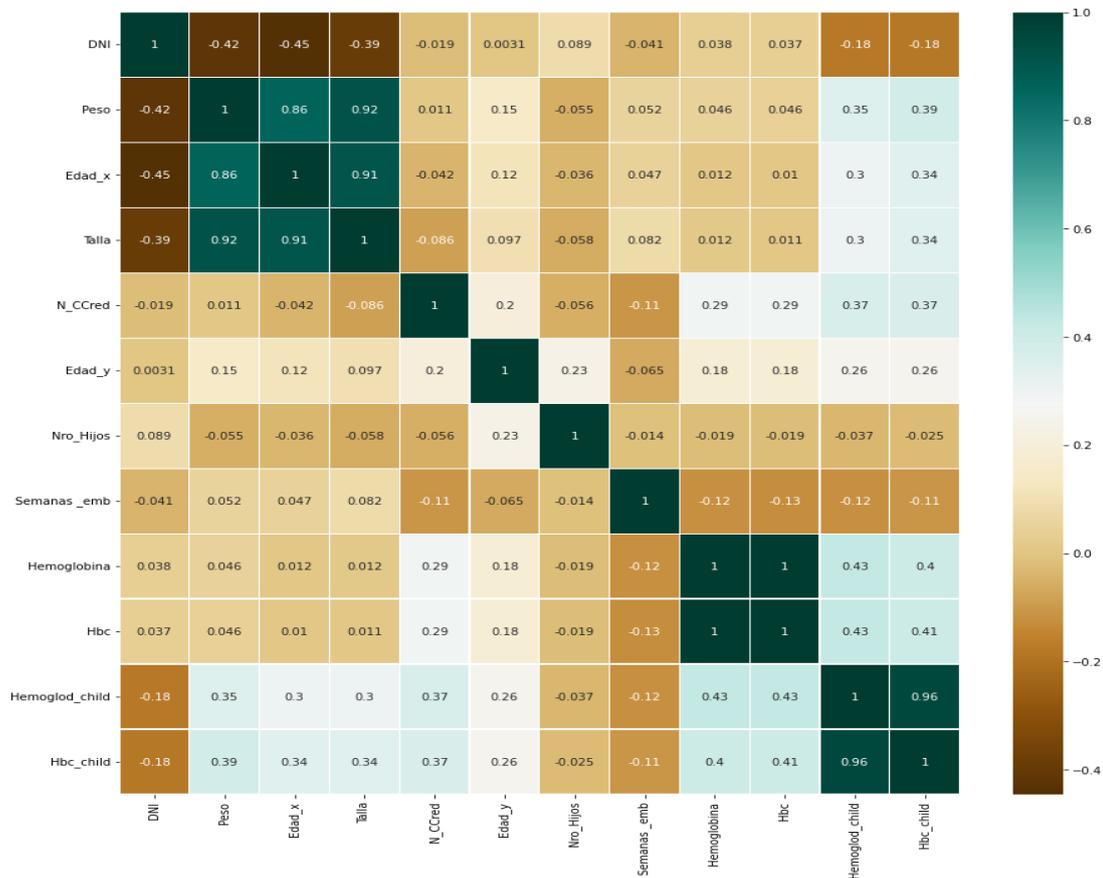
Figura 8. Nivel de Anemia en menores de 3 a 59 meses según Anemia durante el embarazo de las madres atendidos en el centro de salud I-3 Vallecito-Puno-2021.



Fuente: Elaboración propia en base a los datos.

De las madres que, si tuvieron anemia, la mayoría de sus hijos analizados si tuvieron también anemia, mientras que en grupo de madres que no tuvieron anemia no existe una diferencia significativa entre los menores que si hubo o no anemia en el año 2021 del Centro de salud I-3 Vallecito.

Figura 9. Correlaciones según el nivel de hemoglobina y a las variables numéricas en los menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.



Fuente: Elaboración propia en base a los datos.

Del gráfico anterior se evidencia las correlaciones entre todas las variables numéricas existentes, donde el interés radica en la variable “Hbc_child” que es el nivel de hemoglobina corregido de los menores donde se puede evidenciar que tiene un 0.39 de correlación con el peso, seguido del 0.34 con la edad del niño, 0.34 con la talla, 0.37 con el número de controles de crecimiento, 0.26 con la edad de la madre, 0.025 con el número de hijos, 0.11 con las semanas de embarazo, 0.41 con el nivel de hemoglobina corregido de la madre durante el embarazo. Donde estas correlaciones indican que son regulares, pero existen un nivel de correlación.

REALIZANDO MODELADO DE DATOS.

Para el modelamiento de los datos se presentará a continuación cinco modelos diferentes que convergieron con métricas significativas durante la prueba de diferentes modelos y la configuración de diferentes parámetros.

MODELO LOGIT- 0

Para este modelo en primer lugar se utilizaron la técnica de selección de variables con selección de características a priori, es decir que el investigador al explorar los datos verifica y propone la cantidad de características del modelo para la cual al convertir todo el data consolidado incluido las variables dummy, se tuvo 38 variables de los cuales se empezó a buscar desde 30 variables para responder al modelo donde no se tuvo convergencia, hasta llegar a la propuesta de 16 variables, proponiendo así encontrar las 16 variables más significativas antes de hacer el respectivo entrenamiento del modelo.

Tabla 1. Resumen de modelo 00 con 15 variables significativas para la predicción de anemia en menore de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

Logit	Regression	Results
Dep. Variable: Anemia	No. Observations:	287
Model: Logit	Df Residuals:	271
Method: MLE	Df Model:	15
Date: Sat, 14 Jan 2023	Pseudo R-squ.:	0.8171
Time: 00:33:25	Log-Likelihood:	-35.490
converged: True	LL-Null:	-194.01
Covariance Type:nonrobust	LLR p-value:	1.590e-58

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior se evidencia que el resultado del modelo denominado “modelo 0”, se verifica el nombre de la variable dependiente ANEMIA donde se presenta el resumen del modelo entrenando 287 datos de las 410, esta cantidad significó el 70% de entrenamiento y 30% de testing, logrando el modelo en general un

R2 de 81.71% el cual indica la flexibilidad de modelo, Criterio de información Bayesiana (BIC) y Derivación alternativa del índice (AIC). La cantidad de variables entrenada es de 16 menos uno. También es importante mencionar que el modelo en su conjunto converge lo cual nos indica que en 10 iteraciones del método de Newton Raphson se encontró una solución, luego está el residuo del R cuadrado con un 0.091 lo cual indica

Tabla 2. Coeficientes de modelo 00 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	coef	std err	z	P> z	[0.025	0.975]
const	13.1618	13.231	0.995	0.320	-12.771	39.095
Peso	-0.7722	0.187	-4.131	0.000	-1.139	-0.406
cat_Sexo_M	2.0488	0.830	2.469	0.014	0.422	3.676
N_CCred	-0.8522	0.185	-4.594	0.000	-1.216	-0.489
cat_N_Sachets_0	1.6151	0.694	2.328	0.020	0.255	2.975
cat_N_Sesiones_0	-3.3580	11.043	-0.304	0.761	-25.002	18.286
cat_Condicion_Civil_Casada	1.5747	0.734	2.145	0.032	0.136	3.014
cat_Nivel_Instrucción_Secundaria	0.2373	6.824	0.035	0.972	-13.137	13.611
cat_Nivel_Instrucción_Sup No Univ	-2.7530	6.897	-0.399	0.690	-16.271	10.765
cat_Nivel_Instrucción_Sup Univ	-2.4484	6.899	-0.355	0.723	-15.971	11.074
cat_Ingreso_De 1000 a 3000	0.3578	2.858	0.125	0.900	-5.244	5.959
cat_Ingreso_De 3000 a 5000	-1.1053	2.979	-0.371	0.711	-6.943	4.733
cat_Ingreso_Mas de 5000	-1.7102	3.018	-0.567	0.571	-7.625	4.205
cat_Patologia_Materna_Otra	1.3110	0.990	1.324	0.186	-0.630	3.252
cat_Patologia_Materna_Ninguna	-0.3265	0.729	-0.448	0.654	-1.755	1.102
cat_Anemia_mother_Si	5.4824	1.022	5.365	0.000	3.480	7.485

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior encontramos que, de las 16 variables propuestas para entrenamiento del modelo, se presentan sus respectivos coeficientes detallado por cada una de ellas, mencionado que la columna Std.Err. es el error estándar del coeficiente que estima la variabilidad entre las estimaciones del coeficiente que se obtendrían si se toman los datos una y otra vez. De la misma forma el valor z es el equivalente a la prueba estadística de Wald donde es un estadístico de prueba que mide la relación entre el coeficiente y su error estándar, seguidamente nos presenta la columna del valor probabilístico P, esto nos indica que aquellos coeficientes menores que 0.05, son significativos y las siguientes columnas son los intervalos de confianza.

Entonces, del modelo 0 las variables significativas para formar el modelo vienen comprendidas por el peso ($\text{sig}=0.000$), categoría de sexo masculino ($\text{sig}=0.014$), el número de consultas CRED ($\text{sig}=0.00$), si el no consume sachet ($\text{sig}=0.000$), la categoría de condición civil casada con ($\text{sig}=0.032$), y si la madre tuvo anemia durante la etapa gestacional ($\text{sig}=0.000$), el resto de las 16 variables no mencionadas fueron no significativas ya que su nivel “p” fue mayor que 0.05.

Tabla 3. Matriz de confusión con el modelo 00 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Predicción 0	1
Real		
0	52	6
1	4	61

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior se obtiene la confusión de matriz con el modelo 0 donde se valida con el grupo de datos de testing y se comprueba que de los 123 datos separados para test, realizando una tabla cruzada entre los valores reales de la variable Anemia y los proyectados con un umbral de 0.7, se encuentra que 52 datos fueron predichos como Verdaderos Negativos, seguido de 61 datos que fueron Verdaderos positivos, en seguida 6 datos fueron falsos positivos, y 4 datos clasificados como falso negativos.

Tabla 4. Métricas de evaluación con el modelo 00 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Métricas	Valores
0	sensibilidad	0.938462
1	especificidad	0.103448
2	precisión	0.910448
3	Accuracy	0.918699

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior tenemos las métricas de evaluación del modelo 0 donde se obtuvo una sensibilidad de 0.9684, especificidad de 0.1034, precisión del modelo fue de 0.9104 y exactitud del modelo fue de 0.9187.

MODELO LOGIT- 01

De las 38 variables que se tiene en el data consolidada junto con las variables dummy para este modelo se propuso solo 13 variables a encontrar los más significativos incluyendo una constante donde a continuación se presentan los resultados.

Tabla 5. Resumen de modelo 01 con 13 variables significativas para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

```
Logit Regression Results
=====
Dep. Variable:          Anemia      No. Observations:      287
Model:                 Logit       Df Residuals:          273
Method:                MLE        Df Model:              13
Date:                  Sat, 14 Jan 2023   Pseudo R-squ.:         0.8162
Time:                  00:44:26    Log-Likelihood:        -35.665
converged:             True       LL-Null:               -194.01
Covariance Type:      nonrobust   LLR p-value:           7.674e-60
=====
```

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior se evidencia que el resultado del modelo denominado “modelo 1”, se verifica el nombre de la variable dependiente ANEMIA donde se presenta el resumen del modelo entrenando 287 datos de las 410, teniendo 273 grados de libertad, es decir 13 variables están siendo analizadas en el presente modelo, logrando el modelo en general un R2 de 81.62% el cual indica la flexibilidad de modelo. La cantidad de variables entrenada es de 13. También es importante mencionar que el modelo en su conjunto converge lo cual nos indica que en 10 iteraciones del método de Newton Raphson se encontró una solución.

Tabla 6. Coeficientes de modelo 01 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	coef	std err	z	P> z	[0.025	0.975]
const	11.9918	8.637	1.388	0.165	-4.937	28.921
Peso	-0.7779	0.190	-4.105	0.000	-1.149	-0.406
cat_Sexo_F	-2.0649	0.824	-2.505	0.012	-3.680	-0.449
N_Ccred	-0.8686	0.186	-4.670	0.000	-1.233	-0.504
cat_N_Sachets_0	1.6037	0.695	2.307	0.021	0.241	2.966
cat_Condicion_Civil_Casada	1.5483	0.730	2.120	0.034	0.117	2.979
cat_Nivel_Instrucción_Secundaria	0.4328	7.806	0.055	0.956	-14.867	15.733
cat_Nivel_Instrucción_Sup_Univ	-2.3022	7.868	-0.293	0.770	-17.724	13.119
cat_Nivel_Instrucción_Sup_No_Univ	-2.6416	7.867	-0.336	0.737	-18.060	12.776
cat_Ingreso_De_1000_a_3000	0.0758	2.308	0.033	0.974	-4.447	4.599
cat_Ingreso_De_3000_a_5000	-1.3931	2.464	-0.565	0.572	-6.222	3.435
cat_Ingreso_Mas_de_5000	-2.0009	2.520	-0.794	0.427	-6.940	2.938
cat_Patologia_Materna_Otra	1.4516	0.957	1.517	0.129	-0.424	3.327
cat_Anemia_mother_Si	5.5666	1.018	5.469	0.000	3.571	7.562

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior encontramos que, de las 13 variables propuestas para entrenamiento del modelo, se presentan sus respectivos coeficientes detallado por cada una de ellas, mencionando que la columna Std.Err. es el error estándar del coeficiente que estima la variabilidad entre las estimaciones del coeficiente que se obtendrían si se toman los datos una y otra vez. Asimismo, se presenta los coeficientes por cada una de las variables y el valor z es el equivalente a la prueba estadística de Wald donde es un estadístico de prueba que mide la relación entre el coeficiente y su error estándar, seguidamente nos presenta la columna del valor probabilístico P, esto nos indica que

aquellos coeficientes menores que 0.05, son significativos y las siguientes columnas son los intervalos de confianza para cada uno de los coeficientes del modelo.

Entonces, del modelo 01 tiene las variables significativas para formar el modelo que vienen comprendidas por el peso (sig=0.000), categoría de sexo femenino (sig=0.012), el número de consultas CRED (sig=0.00), si el niño hasta ese momento no consume sachet (sig=0.021), la categoría de condición civil casada con (sig=0.034), y si la madre tuvo anemia durante la etapa gestacional (sig=0.000), el resto de las 16 variables no mencionadas fueron no significativas ya que su nivel “p” fue mayor que 0.05.

Tabla 7. Matriz de confusión con el modelo 01 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Predicción 0	1
Real		
0	53	5
1	5	60

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior se obtiene la confusión de matriz con el modelo 01 donde se valida con el grupo de datos de testing y se comprueba que de los 123 datos separados para test, realizando una tabla cruzada entre los valores reales de la variable Anemia y los proyectados con un umbral de 0.75 encontrándose como este umbral las mejores métricas se obtiene que 53 datos fueron predichos como Verdaderos Negativos, seguido de 60 datos que fueron Verdaderos positivos, en seguida 5 datos fueron falsos positivos, y 5 datos clasificados como falso negativos.

Tabla 8. Métricas de evaluación con el modelo 01 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Métricas	Valores
0	Sensibilidad	0.923077
1	Especificidad	0.086207
2	Precision	0.923077
3	Accuracy	0.918699

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior tenemos las métricas de evaluación del modelo 01 donde se obtuvo una sensibilidad de 0.9230, especificidad de 0.0862, precisión del modelo fue de 0.9230 y exactitud del modelo fue de 0.91869.

MODELO LOGIT-02

Para este modelo se planteó escoger proponiendo solo 12 variables, pero al no presentar convergencia entre las variables seleccionadas a priori se determinó plantear el entrenamiento con 10 variables más una constante del total de las 38 variables que contenía el data, para lo cual los resultados se detallan a continuación.



Tabla 9. Resumen de modelo 02 con 10 variables significativas para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

Optimization terminated successfully.
Current function value: 0.186676
Iterations 8

Logit Regression Results			
=====			
==			
Dep. Variable:	Anemia	No. Observations:	287
Model:	Logit	Df Residuals:	276
Method:	MLE	Df Model:	10
Date:	Sat, 14 Jan 2023	Pseudo R-squ.:	0.7239
Time:	01:05:44	Log-Likelihood:	-53.576
converged:	True	LL-Null:	-194.01
Covariance Type:	nonrobust	LLR p-value:	1.706e-54

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior evidencia que el modelo converge a las 8 iteraciones donde al igual que en los anteriores modelos presentados, se entrenó con el 70% de la data y el 30% se dejó para testing o validación, donde se tuvo 276 grados de libertad del residuo, por lo cual se tiene en este modelo denomiado “modelo 02” 10 variables como se presenta en la tabla, asimismo, el modelo en general presentó un R2 de 72.39% el cual indica hasta el momento el menor R3 de que los anteriores modelos.

Tabla 10. Coeficientes de modelo 02 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	coef	std err	z	P> z	[0.025	0.975]

const	-2.1948	1.621	-1.354	0.176	-5.371	0.981
N_CCred	-0.5528	0.107	-5.166	0.000	-0.763	-
0.343						
cat_N_Sachets_0	1.7972	0.558	3.223	0.001	0.704	2.890
cat_Condicion_Civil_Casada	1.6715	0.614	2.721	0.007	0.468	2.875
cat_Nivel_Instrucción_Secundaria	1.9647	1.416	1.388	0.165	-0.810	4.739
cat_Nivel_Instrucción_Sup No Univ	0.6060	1.495	0.405	0.685	-2.323	3.535
cat_Nivel_Instrucción_Sup Univ	1.0355	1.585	0.653	0.514	-2.072	4.143
cat_Ingreso_De 1000 a 3000	1.7586	0.968	1.817	0.069	-0.138	3.655
cat_Ingreso_De 3000 a 5000	-0.2165	1.142	-0.190	0.850	-2.454	2.021
cat_Ingreso_Mas de 5000	-0.7611	1.284	-0.593	0.553	-3.279	1.756
cat_Anemia_mother_Si	4.0167	0.624	6.433	0.000	2.793	5.240

Fuente: Elaboración propia a partir de los datos.

Describiendo los resultados presentados en la tabla anterior, de las 10 variables propuestas para entrenamiento del “modelo 02”, se presentan sus respectivos coeficientes detallado por cada una de ellas en la primera columna denominado coef., mencionando que la columna Std.Err. es el error estándar del coeficiente que estima la variabilidad entre las estimaciones del coeficiente que se obtendrían si se toman los datos una y otra vez. Asimismo, se presenta los coeficientes por cada una de las variables y el valor z es el equivalente a la prueba estadística de Wald menores que 0.05, son significativos y las siguientes columnas son los intervalos de confianza para cada uno de los coeficientes del modelo.

Entonces, del modelo 02 tiene las variables significativas para formar el modelo que vienen comprendidas el número de consultas CRED (sig=0.000), si el niño hasta ese momento no consume sachet(sig=0.001), la categoría de condición civil casada con (sig=0.007), y si la madre tuvo anemia durante la etapa gestacional (sig=0.000), el resto de las variables no mencionadas fueron no significativas ya que su nivel “p” fue mayor que 0.05 por lo cual no se consideran en el modelo.

Tabla 11. Matriz de confusión con el modelo 02 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Predicción 0	1
Real		
0	47	11
1	1	64

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior se obtiene la confusión de matriz con el modelo 02 donde se valida con el grupo de datos de testing y se comprueba que de los 123 datos separados para test, realizando una tabla cruzada entre los valores reales de la variable

Anemia y los proyectados con un umbral de 0.5; encontrándose este umbral entre las mejores métricas se obtiene que 47 datos fueron predichos como Verdaderos Negativos, seguido de 64 datos que fueron Verdaderos positivos, en seguida 11 datos fueron falsos positivos, y 1 dato clasificado como falso negativos.

Tabla 12. Métricas de evaluación con el modelo 02 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Métricas	Valores
0	Sensibilidad	0.984615
1	Especificidad	0.189655
2	Precision	0.853333
3	Accuracy	0.902439

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior tenemos las métricas de evaluación del modelo 02 donde se obtuvo una sensibilidad de 0.984615, especificidad de 0.1896, precisión del modelo fue de 0.8533 y exactitud del modelo fue de 0.9024.

MODELO LOGIT-03

En este modelo se realizó la selección de características por el método de forward selection. En este método se empieza con 0 variables en el modelo donde paso a paso se empieza a ingresar una nueva variable y luego otra variable verificando su rendimiento del modelo para ver si al entrar una nueva variable esta mejora el rendimiento o empeora para no considerarlo en el modelo, es así que al empezar en 0 variables al final este modelo 03 termino convergiendo con 15 variables de las 38 disponibles en el data consolidada.

Tabla 13. Resumen de modelo 03 método de fordward selection para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

Optimization terminated successfully.
Current function value: 0.111402
Iterations 11

Logit Regression Results

Dep. Variable:	Anemia	No. Observations:	287
Model:	Logit	Df Residuals:	271
Method:	MLE	Df Model:	15
Date:	Sat, 14 Jan 2023	Pseudo R-squ.:	0.8352
Time:	02:01:11	Log-Likelihood:	-31.972
converged:	True	LL-Null:	-194.01
Covariance Type:	nonrobust	LLR p-value:	5.438e-60

Fuente: Elaboración propia a partir de los datos.

Interpretando la tabla anterior se muestra que el número de observaciones para el modelo entrenado fue de 287, donde el resto de las 410 fueron para testing el cual significo la división de 70% para entrenamiento y 30% para prueba del modelo, asimismo se interpreta de la tabla que el modelo converge luego de 11 iteraciones con el método de newton Raphson, presentado un R2 de 0.8352.

Tabla 14. Coeficientes de modelo 03 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	coef	std err	z	P> z	[0.025	0.975]
const	11.8174	5.574	2.120	0.034	0.893	22.742
Peso	-0.8212	0.216	-3.796	0.000	-1.245	-0.397
N_CCred	-0.8845	0.206	-4.292	0.000	-1.288	-0.481
Edad_y	-0.1840	0.101	-1.818	0.069	-0.382	0.014
cat_Sexo_M	1.8778	0.836	2.246	0.025	0.239	3.517
cat_N_Sachets_1	-1.3326	0.727	-1.834	0.067	-2.757	0.092
cat_Condicion_Civil_Casada	2.4203	2.860	0.846	0.397	-3.186	8.026
cat_Condicion_Civil_Convivalente	0.7654	2.858	0.268	0.789	-4.836	6.366
cat_Condicion_Civil_Soltera	0.1672	3.040	0.055	0.956	-5.790	6.125
cat_Nivel_Instrucción Primaria	2.6985	14.378	0.188	0.851	-25.483	30.880
cat_Nivel_Instrucción Secundaria	3.1003	1.122	2.764	0.006	0.902	5.299
cat_Ingreso_De 1000 a 3000	1.4174	0.872	1.626	0.104	-0.291	3.126
cat_Ingreso_Menos de 1000	1.1839	2.913	0.406	0.684	-4.526	6.894
cat_Patologia_Materna Enfermedad Infecciosa	1.4767	0.884	1.670	0.095	-0.257	3.210
cat_Patologia_Materna Otra	1.8306	1.025	1.786	0.074	-0.178	3.839
cat_Anemia_mother_Si	5.8121	1.120	5.188	0.000	3.616	8.008

Fuente: Elaboración propia a partir de los datos.

Describiendo los resultados presentados en la tabla anterior, de las variables presentadas para entrenamiento del “modelo 03”, se presentan sus respectivos

coeficientes detallado por cada una de ellas que vienen comprendidas desde la constante siendo significativa con ($\text{sig}=0.034$), seguido del peso($\text{sig}=0.000$), el número de consultas CRED($\text{sig}=0.000$), la edad de la madre al momento de la encuesta($\text{sig}=0.069$), la categoría de sexo masculino($\text{sig}=0.025$), la categoría de nivel de instrucción secundaria ($\text{sig}=0.006$), y si la madre tuvo anemia durante la etapa gestacional ($\text{sig}=0.000$), el resto de las variables no mencionadas fueron no significativas ya que su nivel “p” fue mayor que 0.05 por lo cual no se consideran en el modelo.

Tabla 15. Matriz de confusión con el modelo 03 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Predicción 0	1
Real		
0	52	6
1	7	58

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior se obtiene la confusión de matriz con el modelo 03 donde se valida con el grupo de datos de testing y se comprueba que de los 123 datos separados para validación, realizando una tabla cruzada entre los valores reales de la variable Anemia y los proyectados con un umbral de 0.7; se verificó que este umbral entre las mejores métricas se obtiene que 52 datos fueron predichos como Verdaderos Negativos, seguido de 58 datos que fueron Verdaderos positivos, en seguida 6 datos fueron falsos positivos, y 7 datos clasificados como falso negativos.

Tabla 16. Métricas de evaluación con el modelo 03 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Métricas	Valores
0	Sensibilidad	0.892308
1	Especificidad	0.103448
2	Precision	0.906250
3	Accuracy	0.894309

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior tenemos las métricas de evaluación del modelo 03 donde se obtuvo una sensibilidad de 0.8923, especificidad de 0.1034, precisión del modelo fue de 0.9062 y exactitud del modelo fue de 0.8943.

MODELO LOGIT- 04

En este modelo se realizó la selección de características por el método de Backward selection. En este método se empieza con la totalidad de variables en el modelo donde paso a paso se empieza a eliminar una nueva variable y luego otra variable verificando su rendimiento del modelo para ver si al sacar una variable esta mejora el rendimiento o empeora para no considerarlo en el modelo, es así que al empezar en 38 variables al final este modelo 04 termino convergiendo con 15 variables de las 38 disponibles en el data consolidada.

Tabla 17. Resumen de modelo 04 método de Backward selection para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

```

Optimization terminated successfully.
Current function value: 0.111941
Iterations 12
Logit Regression Results
=====
==
Dep. Variable:          Anemia      No. Observations:      287
Model:                 Logit      Df Residuals:          271
Method:                MLE       Df Model:              15
Date:                  Sat, 14 Jan 2023      Pseudo R-squ.:        0.8344
Time:                  02:12:48      Log-Likelihood:        -32.127
converged:             True       LL-Null:               -194.01
Covariance Type:      nonrobust   LLR p-value:          6.308e-60
=====

```

==**Fuente:** Elaboración propia a partir de los datos.

De la tabla anterior, se muestra que el número de observaciones para el modelo entrenado fue de 287, el cual significa la división de 70% para entrenamiento y 30% para prueba del modelo, asimismo se interpreta de la tabla que el modelo converge luego de 12 iteraciones con el método de Newton Raphson, presentado un R2 de 0.8344.

Tabla 18. Coeficientes de modelo 04 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	coef	std err	z	P> z	[0.025	0.975]
const	12.8831	4.915	2.621	0.009	3.249	22.517
Peso	-0.8545	0.222	-3.841	0.000	-1.290	-0.419
N_Ccred	-0.9308	0.226	-4.120	0.000	-1.374	-0.488
Edad_y	-0.2521	0.106	-2.372	0.018	-0.460	-0.044
cat_Sexo_M	2.2178	0.888	2.498	0.012	0.478	3.958
cat_N_Sachets_0	1.2600	0.714	1.765	0.078	-0.139	2.659
cat_N_Consejerias_0	1.5336	0.920	1.667	0.095	-0.269	3.336
cat_Condicion_Civil_Casada	1.9976	0.830	2.407	0.016	0.371	3.624
cat_Condicion_Civil_Soltera	-1.1755	1.233	-0.954	0.340	-3.592	1.241
cat_Nivel_Instrucción Primaria	3.5655	20.655	0.173	0.863	-36.918	44.049
cat_Nivel_Instrucción Secundaria	3.9362	1.354	2.907	0.004	1.283	6.590
cat_Ingreso_De 1000 a 3000	1.0028	3.047	0.329	0.742	-4.968	6.974
cat_Ingreso_De 3000 a 5000	-0.8129	3.176	-0.256	0.798	-7.037	5.411
cat_Ingreso_Mas de 5000	-0.3398	3.181	-0.107	0.915	-6.574	5.894
cat_Patologia_Materna Enfermedad Infecciosa	1.4488	0.908	1.595	0.111	-0.332	3.229
cat_Anemia_mother_Si	6.1425	1.253	4.901	0.000	3.686	8.599

Fuente: Elaboración propia a partir de los datos.

Describiendo los resultados presentados en la tabla anterior, de las variables presentadas para entrenamiento del “modelo 04”, se presentan sus respectivos

coeficientes detallado por cada una de ellas que vienen comprendidas desde la constante siendo significativa con ($\text{sig}=0.009$), seguido del peso($\text{sig}=0.000$), el número de consultas CRED($\text{sig}=0.000$), la edad de la madre al momento de la encuesta($\text{sig}=0.018$), la categoría de sexo masculino($\text{sig}=0.012$), la categoría de condición civil casado con $\text{sig}(0.016)$, nivel de instrucción secundaria ($\text{sig}=0.004$), y si la madre tuvo anemia durante la etapa gestacional ($\text{sig}=000$), el resto de las variables no mencionadas fueron no significativas ya que su nivel “p” fue mayor que 0.05 por lo cual no se consideran en el modelo.

Tabla 19. Matriz de confusión con el modelo 04 para la predicción de anemia en menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Predicción 0	1
Real		
0	53	5
1	6	59

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior se obtiene la confusión de matriz con el modelo 03 donde se valida con el grupo de datos de testing y se comprueba que de los 123 datos separados para validación, realizando una tabla cruzada entre los valores reales de la variable Anemia y los proyectados con un umbral óptimo de 0.6 se verificó que este umbral entre las mejores métricas y se obtiene que 53 datos fueron predichos como Verdaderos Negativos, seguido de 59 datos que fueron Verdaderos positivos, en seguida 5 datos fueron falsos positivos, y 6 datos clasificados como falso negativos.

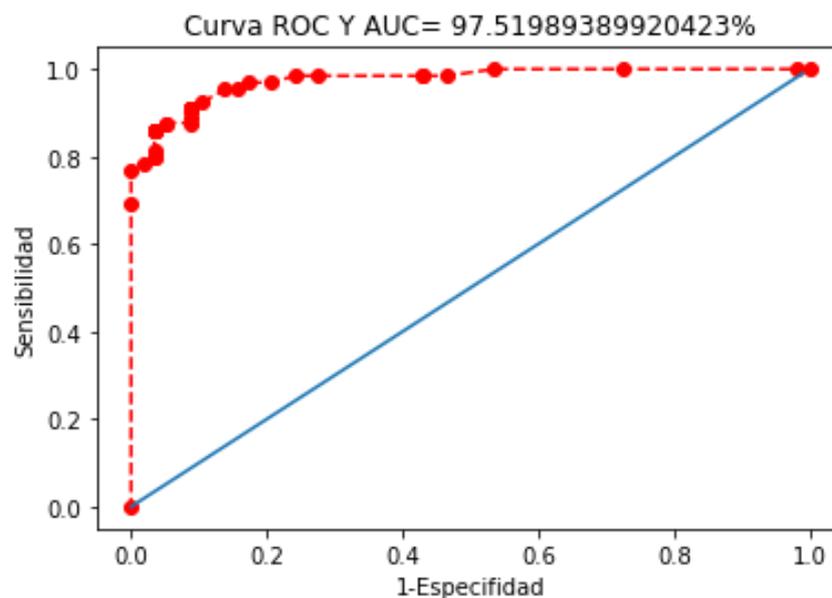
Tabla 20. Métricas de evaluación con el modelo 04 para la predicción de anemia en niños de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021

	Métricas	Valores
0	Sensibilidad	0.907692
1	Especificidad	0.086207
2	Precision	0.921875
3	Accuracy	0.910569

Fuente: Elaboración propia a partir de los datos.

De la tabla anterior tenemos las métricas de evaluación del modelo 04 donde se obtuvo una sensibilidad de 0.9076, especificidad de 0.0862, precisión del modelo fue de 0.9218 y exactitud del modelo fue de 0.9105.

Figura 10. Curva Roc y área bajo la curva del modelo 02 con diferentes umbrales numéricas en los menores de 3 a 59 meses atendidos en el centro de salud I-3 Vallecito-Puno-2021.



Otra de las métricas importantes es luego de haber probado los diferentes modelos que se han presentado y han convergido, y mediante los diferentes umbrales

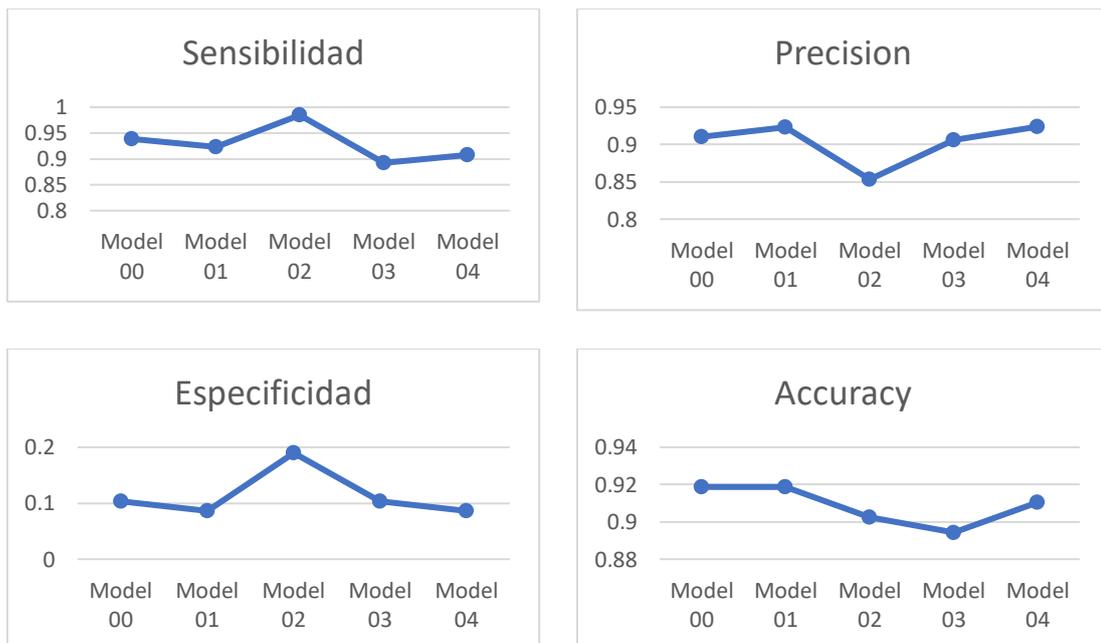
propuestos en uno de los modelos óptimos, se forma una gran cantidad de sensibilidad y especificidad que viene comprendido en el grafico anterior donde se tiene en general del 97.6% de los datos caen debajo de la curva por lo cual se concluye que el modelo propuesto es óptimo y las variables seleccionadas representan significativamente los modelos presentados.

Tabla 21. Resumen de los resultados de métricas de evaluación de los modelos planteados.

Métricas	Sensibilidad	Especificidad	Precision	Accuracy
Model 00	0.938462	0.103448	0.910448	0.918699
Model 01	0.923077	0.086207	0.923077	0.918699
Model 02	0.984615	0.189655	0.853333	0.902439
Model 03	0.892308	0.103448	0.90625	0.894309
Model 04	0.907692	0.086207	0.923875	0.910569

Fuente: Elaboración propia a partir de los datos.

Figura 11. Métricas de evaluación de los diferentes modelos planteados según métrica.



Del cuadro y grafico anterior evidenciamos que el modelo propuesto 01 tiene las mejores métricas de evaluación de modelo ya que en cuanto a la precisión y exactitud tiene los indicadores más altos con 0.923077 y 0.918699 respectivamente mientras que si bien es cierto que la sensibilidad está en tercer lugar, pero en cuanto a la especificidad está en el primer lugar ya que presenta 0.086207 y esta se mide según el valor más bajo porque se está restando 1 – especificidad por lo cual el valor más bajo es el más optimo.

Entonces a partir de ello el modelo 02 tiene las siguientes variables significativas de acuerdo a los factores estudiados en la presente investigación.

Tabla 22. Resumen de los coeficientes significativos según los factores estudiados en el modelo optimo.

Coef	Std	err	z	P> z	[0.025	0.975]
FACTORES BIOLÓGICOS						
Peso	-0.7779	0.19	-4.105	0.000	-1.149	-0.406
cat_Sexo_F	-2.0649	0.824	-2.505	0.012	-3.68	-0.449
FACTORES SIEN CRED						
N_CCred	-0.8686	0.186	-4.67	0.000	-1.233	-0.504
cat_N_Sachets_0	1.6037	0.695	2.307	0.021	0.241	2.966
FACTORES SOCIODEMOGRÁFICOS DE MADRE						
cat_Condicion_Civil_Casada	1.5483	0.73	2.12	0.034	0.117	2.979
FACTORES GINECO-OBSTÉTRICOS						
cat_Anemia_mother_Si	5.5666	1.018	5.469	0.000	3.571	7.562

Fuente: Elaboración propia a partir de los datos.

Del cuadro anterior se infiere que dentro de los factores biológicos influye a la presencia de anemia el peso y el sexo femenino en el niño, dentro de los factores SIEN CRED influye en la presencia de anemia el número de consultas de la madre que tuvo en el establecimiento de salud y el niño que no consume sachet influye a la presencia de anemia, asimismo dentro de los factores sociodemográficos tiene una influencia de las madres que tienen la condición civil casada, y finalmente entre los

factores gineco obstétricos para el presente modelo se tiene influencia sobre el antecedente de la madre que tuvo anemia en la etapa gestacional; todos estos factores descritos influyen en la presencia y/o ausencia de anemia en menores de 3 a 59 meses del centro de salud Vallecito I-3 de la ciudad de Puno en el año 2021. Quedando el modelo de la siguiente manera:

$$p(y = Anemia_{si}) = \frac{1}{1 + e^{-[\beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_n X_{ni}]}} = \frac{1}{1 + e^{-y}}$$

Donde:

$$\begin{aligned} \ln(odds) = & [-0.778(Peso) - 2.0645(Sexo F) - 0.868(N CCred) \\ & + 1.60(Sachet No) + 1.548(Casada) + 5.56(Ane Madre gest)] \end{aligned}$$

Interpretación de Parámetros:

El peso del niño en el momento de su medición de hemoglobina tiene un coeficiente de regresión de -0.778, esto nos indica que a medida que el peso del menor aumenta entonces la presencia de anemia en el menor disminuye. Seguidamente el coeficiente significativo del sexo tiene -2.0645 dándonos a entender que si el niño es de sexo femenino=1 entonces la presencia de anemia disminuye que si fuera varón esta probabilidad incrementaría, seguidamente tenemos el coeficiente de la variable número de consultas de crecimiento que realizó una madre en el establecimiento de salud, donde tiene coeficiente de -0.868, esto nos indica que mientras más consultas CRED tenga una madre disminuye el nivel de probabilidad de que un menor tenga anemia, asimismo se observa la variable consumo de sachet_0 donde se tiene significativo con un coeficiente de +1.60 en el consumo de sachet en los niños, dándonos a entender que tendrá mayor probabilidad de tener anemia un menor si es que no consume sachet, cuando haya un cambio de que si consume sachet entonces



esta probabilidad disminuirá en la presencia de anemia del niño, también se encontró significativo a la variable condición civil de la madre de que sea casada y tenga una familia estable con coeficiente de +1.548 interpretando que la probabilidad de tener anemia de la madre ese incrementa estando en esta condición civil, aquí se debe informar que no importa la condición civil que se tenga para la presencia de anemia, ya que si el estado civil casado es lo mejor para una familia, esto no te asegura la presencia o ausencia de anemia en los menore, ya que al estar casado igual se tiene altas probabilidades de tener anemia en sus hijos recién nacidos, finalmente se tiene la variable anemia en la madre en la etapa gestacional con coeficiente de +5.56 siendo un valor positivo y alto, se interpreta que mientras la madre tenga anemia en la etapa gestacional, entonces existirá una alta probabilidad de que su hijo si tenga anemia también y caso contrario si la madre para a no tener anemia en la etapa gestacional. El menor recién nacido también disminuirá las probabilidades de tener anemia.

4.2. DISCUSIÓN

Según el ministerio de salud en el año 2017 haciendo referencia a la encuesta demográfica se encontró que a nivel nacional se tiene el 43.6% de menores con anemia comprendidos entre 3 y 59 meses, y menciona que la región con mayor prevalencia de anemia es la región de Puno con 69.4% de población estudiada, justamente la presente investigación realizada en el puesto de salud Vallecito en el año 2021 tenemos que del 100% de la población estudiada de menores de entre 3 y 59 meses atendidos durante todo el año se presentó el 57.3% de menores con anemia, esto siendo un puesto de salud ubicado en zona urbana aun así se nota que el nivel de anemia es alto y se corrobora con lo mencionado por el ministerio de salud.



El objetivo del estudio fue determinar el modelo logit óptimo que permita identificar los factores asociados que influyen en la presencia de anemia de los niños menores de 5 años de edad donde estudios parecidos como el de Canaza(2020) que fue realizado en la Microred Yauri de la provincia de espinar, encontró un modelo óptimo con las variables predictoras de sector, altitud de localidad, número de visitas realizadas a la Microred por parte de la madre, donde no se estudió en la presente investigación el sector ni la altura de la localidad ya que el establecimiento de salud vallecito tiene la misma altura en msnm para todos los sectores de su jurisdicción pero si se concuerda con el número de visita o número de consultas CRED como se le denominó en la operacionalización de variables. Otro estudio reciente realizando modelos logit fue de (Navarro & Oliva, 2020) donde realizó el análisis de la base de datos ENDES 2018, encontrando 14 covariables significativas que influyen en su modelo de las cuales se concuerda con las siguientes: el sexo del menor de edad, la presencia de anemia de la madre durante el embarazo, el consumo de suplementos para combatir la deficiencia de hierro que sería el equivalente al consumo de sachet. Asimismo es importante mencionar que de las variables utilizadas para la presente investigación y que también fueron utilizadas las investigadoras, no se pudo entrar significancia para el modelo propuesto pero que ellas si lograron y son las siguientes, la edad del niño en meses, grado de instrucción de la madre y el índice de riqueza en el hogar equivalente al ingreso promedio mensual, quizá no se encontró significancia por la diferencia existente en la captura de datos de sus niveles de riqueza.

En cuando al modelo óptimo que presentaron y el análisis de sus métricas de evaluación se encontró que (Navarro & Oliva, 2020) presentaron su modelo con una exactitud del 69.5% mientras que en la presente investigación se logró ajustar y tener



una exactitud con el modelo 02 de 90% por ende nuestro modelo queda muy superior en esta métrica respecto a la presentado por Navarro y Oliva.

Mientras que Canaza (2021) logro una especificidad de 93.8% y una sensibilidad de 96.2% el cual en el modelo propuesto en la presente investigación se presentó una especificidad de (1-especifi_1) dándonos el 81% de clasificados como verdaderos negativos, siendo más bajo que el presentado por Canaza, pero en cuanto a la sensibilidad se logró obtener 98.46% quedando muy superior el modelo de la presente investigación. En cuanto a la curva ROC encontró un área bajo la curva Canaza (2021) del 97.6/ mientras que en esta investigación se logró obtener la curva ROC de 97.51%. Asimismo, es importante indicar que los investigadores mencionados utilizaron paquetes estadísticos con software no libre y que no están soportados para a realizar ciencia de datos, por lo cual en esta investigación si se utilizaron software libre y paquetes de Python que es era el análisis y la formación de modelos predictivos de gran volumen de datos

V. CONCLUSIONES

El presente trabajo tuvo como finalidad identificar los factores en la prevalencia de anemia mediante un modelo óptimo de regresión logística aplicando machine learning para los menores de 3 a 59 meses de nacido en el Centro de Salud Vallecito I-3 de la ciudad de Puno en el año 2021. Donde en el desarrollo de toda la presente investigación se diseñaron y proyectaron varios modelos, se presentaron y describieron 05 modelos que tuvieron las mejores métricas de evaluación y que todas estas convergieron.

Se determinó mediante la validación de diferentes métricas que el modelo 02 presenta mejor los factores que influyen en la presencia o ausencia de anemia donde fueron los siguientes: de los factores biológicas el peso de niño($b=-0.779$, $\text{sig}=0.000$) y la categoría de sexo($b=-2.0649$, $\text{sig}=0.012$) fueron las más significativas, de los factores de SIEN CRED el número de consultas de crecimiento(-0.8686 , $\text{sig}=0.000$) y el no consumo de sachet($1.6037, 0.021$) fueron significativas, seguido de factores sociodemográficos de la madre donde solo la categoría de condición civil casada($b=1.5483$, $\text{sig}=0.034$) fue significativa y finalmente los factores gineco obstétricos de la madre influyeron en la presencia de anemia de la madre durante la etapa gestacional(5.5666 , $\text{sig}=0.000$),

El modelo final propuesto quedo como la siguiente $\text{Ln(odds)} = [-0.778(\text{Peso}) - 2.0645(\text{Sexo F}) - 0.868(\text{N CCred}) + 1.60(\text{Sachet No}) + 1.548(\text{Casada}) + 5.56(\text{Ane Madre gest})]$, fue este que tuvo las mejores métricas de evaluación con sensibilidad de 0.984615, especificidad -1 de 0.1896, precisión del modelo fue de 0.8533 y exactitud del modelo fue de 0.9024, y teniendo una curva ROC de AUC de 97.5199%.



VI. RECOMENDACIONES

Después de identificar y cumplir con los objetivos propuestos en la prevalencia de anemia se expresan las siguientes sugerencias

Se recomienda poner mayor énfasis en la política de la etapa gestacional de la madre para la ausencia de anemia ya que esta influye directamente con alta probabilidad para la presencia de anemia del niño.

Se recomienda realizar ciencia de datos con las librerías utilizadas en la presente investigación para el análisis de diferentes datos que se recopilan a diario en el sector salud.

Dada la elevada prevalencia de este padecimiento realizar investigaciones en otros centros de salud de nuestra región, en zonas rurales, para así conocer otros factores para la prevalencia de anemia que no se hayan encontrado en este estudio.

Se recomienda utilizar esto como antecedente a futuras investigaciones y a la vez utilizar otras técnicas de machine learning existentes para la predicción y anemia, así como otros estudios que se requieran realizar modelos predictivos.



VII. REFERENCIAS BIBLIOGRÁFICAS

- Aignerren, M., (2009) *El cuestionario el instrumento de recolección de información de la técnica de la encuesta social*. Centro de Estudios de OPINIÓN [En línea], pp. 1-79. Disponible en: <http://aprendeenlinea.udea.edu.co/revistas/index.php/ceo/article/viewFile/1696/134>
- Aguirre Andrade, M. (2015). *Factores asociados con la prevalencia de la anemia y la desnutrición, en niños que acuden al Centro de Salud San Juan Bautista y Socos, Huamanga, 2015*. Universidad Nacional de San Cristóbal de Huamanga.
- Alderete, A. M. (2006). *Fundamentos del Análisis de Regresión Logística en la Investigación Psicológica*.
- Álvarez Cáceres, R. (1995). *Estadística multivariante y no paramétrica con SPSS*. Madrid.
- Álvarez Jiménez, K. F. (2019). *Conocimientos y prácticas sobre prevención de la anemia ferropénica en madres de niños menores de 5 años de la Institución Educativa José Abelardo Quiñones del distrito de Los Olivos-2019*. Universidad Privada del Norte.
- Boddy, Clive. *Sample size for qualitative research* [En línea]. *Qualitative Market Research*. Reino Unido: Emerald Group Publishing Limited, Vol. 19 No. 4, pp. 426-432, 2016. Disponible en: doi.org/10.1108/QMR-06-2016-0053.
- Casella, G., & Berger, R. L. (2002). *Statistical inference* (2nd ed.). Duxbury Press.
- Cohen, J. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Lawrence Erlbaum Associates.
- Castillo, F. (2022). *Regresión logística: factores de riesgo que inciden al trabajo infantil - Encuesta Nacional de Hogares, Perú 2021*. [Trabajo de suficiencia profesional de pregrado, Universidad Nacional Mayor de San Marcos, Facultad de Ciencias Matemáticas, Escuela Profesional de Estadística]. Repositorio institucional Cybertesis UNMSM.
- Centro Nacional de Alimentación y Nutrición. (2007). *Tabla de valoración nutricional antropométrica: niños < 5 años*.



- Cevallos Melo, D. O., & Vélez Gonzáles, J. R. (2018). *Prevalencia de anemia en niños menores de cinco años atendidos en el área de pediatría del Hospital General Guasmo Sur, periodo enero – junio 2017*. Universidad Católica de Santiago de Guayaquil, Guayaquil.
- Charapaqui Jacinto, R. A. (2019). *Variables sociodemográficas y materno infantiles asociados a*. Universidad Ricardo Palma.
- Covri Rivera, D., & Maldonado Ambrosi, A. E. (2023). *Puntuación de crédito en la Cooperativa de Ahorro y Crédito Jardín Azuayo mediante regresión logística*. Revista Ñeque, 6(14), 71–83. <https://doi.org/10.33996/revistaneque.v6i14.108>
- Cutipa Huarsaya, N. M. (2019). *Factores de riesgo asociados a la anemia ferropénica en niños menores de 36 meses de los Establecimientos de Salud Simón Bolívar I-3 y 4 de noviembre I - 3 - 2016*. Universidad Nacional del Altiplano.
- Díaz Cayro, L. R. (2019). *Factores Asociados a la Presencia de Anemia Nutricional en niños de 1 a 5 años, Micro red Cono Sur diciembre 2015 marzo 2016*. Universidad Andina Néstor Cáceres Velasquez.
- Flores Manrique, L. (2002). *Análisis estadístico de los factores de riesgo que influyen en la enfermedad Angina de Pecho*. UNMSM, Lima.
- García, T., (2003). El cuestionario como instrumento de investigación/evaluación. Página del proyecto de apoyo para profesionales de la formación (PROMETEO) de la Junta de Andalucía [En línea], pp. 28. Disponible en: http://www.univsantana.com/sociologia/El_Cuestionario.pdf.
- Gómez Restrepo, M. D., & García Granda, M. G. (2015). *Prevalencia de anemia y factores asociados en la población infantil que acude a la consulta externa del Hospital Moreno Vázquez, Gualaceo 2014*. Universidad de Cuenca.
- Guibert López, L. A. (2014). *Factores de riesgo asociados a anemia en niños a los seis meses de edad atendidos en el hospital belén de Trujillo*. Universidad Privada Antenor Orrego - UPAO.
- Hernández, R., Fernández, C., & Baptista, M. (2014). *Metodología de la Investigación*. México DF, México: McGraw-Hill
- Huacan Mamani, K. Y., & Yepéz Zúñiga, L. F. (2017). *Factores que Influyen en la Presencia de Anemia Ferropénica en niños de 6 a 35 meses en el Centro de Salud José Antonio Encinas Puno diciembre 2016 – febrero 2017*. Universidad Andina Néstor Cáceres Velasquez.



- Instituto Nacional de Estadística e Informática. (1995). *DEMANDA DE LA PLANIFICACION FAMILIAR EN EL PERU*. Lima.
- Instituto Nacional de Estadística e Informática. (2018). *Lactancia y nutrición de niñas, niños y mujeres. Encuesta Demográfica y de Salud Familiar*.
- Instituto Nacional de Estadística e Informática. (2019). *Perú: Indicadores de resultados de los programas presupuestales, primer semestre 2019. Encuesta demográfica y de salud familiar*.
- Instituto Nacional de Estadística e Informática. (2021). *Encuesta Demográfica y de Salud Familiar ENDES 2020 en Perú*.
- Irma Yupari-Azabache, Lucia Bardales-Aguirre, Julio Rodriguez-Azabache, J. Shamir Barros-Sevillano, Ángela Rodríguez-Díaz. *Factores de riesgo de mortalidad por covid-19 en pacientes hospitalizados: un modelo de regresión logística*. Rev. Fac. Med. Hum. Enero 2021; 21(1):19-27. DOI 10.25176/RFMH.v21i1.3264
- Jesus Pacheco, E. E. (2017). *Factores socioculturales de madres y prevención de anemia en niños de 6 a 36 meses, centro de salud San Luis, 2017*. UNIVERSIDAD INCA GARCILASO DE LA VEGA.
- Johnson, A., & Smith, J. (2018). Non-experimental Research Design. In Handbook of Research Methods in Psychology (2nd ed., pp. 123-145). Wiley-Blackwell.
- López, P., & Fachelli, S. (2015). *Metodología de la Investigación Social Cuantitativa*. Barcelona.
- Maquera Layme, E. (2017). *Prevalencia de anemia en niños de 6 a 35 meses de edad en el Centro de Salud Metropolitano de Ilave, Puno 2016*. Universidad Nacional del Altiplano.
- Mayo clinic. (2022). *Anemia*.
- Mayhua Quenta, R. (2023). *Factores que influyen en las preferencias del consumo de quinua orgánica en los hogares de la ciudad de Puno, 2020*. Semestre Económico, 12(1), 27–38. <https://doi.org/10.26867/se.2023.v12i1.142>
- Ministerio de Desarrollo e Inclusión Social. (2018). *Plan multisectorial de lucha contra la Anemia*.
- Ministerio de Salud. (2016). *Guía de práctica clínica para el diagnóstico y tratamiento de la anemia por deficiencia de hierro en niñas, niños y adolescentes en establecimientos de salud de primer nivel de atención*. Lima.



- Ministerio de Salud. (2017). *Norma Técnica - Manejo terapéutico y preventivo de la Anemia en niños, adolescentes, mujeres gestantes y puérperas*. Lima.
- Ministerio de Salud. (2017). *Plan Nacional para la Reducción y Control de la Anemia Materno Infantil y la Desnutrición Crónica Infantil en el Perú: 2017-2021*.
- Organización Mundial de la Salud. (1975). *Lucha contra la anemia nutricional, especialmente con la carencia de hierro*.
- Organización Mundial de la Salud. (2011). *Concentraciones de hemoglobina para diagnosticar la anemia y evaluar su gravedad*.
- Parra, J. (2022, enero 13). *Ecoknowmic, Econometría - Introducción a las variables ficticias o variables dummy*. Javier Parra | Economía y finanzas digitales. <https://www.javierparra.net/ecoknowmic/introduccion-a-las-variables-ficticias-o-variables-dummy/>
- Paredes Flores, D. (2019). *Factores relacionados a la anemia en niños de 6 a 23 meses de edad, atendidos en el puesto de salud intiorko, Tacna año 2014*.
- Paz Fernández, A. C. (2017). *Factores de Riesgo Asociados a Anemia en Niños de 6 a 24 Meses Atendidos en el Hospital I Edmundo Escobel Año 2016*. Universidad Católica de Santa María.
- Pedrero, V., Reynaldos-Grandón, K., Ureta-Achurra, J., & Cortez-Pinto, E. (2021). Generalidades del Machine Learning y su aplicación en la gestión sanitaria en Servicios de Urgencia. *Revista medica de Chile*, 149(2), 248–254. <https://doi.org/10.4067/s0034-98872021000200248>
- ¿Qué es Machine Learning? (s/f). IBM.com. Recuperado el 12 de julio de 2023, de <https://www.ibm.com/mx-es/analytics/machine-learning>.
- Ravasco, P., Anderson, H., & Mardones, F. (2010). Métodos de valoración del estado nutricional.
- Rea, C. M. G. (2023). *Modelo matemático para establecer los factores asociados a las calificaciones escolares utilizando regresión logística en los estudiantes del primer ciclo del Instituto Superior Tecnológico Tres de Marzo de la Provincia Bolívar [Universidad Técnica de Ambato]*. <http://repositorio.uta.edu.ec/bitstream/123456789/38614/1/t2302mma.pdf>
- Rivera Palomino, E. F. (2019). *Factores sociodemográficos asociados a la anemia ferropénica en niños menores de 5 años en el Hospital Arzobispo Loayza durante el periodo de enero a noviembre del 2028*. Lima.



- Roque, M. (2018). *Modelos de regresión logística multinomial de la calidad de fibra de alpaca Huacaya en función de sus características: sexo y edad - Corani, Carabaya, Puno - 2017* [Tesis, Universidad Nacional del Altiplano]. <http://repositorio.unap.edu.pe/handle/UNAP/7755>
- Silva Rojas, M., Retureta Rodríguez, E., & Panique Benítez, N. (2015). Incidencia de factores de riesgo asociados a la anemia ferropénica en niños menores de cinco años. *Universidad de Ciencias Médicas de Las Tunas*.
- Silva, L. C., & Barroso Utra, I. M. (2000). Cuaderno Regresion Logistica.
- Smith, J., Johnson, A., & Brown, C. (2019). Machine Learning in Healthcare: A Review. *Journal of Biomedical Informatics*, 95, 103208. doi: 10.1016/j.jbi.2019.103208
- Solís, J., Quiroz, S., Fosado, O. (2022). *Modelo de regresión logística para la estimación de la deserción escolar del posgrado en la Universidad Técnica de Manabí, Ecuador*. *Revista Bases de la Ciencia*, 7(3),1-14. DOI: <https://doi.org/10.33936/revbasdelaciencia.v7i3.5197>
- Torres Garcia, M. G. (2020). *Factores sociodemográficos para la anemia Ferropénica en niños menores de 5 años, MICRORED Cajaruro, Amazonas, 2020*. Chachapoyas.
- Urquidi, C., Vera, C., Trujillo, N., & Mejía, H. (2006). *Prevalencia de anemia en niños de 6 a 24 meses de edad de tres centros de salud de la ciudad de La Paz*. *Revista de la Sociedad Boliviana de Pediatría*.
- Vall d'Hebron Barcelona Hospital Campus. (2020). *Servicio de Obstetricia y Medicina reproductiva del Hospital Vall d'Hebron*.
- Várela, R. (2019). Prevalencia de anemia en niñas/os de 6 a 48 meses que concurren a dos CAIF de la ciudad de Salto. *Universidad de Buenos Aires. Departamento de Medicina Familiar y Comunitaria*.
- Vásquez Servín, A. G. (2017). *Anemia ferropénica en niños menores de 5 años*.



ANEXOS

ANEXOS 01. DATOS PROCEDADOS

n°	nivel anemia	Sexo	Edad (meses)	Peso	Talla	WHZ	HAZ	WAZ	N_Sac	N_Sachets	N_Cons	N_Ser	N_Vis	Edad	cond	nivel	N°	SIS	sis	ING	seman	Dx_CL	DX_C	tipo	patol	ane
1	1	M	1	12.98	12	75	2.79	-0.8	1.87	30	1	0	0	27	3	4	1	Si	1	2	37	Normal	1	1	3	0
2	1	M	1	9.95	9.3	72.2	0.45	-0.4	0.1	30	1	1	0	27	1	4	1	Si	1	2	38	Normal	1	1	2	0
3	0	M	1	48.1	14	95.5	-0.34	-1.9	-1.35	0	0	0	0	17	1	2	1	Si	1	4	37	Deficit	0	1	2	0
4	1	F	0	36.01	16	96	1.4	0.25	1.14	0	0	0	0	29	1	3	1	Si	1	2	40	Normal	1	1	1	0
5	0	F	0	47.8	16	100.5	0.3	-0.5	-0.1	0	0	0	0	17	2	3	1	Si	1	1	37	Deficit	0	1	2	0
6	0	F	0	49.45	16	100.5	0.44	-0.7	-0.15	0	0	0	0	18	2	3	1	Si	1	4	39	Deficit	0	1	2	0
7	1	M	1	54.34	16	105	0.11	-0.4	-0.18	0	0	0	0	30	3	3	2	Si	1	2	40	Normal	1	1	1	0
8	0	F	0	48.16	18	94.3	2.99	-2	0.87	0	0	0	0	18	2	3	1	No	0	1	38	Deficit	0	1	2	0
9	1	M	1	51.98	17	100	1.21	-1.3	0	0	0	0	0	32	2	3	4	Si	1	2	38	Normal	1	1	2	0
10	1	F	0	36.11	17	97.6	2.53	0.65	2.13	0	0	0	0	33	3	3	1	Si	1	2	39	Deficit	0	1	3	0
11	1	F	0	50.04	18	101.5	1.37	-0.6	0.56	0	0	0	0	35	1	3	1	Si	1	2	40	Deficit	0	1	1	0
12	1	F	0	9.1	7.8	68.9	-0.19	-0.6	-0.45	0	0	0	0	35	1	2	4	No	0	2	39	Deficit	0	1	3	0
13	1	M	1	24.02	12	85.5	0.31	-0.5	-0.04	0	0	0	0	37	3	3	4	Si	1	2	38	Normal	1	1	2	0
14	1	F	0	12.09	8.8	74	-0.2	-0	-0.15	30	1	0	0	16	2	3	1	Si	1	2	41	Normal	1	1	2	1
15	1	F	0	13.24	9.9	74.5	0.92	-0.4	0.54	30	1	0	0	17	2	3	1	Si	1	3	38	Normal	1	1	2	0
16	1	M	1	6.21	8	69	-0.3	0.49	-0.02	0	0	0	0	17	1	3	1	Si	1	3	40	Normal	1	1	3	1
17	1	F	0	12.32	13	87	1.14	4.88	2.87	30	1	0	0	19	2	3	1	Si	1	3	40	Normal	1	1	4	1
18	0	F	0	18.1	13	81	2.37	0.08	1.78	30	1	0	0	18	2	3	1	No	0	2	36	Deficit	0	1	2	0
19	1	M	1	24.02	12	89	-0.45	0.61	0.03	0	0	1	0	20	2	3	1	Si	1	1	39	Normal	1	1	2	1
20	1	F	0	6.11	6.1	65	-1.59	-0.4	-1.44	30	1	0	0	20	2	3	1	Si	1	1	37	Normal	1	1	2	1
21	1	F	0	5.03	6.8	63.2	0.29	-0.4	-0.08	0	0	0	0	20	3	3	1	Si	1	3	37	Normal	1	1	3	1
22	1	M	1	18.1	10	78.5	-0.09	-1.4	-0.73	30	1	0	0	21	2	4	1	Si	1	3	37	Normal	1	1	2	1
23	0	M	1	36.01	17	96	2.07	-0	1.39	0	0	0	0	40	2	4	1	Si	1	4	36	Normal	1	1	1	0
24	1	F	0	7.03	8.2	71	0.35	1.6	1	30	1	0	0	21	1	3	1	Si	1	1	37	Normal	1	1	2	1
25	0	M	1	48.16	16	99	0.79	-1.1	-0.13	0	0	1	0	19	1	3	1	Si	1	4	40	Normal	1	1	2	1
26	1	M	1	12.16	10	77.5	0.57	0.66	0.72	0	0	0	0	21	2	3	1	Si	1	3	37	Normal	1	1	1	0
27	0	F	0	5.72	7.7	66	0.46	0.33	0.47	0	0	1	0	20	3	4	1	No	0	4	37	Normal	1	1	4	0
28	1	F	0	24.05	9.1	81	-1.09	-1.5	-1.58	0	0	1	0	21	1	3	1	Si	1	1	38	Normal	1	1	2	1
29	0	F	0	11.04	7.8	73	-1.32	0.07	-0.93	30	1	0	0	20	1	5	1	Si	1	4	40	Normal	1	1	1	0
30	1	M	1	4.44	5.5	61	-1.55	-1.5	-2.17	0	0	0	0	21	2	3	1	Si	1	2	38	Normal	1	1	2	0
31	0	F	0	41.86	16	98.4	0.59	-0.1	0.32	0	0	0	0	20	2	4	1	Si	1	3	39	Normal	1	2	2	0
32	0	F	0	24.02	12	86	0.34	0.08	0.35	0	0	1	0	20	1	5	1	Si	1	4	39	Normal	1	1	2	0
33	1	F	0	9.99	9.2	72.5	0.59	0.42	0.62	30	1	1	0	21	1	3	1	Si	1	2	38	Normal	1	1	2	0
34	1	F	0	30	15	93.5	1.35	0.8	1.44	30	1	0	0	21	2	3	1	Si	1	3	39	Normal	1	1	2	1
35	1	F	0	12.02	7.2	71	-1.08	-1.2	-1.35	30	1	0	0	21	1	3	1	No	0	1	36	Normal	1	1	2	1
36	1	M	1	4.04	7.8	68	-0.26	1.97	0.96	0	0	0	0	21	2	3	1	No	0	3	37	Normal	1	1	2	1
37	0	M	1	50.37	17	107.2	-0.73	0.6	-0.11	0	0	1	0	21	2	4	1	Si	1	3	38	Normal	1	1	3	0
38	1	M	1	6.24	8	67	0.4	-0.5	-0.02	0	0	0	0	22	2	3	1	Si	1	3	42	Normal	1	1	4	0
39	0	M	1	6.08	7.3	63.2	0.83	-2.1	-0.77	30	1	0	0	22	1	4	1	No	0	3	37	Normal	1	1	2	0
40	0	M	1	9.03	9.4	71	1.02	-0.4	0.52	30	1	0	0	23	2	4	1	Si	1	4	36	Deficit	0	1	1	0
41	0	M	1	43.2	13	98	-1.45	-0.6	-1.32	0	0	0	0	33	2	4	1	Si	1	1	39	Deficit	0	1	3	0
42	1	F	0	12.65	9	72	0.54	-1.1	-0.08	30	1	1	0	22	2	3	1	Si	1	2	36	Deficit	0	1	1	1
43	0	F	0	42.12	18	103.3	0.81	1.03	1.13	0	0	0	0	23	2	5	2	Si	1	2	40	Deficit	0	1	3	0
44	1	M	1	11.99	11	77	0.78	0.55	0.85	30	1	1	0	22	2	3	1	Si	1	2	39	Deficit	0	1	4	0
45	0	F	0	35.98	14	97	0	0.52	0.3	0	0	0	0	29	2	4	1	Si	1	4	36	Deficit	0	1	3	0
46	0	M	1	6.97	9.6	72.3	0.91	1.47	1.4	0	0	1	0	29	3	4	2	Si	1	3	37	Deficit	0	1	3	0
47	1	F	0	24.67	11	83	0.4	-1	-0.22	30	1	0	0	22	1	3	1	Si	1	1	40	Deficit	0	1	1	1
48	0	F	0	36.07	13	91	-0.35	-1.1	-0.81	0	0	0	1	31	3	4	1	Si	1	3	38	Deficit	0	1	4	0
49	0	F	0	50.56	15	99.5	0.23	-1.1	-0.51	0	0	0	0	31	2	5	1	Si	1	2	40	Deficit	0	1	2	0
50	1	M	1	6.05	5.1	57	0.19	-5	-3.7	30	1	0	0	22	1	3	1	Si	1	1	40	Normal	1	1	4	1
51	1	M	1	5.95	8.3	68.9	0.19	0.64	0.44	0	0	0	0	22	1	3	2	Si	1	2	39	Normal	1	1	1	1
52	0	F	0	24.02	15	90	1.85	1.32	2.08	0	0	0	0	33	2	4	1	Si	1	2	38	Deficit	0	1	1	0
53	0	M	1	6.08	8.9	67.5	1.44	-0.1	0.98	30	1	0	0	34	3	5	4	Si	1	4	37	Deficit	0	1	1	0
54	0	F	0	36.34	14	95.3	-0.22	0	-0.12	0	0	1	0	24	1	4	1	Si	1	4	37	Deficit	0	1	1	0
55	1	M	1	27.04	10	82	-0.9	-2.4	-1.93	30	1	0	0	22	1	4	1	Si	1	2	38	Normal	1	1	1	0
56	0	M	1	18.46	13	81	2.16	-0.6	1.34	0	0	1	0	22	2	4	1	Si	1	4	40	Deficit	0	1	2	0
57	1	M	1	5.98	8.2	66	1.06	-0.7	0.31	0	0	0	0	23	1	3	2	Si	1	2	38	Normal	1	2	4	0
58	0	F	0	32.16	13	87.6	0.5	-1.3	-0.32	0	0	0	0	19	2	4	1	Si	1	1	38	Deficit	0	1	2	0
59	1	M	1	6.01	11	66	4.32	-0.7	2.88	30	1	0	0	23	1	3	2	Si	1	2	38	Normal	1	1	2	0
60	1	M	1	6.05	7.9	66.2	0.55	-0.7	-0.06	0	0	1	0	23	2	5	1	Si	1	3	38	Normal	1	1	1	0
61	1	M	1	54.08	17	106.5	-0.02	-0.1	-0.03	0	0	0	1	23	2	3	1	Si	1	2	39	Normal	1	1	2	0
62	1	M	1	30.06	14	90.5	0.46	-0.4	0.12	0	0	1	1	24	2	3	1	Si	1	3	38	Normal	1	1	4	0
63	0	M	1	24.05	12	86	0.09	-0.4	-0.12	0	0	0	0	20	2	5	1	Si	1	4	39	Deficit	0	1	2	0
64	1	F	0	36.01	12	86	0.15	-2.4	-1.27	0	0	1	0	17	2	2	1	Si	1	3	40	Normal	1	1	2	1
65																										



81	1	F	0	4.4	5.9	61.8	-0.71	-0.5	-0.88	0	0	0	0	0	25	2	3	2	Si	1	3	37	Normal	1	1	4	0
82	0	F	0	36.01	15	95.5	0.37	0.12	0.35	0	0	0	0	0	35	3	5	2	Si	1	4	37	Deficit	0	1	4	0
83	1	F	0	6.44	7.9	66.5	0.68	0.03	0.48	0	0	0	0	0	25	2	3	1	Si	1	3	39	Normal	1	1	4	0
84	1	M	1	9.99	8.5	69.5	0.27	-1.7	-0.69	30	1	0	0	0	25	2	5	1	Si	1	3	37	Normal	1	1	4	1
85	1	F	0	18.92	9.5	78.3	-0.25	-1.1	-0.72	30	1	0	0	0	25	2	3	1	Si	1	3	38	Normal	1	1	4	0
86	0	F	0	4.21	6.6	58	0.49	-2.1	-1.24	0	0	0	0	0	39	3	5	4	Si	1	4	36	Deficit	0	1	3	0
87	1	M	1	42.05	13	91.3	-0.48	-2.2	-1.54	0	0	1	0	0	25	2	5	1	Si	1	2	38	Normal	1	1	1	0
88	1	F	0	36.07	13	93	-0.33	-0.6	-0.5	0	0	0	0	0	25	2	4	1	Si	1	3	37	Normal	1	1	2	0
89	1	M	1	18.5	10	79	0.16	-1.4	-0.54	28	1	0	0	0	26	2	3	3	Si	1	3	39	Normal	1	1	1	1
90	1	M	1	10.12	11	75	1.72	0.68	1.66	0	0	1	0	0	26	2	3	3	Si	1	3	40	Normal	1	1	4	1
91	1	F	0	6.01	7.6	64.7	0.8	-0.4	0.29	30	1	0	0	0	26	2	3	1	Si	1	3	40	Normal	1	1	1	0
92	1	F	0	6.11	8.1	63.5	1.96	-1	0.82	3	1	1	0	0	26	2	3	3	Si	1	3	38	Normal	1	1	2	0
93	1	F	0	11.96	9.9	73	1.26	-0.4	0.78	30	1	1	0	0	26	2	3	1	Si	1	2	42	Normal	1	1	3	0
94	0	F	0	12.06	11	74	1.88	-0	1.41	30	1	0	0	0	41	5	4	3	Si	1	4	38	Deficit	0	1	4	0
95	0	M	1	30.06	13	88	0.57	-1.2	-0.2	0	0	0	0	0	24	3	4	1	Si	1	4	40	Deficit	0	1	4	0
96	0	F	0	48.13	17	98.7	1.45	-1	0.39	0	0	0	0	0	16	1	3	1	Si	1	2	34	Deficit	0	1	3	0
97	1	F	0	12.94	9.8	77	0.32	0.72	0.55	30	1	0	0	0	26	3	5	1	Si	1	3	40	Normal	1	2	3	0
98	1	F	0	29.63	10	86.6	-1.92	-1.1	-1.92	0	0	0	0	0	26	2	3	2	Si	1	3	40	Normal	1	1	2	0
99	1	F	0	5.95	8.2	68.5	0.41	1.26	0.92	0	0	0	0	0	27	3	4	2	Si	1	2	38	Normal	1	1	2	0
100	0	M	1	31.8	12	85.6	-0.34	-2.2	-1.43	0	0	1	0	0	16	2	2	1	Si	1	2	39	Normal	1	1	3	0
101	1	M	1	36.01	12	93.5	-1.48	-0.7	-1.34	0	0	0	0	0	27	3	3	1	Si	1	3	38	Normal	1	1	4	0
102	1	F	0	11.86	9.2	73.2	0.47	-0.2	0.25	30	1	0	0	0	27	3	4	1	No	0	3	40	Normal	1	1	3	0
103	1	M	1	12.06	8.8	78.5	-1.51	1.13	-0.65	30	1	0	0	0	27	3	3	3	Si	1	4	36	Normal	1	1	3	0
104	0	M	1	12.55	11	76	1.24	-0.2	0.91	30	1	1	0	0	17	2	3	1	Si	1	3	38	Normal	1	1	4	1
105	0	F	0	12.32	11	75.4	1.74	0.4	1.48	0	0	0	0	0	17	1	3	1	Si	1	2	37	Normal	1	1	4	1
106	1	F	0	33.68	12	89	-0.37	-1.2	-0.89	30	1	0	0	0	27	3	4	1	Si	1	3	36	Normal	1	1	4	0
107	1	F	0	5.95	6.2	65.3	-0.39	-0.2	-0.44	0	0	0	0	0	27	3	3	2	Si	1	1	40	Deficit	0	1	1	1
108	1	F	0	19.09	10	78	0.45	-1.3	-0.28	30	1	0	0	0	27	3	3	2	Si	1	1	38	Deficit	0	1	4	1
109	1	F	0	6.08	8.4	64	2.15	-0.8	1.12	30	1	0	0	0	27	3	3	1	Si	1	2	37	Deficit	0	1	2	0
110	0	F	0	36.04	14	93	0.65	-0.5	0.19	0	0	0	0	0	17	2	3	1	Si	1	2	39	Normal	1	1	4	0
111	0	F	0	6.05	7.9	67	0.52	0.53	0.63	30	1	0	0	0	18	1	3	1	Si	1	2	40	Normal	1	1	3	0
112	0	F	0	24.94	13	86.2	-0.39	-0.1	-0.26	0	0	0	0	0	19	3	5	1	No	0	3	33	Normal	1	1	3	0
113	1	M	1	10.05	9.8	72	1.13	-0.6	0.56	30	1	1	0	0	27	2	3	3	Si	1	3	38	Normal	1	1	2	0
114	0	M	1	30.98	13	88.5	0.45	-1.2	-0.3	0	0	0	0	0	19	2	4	1	Si	1	3	36	Normal	1	1	4	0
115	0	F	0	24.54	13	90	-0.04	1.18	0.67	0	0	0	0	0	19	3	5	1	Si	1	3	37	Normal	1	1	2	0
116	1	M	1	12.16	9.5	73.4	0.47	-1.1	-0.14	30	1	0	0	0	27	3	3	1	Si	1	3	38	Normal	1	1	1	0
117	1	M	1	18	12	78	1.63	-1.6	0.53	30	1	1	0	0	27	3	3	2	Si	1	3	40	Normal	1	1	3	0
118	0	M	1	43.99	17	102.5	0.86	0.37	0.78	0	0	0	0	0	19	2	3	1	Si	1	2	38	Normal	1	1	4	0
119	1	M	1	4.07	7.5	65.5	0.19	0.73	0.58	0	0	1	0	0	27	3	3	1	Si	1	3	40	obrePes	2	1	4	1
120	1	M	1	6.05	5.6	61	-1.37	-3.1	-3.12	0	0	0	0	0	27	3	4	2	Si	1	3	40	obrePes	2	1	3	1
121	1	F	0	6.08	7.7	65	0.9	-0.4	0.41	0	0	0	0	0	27	3	3	1	Si	1	3	38	obrePes	2	1	1	1
122	1	F	0	32.13	13	87.5	0.87	-1.3	-0.07	0	0	0	0	0	28	3	3	1	Si	1	3	40	obrePes	2	1	2	1
123	1	M	1	24.08	12	84.7	0.66	-0.8	0.06	0	0	0	0	0	28	3	4	1	Si	1	2	38	obrePes	2	1	4	0
124	1	F	0	20.17	13	86.1	1.35	1.08	1.56	30	1	0	0	0	28	3	3	1	Si	1	3	39	obrePes	2	1	4	1
125	1	M	1	7.26	10	72	1.84	1.15	1.99	30	1	1	0	0	28	3	3	3	Si	1	3	38	obrePes	2	1	2	1
126	1	M	1	12.06	9.8	73.5	0.7	-1	0.08	30	1	0	0	0	28	2	3	2	Si	1	2	39	obrePes	2	1	3	0
127	1	F	0	12.12	10	74	1.19	-0.1	0.88	30	1	0	0	0	28	3	5	3	Si	1	3	38	obrePes	2	1	1	1
128	1	M	1	12.42	9.5	75	0	-0.5	-0.23	30	1	0	0	0	28	3	3	1	Si	1	4	38	obrePes	2	1	4	0
129	1	M	1	11.99	11	76	1.78	0.13	1.46	30	1	1	0	0	28	3	5	1	Si	1	3	39	obrePes	2	1	3	1
130	0	M	1	18.04	11	84	0.2	0.63	0.4	30	1	1	0	0	20	2	4	1	Si	1	3	38	Normal	1	1	4	0
131	1	F	0	6.01	8.5	67.5	1.15	0.8	1.25	30	1	0	0	0	28	3	4	1	Si	1	3	40	obrePes	2	1	1	1
132	1	F	0	30.06	10	83.4	-1.09	-2.1	-1.98	30	1	0	0	0	28	3	3	2	Si	1	2	37	obrePes	2	1	1	0
133	1	M	1	33.91	14	91.5	0.86	-0.9	0.13	0	0	1	0	0	28	3	4	1	Si	1	3	38	obrePes	2	1	4	1
134	1	F	0	25.69	11	82.3	0.22	-1.5	-0.62	0	0	0	0	0	29	3	3	1	Si	1	3	36	obrePes	2	1	1	1
135	1	F	0	36.01	14	91.7	0.84	-0.9	0.14	0	0	0	0	0	29	3	3	3	Si	1	2	37	obrePes	2	1	1	1
136	0	M	1	11.99	9	72	0.18	-1.6	-0.63	30	1	1	0	0	20	3	3	1	Si	1	2	39	obrePes	2	1	2	0
137	0	M	1	48.03	15	100	-0.29	-0.8	-0.68	0	0	0	0	0	20	2	4	1	Si	1	4	37	Deficit	0	1	4	0
138	1	M	1	12.06	10	74	0.97	-0.8	0.4	30	1	0	0	0	29	3	4	2	Si	1	3	38	obrePes	2	1	4	1
139	1	M	1	11.93	9.2	74	-0.19	-0.7	-0.47	0	0	0	0	0	29	3	3	1	Si	1	2	40	Deficit	0	1	1	1
140	1	F	0	5.78	7.6	64.3	1.01	-0.5	0.44	0	0	0	0	0	29	3	3	1	Si	1	2	40	Deficit	0	1	1	1
141	1	M	1	19.75	10	81	-0.51	-1.1	-0.91	30	1	0	0	0	29	3	3	2	Si	1	1	39	Deficit	0	1	2	1
142	0	M	1	11.99	9.4	73	0.34	-1.1	-0.28	30	1	0	0	0	20	2	3	1	Si	1	3	37	Deficit	0	1	3	1
143	1	F	0	7.13	7.4	68	-0.51	0.22	-0.31	30	1	0	0	0	29	3	3	2	Si	1	3	38	obrePes	2	2	3	1
144	0	M	1	13.01	11	75	1.44	-0.8	0.77	30	1	0	0	0	20	2	5	1	Si	1	2	38	Deficit	0	1		



162	1	F	0	15.24	10	78.2	0.72	0.17	0.61	0	0	0	0	0	0	31	2	3	1	No	0	3	37	obrePes	2	1	2	1
163	1	F	0	7.16	8.4	68.5	0.73	0.44	0.73	30	1	1	1	0	0	31	3	3	1	Si	1	3	37	obrePes	2	1	1	1
164	0	M	1	36.17	14	92.5	0.56	-1	-0.15	0	0	0	0	0	0	22	2	5	2	Si	1	3	37	Normal	1	1	4	0
165	1	M	1	18	9.2	72.3	0.35	-3.7	-1.55	30	1	0	0	0	0	32	2	3	1	Si	1	2	37	obrePes	2	2	3	0
166	1	F	0	25.46	11	82	0.03	-1.5	-0.78	30	1	1	0	0	0	33	3	3	1	Si	1	3	40	obrePes	2	1	1	1
167	0	F	0	6.01	7.1	68	-0.97	1.02	-0.22	0	0	0	0	0	0	22	2	4	1	Si	1	3	39	Normal	1	1	3	0
168	1	M	1	25	12	85	0.49	-1	-0.14	0	0	0	0	0	0	33	3	3	1	Si	1	2	41	Deficit	0	1	1	1
169	0	F	0	30.03	13	91	0.4	0.08	0.4	0	0	0	0	0	0	22	1	3	1	Si	1	3	36	Normal	1	1	3	0
170	0	F	0	48.16	15	95	0.84	-1.8	-0.51	0	0	0	0	0	0	22	2	4	1	Si	1	3	38	Normal	1	1	2	0
171	1	F	0	10.12	7.2	67	-0.43	-1.9	-1.34	30	1	0	0	0	0	33	3	3	1	Si	1	1	40	Deficit	0	1	1	1
172	0	M	1	57.07	18	101.5	1.48	-1.5	0.02	0	0	0	0	0	0	23	1	4	1	Si	1	3	37	Normal	1	1	4	0
173	1	F	0	30.09	15	91	1.29	0.08	1.02	30	1	0	0	0	0	34	4	3	1	Si	1	2	37	Deficit	0	1	3	1
174	1	F	0	18.53	12	80	1.61	-0.4	0.99	30	1	1	0	0	0	34	4	3	1	Si	1	3	40	obrePes	2	1	4	1
175	1	F	0	6.01	8.3	66.8	1.16	0.49	1.1	30	1	1	0	0	0	35	4	3	2	No	0	2	37	obrePes	2	1	2	1
176	0	F	0	42.18	14	95	0.09	-1	-0.52	0	0	0	0	0	0	23	2	4	1	Si	1	4	36	Normal	1	1	3	0
177	0	F	0	48.76	16	98	0.94	-1.2	-0.09	0	0	0	0	0	0	23	2	5	1	Si	1	3	39	Normal	1	1	3	0
178	1	M	1	6.05	9.6	70	1.54	1.08	1.73	30	1	1	0	0	0	36	3	3	1	Si	1	2	38	obrePes	2	1	2	1
179	1	M	1	30.13	12	87	0.04	-1.5	-0.75	30	1	0	0	0	0	40	3	3	1	Si	1	3	38	obrePes	2	1	2	1
180	0	M	1	27.7	13	89	0.71	-0.4	0.34	30	1	0	0	0	0	23	2	4	1	Si	1	3	40	Normal	1	1	1	0
181	1	F	0	12.19	9.3	73	0.65	-0.5	0.27	0	0	0	0	0	0	40	3	4	2	Si	1	3	39	obrePes	2	1	1	1
182	0	M	1	2.07	5.8	60	-0.33	0.67	0.28	0	0	0	0	0	0	23	2	4	1	Si	1	3	38	Normal	1	1	1	0
183	1	F	0	14.85	10	77	0.76	-0.1	0.52	30	1	0	0	0	0	16	2	2	1	Si	1	2	36	obrePes	2	1	3	1
184	0	F	0	18.96	10	81	-0.16	-0.2	-0.22	30	1	1	0	0	0	23	2	4	2	Si	1	3	36	Normal	1	1	2	0
185	0	M	1	42.12	13	92	-0.34	-2	-1.36	0	0	1	0	0	0	23	2	4	1	Si	1	3	38	Normal	1	1	3	0
186	1	M	1	11.99	8.7	75	-1.08	-0.3	-0.93	0	0	0	0	0	0	18	2	3	1	Si	1	2	39	obrePes	2	1	2	1
187	0	M	1	36.01	13	91.5	-0.13	-1.2	-0.75	0	0	0	0	0	0	23	2	4	1	Si	1	3	38	Normal	1	1	2	0
188	1	M	1	8.97	8.2	70	-0.33	-0.9	-0.74	30	1	0	0	0	0	20	1	3	1	Si	1	2	40	obrePes	2	1	4	1
189	1	F	0	12.09	10	76	0.92	0.74	-1	0	0	0	0	0	0	21	2	3	1	Si	1	2	39	obrePes	2	1	1	1
190	0	M	1	24.18	11	87	-1.15	-0.1	-0.81	30	1	0	0	0	0	23	2	4	1	Si	1	3	38	Normal	1	1	1	0
191	1	M	1	6.21	7.6	66	0.16	-0.9	-0.48	0	0	0	0	0	0	21	1	3	1	Si	1	2	38	obrePes	2	1	1	1
192	1	F	0	11.01	9.4	75.5	0.13	1.08	0.57	30	1	0	0	0	0	21	2	3	1	Si	1	2	37	obrePes	2	2	3	1
193	1	M	1	32.99	13	90	-0.28	-1.1	-0.78	0	0	1	0	0	0	22	1	4	1	Si	1	3	39	obrePes	2	1	4	1
194	1	M	1	7.03	9.4	70	1.31	0.39	1.15	30	1	1	0	0	0	22	3	3	1	Si	1	2	38	obrePes	2	1	3	1
195	1	F	0	6.08	9.7	67	2.6	0.5	2.23	30	1	0	0	0	0	22	2	3	1	Si	1	2	38	obrePes	2	1	2	1
196	1	M	1	4.07	6.9	64	-0.22	0.01	-0.16	0	0	0	0	0	0	22	1	4	1	Si	1	3	38	obrePes	2	1	1	1
197	1	M	1	12.48	10	77.5	0.31	0.51	0.46	30	1	0	0	0	0	23	1	3	2	Si	1	2	37	obrePes	2	1	4	1
198	0	F	0	23.95	11	83	-0.26	-1	-0.72	30	1	0	0	0	0	23	2	4	2	Si	1	3	39	Normal	1	1	1	0
199	1	M	1	59.96	20	107.1	1.5	-0.6	0.64	0	0	0	0	0	0	23	2	4	1	Si	1	2	38	obrePes	2	1	1	1
200	0	M	1	12.12	9	73	-0.12	-1.2	-0.65	30	1	1	0	0	0	23	2	4	1	Si	1	4	37	Normal	1	1	2	0
201	0	F	0	11.99	9	74.5	-0.07	0.21	0.06	30	1	1	0	0	0	23	2	4	1	Si	1	2	36	Deficit	0	1	4	0
202	1	F	0	4.04	7	64.2	0.16	0.97	0.69	0	0	0	0	0	0	24	2	3	1	Si	1	3	40	obrePes	2	1	2	1
203	1	F	0	12.09	9.6	76	0.31	0.74	0.56	0	0	0	0	0	0	15	2	3	1	Si	1	2	35	Deficit	0	1	4	1
204	0	F	0	6.41	7.9	64	1.49	-1	0.47	30	1	1	0	0	0	23	2	4	1	Si	1	2	39	Deficit	0	1	1	0
205	1	M	1	6.08	9.1	73	0.02	2.45	1.23	30	1	1	0	0	0	24	2	4	1	Si	1	2	36	Deficit	0	1	4	1
206	1	M	1	13.34	11	75	1.29	-0.9	0.58	30	1	1	1	0	0	24	2	3	1	Si	1	2	40	Deficit	0	1	2	1
207	1	M	1	5.98	9.6	76.5	-0.31	4.16	1.72	0	0	0	0	0	0	25	2	3	1	Si	1	3	40	obrePes	2	1	2	1
208	0	M	1	49.45	16	100.4	0.41	-0.9	-0.28	0	0	0	0	0	0	23	2	4	2	Si	1	2	37	Deficit	0	1	3	0
209	1	M	1	30.09	13	89.5	0.41	-0.7	-0.07	0	0	1	0	0	0	25	2	3	1	Si	1	3	38	obrePes	2	1	1	1
210	0	M	1	31.9	14	88.2	0.98	-1.5	-0.08	0	0	0	0	0	0	24	2	4	2	Si	1	3	39	Normal	1	1	1	0
211	1	M	1	18.04	10	78.8	-0.13	-1.3	-0.71	30	1	1	0	0	0	25	2	5	2	Si	1	3	40	obrePes	2	1	3	1
212	0	M	1	48.56	18	102.5	1.06	-0.3	0.53	0	0	0	0	0	0	24	2	5	1	Si	1	4	36	Normal	1	1	2	0
213	1	F	0	14.06	8.8	73.8	-0.18	-1	-0.55	30	1	1	0	0	0	25	2	5	2	Si	1	3	38	obrePes	2	1	1	1
214	1	M	1	5.03	7.6	66.5	0	0.26	0.12	0	1	0	0	0	0	25	2	3	1	Si	1	3	39	obrePes	2	1	1	1
215	1	F	0	30.75	14	91	0.51	-0.1	0.39	28	1	0	0	0	0	26	2	3	3	Si	1	3	40	obrePes	2	2	1	1
216	0	M	1	30.13	11	86.2	-0.83	-1.7	-1.48	30	1	0	0	0	0	24	2	5	1	Si	1	4	40	Normal	1	1	2	0
217	0	M	1	55.1	16	103.2	0.08	-0.9	-0.51	0	0	0	0	0	0	24	2	5	2	Si	1	3	37	Normal	1	1	2	0
218	1	M	1	5.03	7.4	65.2	0.06	-0.4	-0.21	0	1	1	0	0	0	26	2	4	1	Si	1	3	41	obrePes	2	1	2	1
219	0	F	0	22.01	12	83	1.22	-0.5	0.66	0	0	0	0	0	0	24	2	5	2	Si	1	3	36	Normal	1	1	4	0
220	0	F	0	30.03	13	91.5	-0.1	0.23	0.12	0	0	1	0	0	0	24	2	5	1	Si	1	3	38	Normal	1	1	2	0
221	1	F	0	6.08	7.8	66	0.71	0.06	0.51	30	1	0	0	0	0	27	3	3	1	Si	1	2	36	obrePes	2	1	1	1
222	1	F	0	12.45	8.4	73.2	-0.53	-0.5	-0.6	30	1	1	0	0	0	27	3	3	3	Si	1	3	36	obrePes	2	1	2	1
223	1	M	1	30.03	12	84.7	0.42	-2.1	-0.88	30	1	0	0	0	0	27	3	3	1	Si	1	3	38	obrePes	2	1		



241	0	F	0	27.66	12	88	-0.14	-0.2	-0.15	30	1	0	0	0	26	2	4	3	Si	1	4	37	Normal	1	2	1	0
242	1	F	0	6.14	9	68	1.61	0.92	1.66	30	1	0	0	0	29	2	3	3	Si	1	2	40	bbrePes	2	1	3	1
243	0	F	0	13.86	11	78	1.37	0.68	1.3	0	0	0	0	0	26	2	5	3	Si	1	3	38	Normal	1	1	3	0
244	0	F	0	43.63	13	92	0.05	-2	-1.1	0	0	0	0	0	26	3	4	1	Si	1	3	38	Normal	1	1	2	0
245	1	F	0	6.21	7.2	63	0.91	-1.4	-0.2	30	1	1	0	0	29	3	3	1	Si	1	3	36	bbrePes	2	1	3	1
246	0	M	1	6.01	9.8	69	1.12	0.66	1.17	0	0	0	0	0	27	3	4	1	Si	1	4	36	Normal	1	1	3	0
247	1	M	1	48.53	15	98	0.15	-1.3	-0.72	0	1	0	0	0	30	3	4	1	Si	1	3	36	bbrePes	2	1	2	1
248	0	F	0	15.11	8.3	73	-0.58	-1.7	-1.2	28	1	0	0	0	27	3	5	1	Si	1	2	38	Deficit	0	1	4	0
249	1	F	0	38.44	16	96	1.47	-0.2	0.93	0	1	0	0	0	30	2	3	4	Si	1	3	38	bbrePes	2	1	1	1
250	1	M	1	30	12	87.5	-0.14	-1.3	-0.77	30	1	0	0	0	34	2	3	1	Si	1	3	37	bbrePes	2	1	4	1
251	1	F	0	6.51	8.4	67.8	0.87	0.55	0.9	30	1	1	0	0	36	3	5	2	Si	1	3	40	bbrePes	2	1	2	1
252	0	M	1	19.42	11	81	-0.15	-1	-0.6	0	0	0	0	0	27	3	5	1	Si	1	2	33	Deficit	0	1	4	0
253	1	M	1	23.98	12	87.9	-0.43	0.04	-0.25	0	1	0	0	0	38	3	5	1	Si	1	2	39	bbrePes	2	2	2	1
254	1	F	0	7.03	7.3	67	-0.4	-0.1	-0.42	30	1	1	0	0	39	3	4	2	Si	1	3	39	bbrePes	2	1	4	1
255	1	F	0	4.14	6.7	60.4	1.15	-0.9	0.19	0	1	0	0	0	41	3	3	2	Si	1	3	37	bbrePes	2	1	3	1
256	1	M	1	24.02	11	84.3	0.1	-0.9	-0.41	0	1	0	0	0	17	2	3	1	Si	1	1	36	bbrePes	2	1	3	1
257	1	F	0	29.93	10	85	-0.71	-1.6	-1.39	30	1	0	0	0	17	2	2	1	Si	1	1	38	bbrePes	2	1	1	1
258	1	F	0	12.06	9.5	74	1.19	-0	0.88	30	1	0	0	0	19	1	3	1	Si	1	1	40	Deficit	0	1	4	1
259	1	M	1	9.99	7.5	69	-0.85	-1.9	-1.64	30	1	0	0	0	21	2	2	1	Si	1	1	36	Deficit	0	1	3	1
260	1	M	1	18	9.6	79.5	-0.92	-1	-1.17	30	1	1	0	0	22	2	3	1	Si	1	2	42	Deficit	0	1	1	1
261	1	M	1	12.55	11	76	1.48	-0.2	1.09	30	1	0	0	0	23	2	3	1	Si	1	1	37	bbrePes	2	1	1	1
262	0	F	0	18.04	12	80	1.88	-0.3	1.26	30	1	0	0	0	27	3	4	1	Si	1	2	40	Deficit	0	1	4	0
263	1	M	1	33.51	11	87.9	-0.93	-1.8	-1.6	0	1	0	0	0	23	2	3	1	Si	1	1	39	bbrePes	2	1	3	1
264	1	M	1	12.35	9.5	79.6	-1.08	1.45	-0.21	30	1	0	0	0	23	2	3	1	Si	1	1	40	bbrePes	2	1	4	1
265	1	F	0	37.72	11.8	88	0.14	-2.1	-1.11	30	1	1	0	0	24	2	3	1	Si	1	1	39	bbrePes	2	1	1	1
266	1	F	0	11.96	11	75.6	1.82	0.64	1.64	30	1	1	0	0	24	2	3	1	Si	1	2	39	bbrePes	2	1	2	1
267	0	M	1	33.58	19	91.5	4.4	-0.8	2.57	0	0	0	0	0	27	3	5	1	Si	1	3	40	Normal	1	1	2	0
268	0	M	1	22.21	13	85.7	-0.62	-0.2	-0.53	0	0	1	0	0	28	3	4	1	Si	1	3	40	Normal	1	1	1	0
269	1	M	1	18.04	10	78.4	-0.09	-1.4	-0.74	30	1	0	0	0	25	2	4	2	No	0	4	40	Normal	1	1	2	1
270	1	F	0	18.04	10	78.8	0.6	-0.7	0.13	0	1	1	0	0	25	2	4	1	Si	1	4	40	bbrePes	2	1	2	1
271	1	F	0	12.06	7.8	69.2	0	-1.9	-0.94	30	1	0	0	0	25	2	3	2	Si	1	1	40	bbrePes	2	1	4	1
272	0	M	1	33.68	13	90.5	-1.01	-1.1	-1.26	30	1	1	0	0	28	3	4	1	Si	1	4	38	Normal	1	1	1	0
273	1	M	1	12.02	9.8	74.8	0.82	-0.4	0.46	0	1	1	0	0	25	2	3	1	Si	1	1	39	bbrePes	2	1	2	1
274	0	F	0	52.04	17	104	0.54	-0.2	0.21	30	1	1	0	0	29	3	4	3	Si	1	3	40	Normal	1	1	3	0
275	1	F	0	36.01	13	88	0.67	-1.9	-0.56	0	1	1	1	0	25	2	4	1	Si	1	4	39	bbrePes	2	1	3	1
276	1	F	0	36.17	13.9	92.5	0.67	-0.7	0.12	0	1	1	0	0	25	2	3	1	Si	1	1	38	bbrePes	2	1	4	1
277	0	F	0	42.18	14	91.1	1.12	-2	-0.36	0	0	0	0	0	29	3	5	3	No	0	3	39	Normal	1	1	3	0
278	1	F	0	25.69	10	84.3	-1.09	-0.9	-1.22	30	1	1	0	0	25	2	3	1	Si	1	3	38	bbrePes	2	1	3	1
279	0	F	0	11.99	10	76	1.13	0.79	1.18	30	1	0	0	0	29	2	4	2	Si	1	4	40	Normal	1	1	4	0
280	0	M	0	36.14	16	93	1.92	-0.6	1.08	0	0	0	0	0	30	3	4	4	Si	1	3	38	Normal	1	1	4	0
281	0	M	1	21.88	12	82	0.92	-1.3	-0.01	0	0	0	0	0	31	2	5	4	Si	1	3	38	Normal	1	1	1	0
282	1	M	1	44.42	10	93	-3.4	-2	-3.31	0	1	1	0	0	26	2	3	2	Si	1	2	38	Deficit	0	1	4	1
283	0	M	0	48.16	17	100.5	1.27	-0.5	0.51	0	0	0	0	0	32	2	5	1	Si	1	3	38	Normal	1	1	2	0
284	0	M	1	6.01	8.2	76.5	-2.23	4.16	0.3	30	1	1	0	0	32	2	5	4	Si	1	3	40	Normal	1	1	3	0
285	0	F	0	6.14	6.9	64.4	-1.65	-0.7	-1.66	0	0	1	0	0	32	2	4	3	Si	1	3	40	Normal	1	1	2	0
286	1	F	0	5.85	6.9	64	0.08	-0.7	-0.39	30	1	0	0	0	26	2	3	2	Si	1	2	36	Deficit	0	1	4	1
287	1	F	0	12.65	9.4	74.4	0.43	-0.2	0.27	30	1	0	0	0	26	3	3	2	Si	1	2	38	Deficit	0	1	1	1
288	1	F	0	18	9.3	81.5	-1.28	0.28	-0.78	25	1	0	0	0	26	2	3	1	Si	1	3	40	bbrePes	2	1	1	1
289	1	M	1	6.01	7.3	68	-0.58	0.19	-0.38	0	1	0	0	0	26	2	3	1	Si	1	1	39	bbrePes	2	1	3	1
290	1	M	1	53.95	16.6	101	0.92	-1.3	-0.19	30	1	0	0	0	26	2	3	1	Si	1	1	38	bbrePes	2	1	2	1
291	1	M	1	12.09	9.7	73	0.72	-1.2	-0.01	30	1	0	0	0	26	2	4	1	Si	1	4	40	bbrePes	2	1	1	1
292	0	M	1	43.33	13	92	-0.34	-2.2	-1.45	0	0	0	0	0	32	2	5	4	Si	1	3	38	Normal	1	1	3	0
293	1	F	0	29.9	11.1	86.7	-0.21	-1.1	-0.72	0	1	0	0	0	26	2	3	1	Si	1	1	40	bbrePes	2	1	1	1
294	1	M	1	5.06	7.6	68	-0.66	0.94	0.01	0	1	0	0	0	26	2	5	1	Si	1	2	37	bbrePes	2	1	2	1
295	0	M	1	18.1	12	83	0.63	0.25	0.55	0	0	0	0	0	32	4	5	1	Si	1	3	38	Normal	1	1	4	0
296	1	F	0	2.04	5.3	55.5	1.23	-0.8	0.14	0	1	0	0	0	26	2	4	1	Si	1	4	37	bbrePes	2	1	3	1
297	1	M	1	6.01	7.7	70.2	-0.1	1.22	0.53	30	1	0	0	0	26	2	3	2	Si	1	1	37	bbrePes	2	1	4	1
298	1	M	1	30.29	13	86.5	0.84	-1.6	-0.29	30	1	1	0	0	26	3	3	3	Si	1	1	40	bbrePes	2	1	1	1
299	1	M	1	6.05	7.5	65.4	0.23	-1.1	-0.53	30	1	1	0	0	26	3	3	1	Si	1	2	40	bbrePes	2	1	2	1
300	1	M	1	7.23	7.7	67.8	-0.43	-0.8	-0.82	30	1	0	0	0	26	3	5	3	Si	1	4	39	bbrePes	2	1	3	1
301	0	M	1	30.09	14	92	0.41	0.01	0.31	0	0	1	0	0	32	2	5	4	Si	1	3	40	Normal	1	1	1	0
302	1	F	0	30.69	11	86	-0.29	-1.5	-1	0	1	1	0	0	27	3	3	1	Si	1	3	41	bbrePes	2	1	3	1
303	0	M	1	24.64	12	84	0.59	-1.2	-0.2	30	1	0	0	0	33	2	5	1	Si	1	2	40	Normal	1	1	4	1
304	1	F	0	12.65	9.6	70.4	1.63	-1.7	0.44	30	1	0	0	0	27	3											



321	0	M	1	20.17	11	81.2	-0.13	-1.1	-0.66	0	0	0	0	0	39	3	4	1	No	0	3	39	Normal	1	1	2	0
322	0	M	1	36.21	14	94	-0.02	-0.6	-0.34	0	0	1	0	0	40	5	4	2	Si	1	3	40	Normal	1	1	1	0
323	1	M	1	30	13	89	0.15	-0.9	-0.33	30	1	1	0	0	31	4	3	1	Si	1	1	37	pbrePes	2	1	4	1
324	1	M	1	25.69	13	87.2	0.38	-0.4	0.08	30	1	1	0	0	31	3	5	2	Si	1	4	37	pbrePes	2	1	2	1
325	1	M	1	24.08	12	84	0.48	-1	-0.2	30	1	1	0	0	31	2	5	4	Si	1	3	36	pbrePes	2	1	2	1
326	0	M	1	42.05	16	101.3	0.5	0.36	0.53	0	0	1	0	0	18	2	3	1	Si	1	2	38	Normal	1	1	2	0
327	0	M	1	27.1	10	88.6	-2.42	-0.3	-1.76	0	0	0	0	0	18	1	2	1	Si	1	2	36	Normal	1	1	1	0
328	1	M	1	18.07	9.7	80	0.29	-0.9	-0.21	30	1	0	0	0	32	2	3	3	Si	1	1	37	pbrePes	2	1	2	1
329	1	M	1	24.05	13	85.5	0.9	-0.5	0.38	30	1	0	0	0	32	2	3	1	Si	1	2	32	pbrePes	2	2	3	1
330	0	M	1	25.03	12	83	0.83	-1.6	-0.25	0	0	0	0	0	18	2	3	1	Si	1	2	36	Normal	1	1	3	0
331	1	M	1	24.11	11	85	-0.73	-0.7	-0.88	30	1	0	0	0	33	3	4	1	Si	1	4	36	pbrePes	2	1	3	1
332	1	M	1	36.11	14	88	1.46	-2.2	-0.21	30	1	0	0	0	33	3	3	1	Si	1	1	40	pbrePes	2	2	4	1
333	1	M	1	48.89	11	82.5	-0.37	-5.1	-3.39	30	1	0	0	0	33	3	4	4	Si	1	4	40	pbrePes	2	1	1	1
334	0	M	1	38.44	12	92	-1.25	-1.5	-1.63	30	1	1	0	0	18	2	3	1	Si	1	2	39	Normal	1	1	1	0
335	1	M	1	24.77	13	87	0.33	-0.2	0.14	30	1	0	0	0	33	4	3	1	Si	1	1	37	pbrePes	2	1	4	1
336	1	F	0	18.1	12	82.2	1.36	0.49	1.22	30	1	1	0	0	33	5	3	1	Si	1	2	38	pbrePes	2	1	1	1
337	1	F	0	42.25	13.9	97	0.15	-0.5	-0.21	30	1	1	0	0	34	2	3	1	Si	1	1	40	pbrePes	2	1	1	1
338	0	M	1	36.73	12	87	-0.17	-2.6	-1.55	0	0	0	0	0	18	2	3	1	Si	1	2	38	Normal	1	1	3	0
339	1	F	0	20.11	9.9	78.5	0.15	-1.4	-0.59	30	1	1	0	0	34	3	4	1	Si	1	4	39	pbrePes	2	1	4	1
340	0	F	0	30.06	12	86.5	0.22	-1.2	-0.46	0	0	0	0	0	19	2	3	1	Si	1	2	38	Normal	1	1	1	0
341	0	F	0	48.26	18	106.5	0.43	0.84	0.78	0	0	0	0	0	20	1	4	1	Si	1	3	40	Normal	1	1	1	0
342	1	F	0	25.2	12	84	0.91	-0.8	0.25	30	1	0	0	0	34	3	2	2	Si	1	2	38	Normal	1	1	3	1
343	1	M	1	24.71	14	86.5	1.81	-0.4	1.13	30	1	1	0	0	34	2	4	1	Si	1	4	38	pbrePes	2	1	3	1
344	1	M	1	17.15	10	78.9	-0.07	-0.9	-0.49	30	1	0	0	0	34	1	3	1	Si	1	2	39	pbrePes	2	1	4	1
345	1	M	1	24.21	10	81	-0.43	-2.1	-1.39	30	1	0	0	0	35	4	3	1	Si	1	2	39	Deficit	0	1	3	1
346	0	M	1	41.92	16	103	-0.17	0.81	0.34	0	0	0	0	0	20	2	5	1	Si	1	4	36	Normal	1	1	2	0
347	1	F	0	30.13	11	90.5	-1.22	-0.1	-0.83	30	1	0	0	0	35	2	3	2	No	0	2	37	Deficit	0	1	3	1
348	0	M	1	25.36	12	85	0.54	-1	-0.15	0	0	0	0	0	21	2	4	1	Si	1	3	39	Normal	1	1	4	0
349	1	F	0	19.22	11	76	1.53	-2	0.2	30	1	0	0	0	35	3	4	2	Si	1	4	42	Deficit	0	1	1	1
350	1	F	0	17.97	14	75	4.59	-2	2.44	30	1	0	0	0	35	4	3	1	Si	1	2	40	pbrePes	2	1	4	1
351	1	M	1	19.06	12	83	0.83	-0.1	0.53	30	1	0	0	0	36	3	3	2	Si	1	1	39	pbrePes	2	1	4	1
352	1	M	1	27.27	10.6	83	0.93	-2.1	-0.48	30	1	1	0	0	36	2	2	2	Si	1	1	40	pbrePes	2	1	4	1
353	1	F	0	18.07	10	81	0.06	0.08	0.08	30	1	1	0	0	36	2	3	2	Si	1	2	38	pbrePes	2	1	1	1
354	1	M	1	36.34	15	95.3	0.32	-0.3	0.06	30	1	0	0	0	36	3	3	2	Si	1	1	39	pbrePes	2	1	1	1
355	1	M	1	30.19	13	92.5	-0.07	0.13	0.05	30	1	1	0	0	36	3	3	2	Si	1	1	37	pbrePes	2	2	4	1
356	0	F	0	34.1	13	94.8	-0.36	0.3	-0.05	30	1	0	0	0	21	1	3	1	Si	1	2	40	Normal	1	1	1	0
357	0	F	0	27.04	11	82	0.34	-1.9	-0.78	0	0	0	0	0	21	2	4	1	Si	1	4	37	Normal	1	1	4	0
358	1	M	1	18.37	10	81	-0.45	-0.6	-0.64	30	1	1	0	0	36	4	4	1	Si	1	4	41	pbrePes	2	1	2	1
359	1	M	1	30.09	12	86.3	-0.09	-1.7	-0.95	30	1	1	1	0	37	3	3	1	Si	1	2	40	pbrePes	2	1	1	1
360	0	M	1	36.04	14	95	0.21	-0.3	-0.03	0	0	0	0	0	21	1	5	1	Si	1	3	39	Normal	1	2	3	0
361	1	F	0	24.15	10.4	84	0.05	-0.6	-0.22	30	1	1	0	0	38	2	2	1	Si	1	1	40	pbrePes	2	1	4	1
362	1	M	1	20.93	11	85.2	-0.51	0.05	-0.36	30	1	1	0	0	38	3	3	4	Si	1	2	40	pbrePes	2	1	3	1
363	0	M	1	17.71	11	79	0.66	-1.1	-0.02	30	1	0	0	0	23	3	4	1	Si	1	3	37	Normal	1	1	4	0
364	1	M	1	18.04	9.9	76.3	0.17	-2.2	-0.9	30	1	1	0	0	39	2	2	1	Si	1	2	37	pbrePes	2	1	2	1
365	0	M	1	46.39	17	107.5	-0.3	1.23	0.53	0	0	0	0	0	23	2	4	2	Si	1	3	40	Normal	1	1	4	0
366	1	F	0	34.76	13	91	0.01	-0.9	-0.42	30	1	0	0	0	39	3	4	1	Si	1	4	40	pbrePes	2	1	3	1
367	0	M	1	25.03	12	85.5	0.41	-0.8	-0.11	30	1	0	0	0	23	2	4	1	Si	1	4	37	Normal	1	1	3	0
368	1	M	1	19.02	9.4	78.3	-0.97	-1.8	-1.56	30	1	1	0	0	40	3	3	2	Si	1	2	38	pbrePes	2	1	2	1
369	1	F	0	18.4	9.9	76.2	0.61	-1.7	-0.33	30	1	1	0	0	41	3	3	2	No	0	2	32	pbrePes	2	1	1	1
370	0	F	0	25	12	88.6	-0.37	0.62	0.15	0	0	0	0	0	23	2	4	2	Si	1	2	36	Deficit	0	1	1	0
371	1	M	1	18.04	14	85	2.4	1	2.28	30	1	0	0	0	41	3	5	2	Si	1	4	40	pbrePes	2	1	4	1
372	1	M	1	20.04	12	81.1	1.5	-1.1	0.57	30	1	0	0	0	43	3	3	1	Si	1	2	37	pbrePes	2	1	3	1
373	1	F	0	4.04	6.4	62	0.04	-0.1	-0.04	30	1	1	0	0	23	2	3	1	Si	1	2	38	pbrePes	2	2	2	1
374	1	F	0	28.68	11	84.4	0.06	-1.5	-0.78	30	1	0	0	0	24	3	3	1	Si	1	2	40	pbrePes	2	1	2	1
375	0	F	0	24.28	10	82	-0.4	-1.2	-0.93	0	0	0	0	0	24	2	5	1	Si	1	2	36	Deficit	0	1	1	0
376	1	M	1	24.08	11	82.2	0.09	-1.6	-0.8	30	1	1	0	0	24	2	3	1	Si	1	2	38	Deficit	0	1	2	1
377	1	F	0	24.08	11	83	0.1	-0.9	-0.35	30	1	0	0	0	24	2	3	1	Si	1	2	40	Deficit	0	1	2	1
378	1	M	1	47.05	14	94.5	0.48	-2	-0.88	30	1	0	0	0	26	2	3	1	Si	1	2	38	Deficit	0	1	2	1
379	1	M	1	36.8	13	89.5	0.32	-1.9	-0.83	30	1	0	0	0	26	2	4	1	Si	1	4	38	pbrePes	2	1	4	1
380	0	F	0	19.15	13	89.8	0.25	2.69	1.53	30	1	0	0	0	24	2	4	2	Si	1	2	40	Deficit	0	1	4	0
381	0	F	0	36.24	13	89	0.53	-1.6	-0.52	0	0	0	0	0	24	3	4	1	No	0	2	40	Normal	1	1	4	0
382	1	M	1	28.78	13	90	-0.18	-0.3	-0.25	30	1	1	0	0	26	2	3	1	Si	1	2	38	pbrePes	2	1	3	1
383	1	M	1	24.11	11	81.1	-0.1	-2	-1.13	30	1	0	0	0	26	2	3	1	Si	1	2	38	pbrePes	2	1	2	1
384	1	M	1	19.29	12	81	0.93	-0.9	0.24	30	1	1	0	0													



ANEXOS 02. VALIDACION POR JUICIO DE EXPERTOS

VALIDACION POR PROFESIONAL DE LA RED DE SALUD PUNO

I DATOS GENERALES

- 1.1 Título de investigación: MODELO LOGIT PARA IDENTIFICAR LOS FACTORES ASOCIADOS EN LA PREVALENCIA DE ANEMIA EN MENORES DE 3 A 59 MESES DEL CENTRO DE SALUD I-3 VALLECITO PUNO-2021.
- 1.2 Nombre de los instrumentos motivo de evaluación: ENCUESTA SOBRE VALIDEZ

II ASPECTOS DE VALIDACION

INDICADORES	CRITERIOS	DEFICIENTE	REGULAR	BUENO	MUY BUENO
		0-5	6-10	11-15	16-20
1. Claridad	Esta formulado con lenguaje apropiado				19
2. Objetividad	Esta expresado en conductas observables				18
3. Organización	Existe una organización lógica				20
4. Intencionalidad	Adecuado para valorar los instrumentos de investigación				18
5. Consistencia	Basado en aspectos teóricos científicos				20
6. Coherencia	Entre lo Items e Indicadores				20
7. Metodología	La estrategia responde al propósito del diagnóstico				20
8. Pertinencia	Es útil y adecuado para la investigación				20

155

OPINION DE APLICABILIDAD

a) Deficiente b) Regular c) Bueno d) Muy bueno

Nombres y apellidos	Zuani Mariuz Pineda Ramos		
Título profesional	Enfermera		
Grado académico	Licenciada en Enfermería		
Mención	Gestión y garantía de los servicios de salud		
Lugar y fecha	Puno, enero 2021		
DNI	73811985	TLF/CEL:	930596608


Zuani Mariuz Pineda Ramos
LIC. EN ENFERMERIA
CEP. 06578

SELLO Y FIRMA DE VALIDADOR



I DATOS GENERALES

1.1 Título de investigación: MODELO LOGIT PARA IDENTIFICAR LOS FACTORES ASOCIADOS EN LA PREVALENCIA DE ANEMIA EN MENORES DE 3 A 59 MESES DEL CENTRO DE SALUD I-3 VALLECITO PUNO-2021.

1.2 Nombre de los instrumentos motivo de evaluación: ENCUESTA SOBRE VALIDEZ

II ASPECTOS DE VALIDACION

INDICADORES	CRITERIOS	DEFICIENTE	REGULAR	BUENO	MUY BUENO
		0-5	6-10	11-15	16-20
1. Claridad	Esta formulado con leguaje apropiado				20
2. Objetividad	Esta expresado en conductas observables				20
3. Organización	Existe una organización lógica				20
4. Intencionalidad	Adecuado para valorar los instrumentos de investigación				18
5. Consistencia	Basado en aspectos teóricos científicos				18
6. Coherencia	Entre lo Items e Indicadores				20
7. Metodología	La estrategia responde al propósito del diagnóstico				20
8. Pertinencia	Es útil y adecuado para la investigación				20

156

OPINION DE APLICABILIDAD

a) Deficiente b) Regular c) Bueno d) Muy bueno

Nombres y apellidos	RENE ERRAIN COLCA CHAVEZ
Título profesional	MEDICO CIRUJANO
Grado académico	MEDICO
Mención	PEDIATRIA
Lugar y fecha	PUNO, ENERO 2021
DNI 80094110	TLF/CEL: 926 55 6050


Dr. Rene Errain Colca Chavez
MEDICO CIRUJANO
CMP: 82077

FIRMA DE VALIDADOR



I DATOS GENERALES

1.1 Título de investigación: MODELO LOGIT PARA IDENTIFICAR LOS FACTORES ASOCIADOS EN LA PREVALENCIA DE ANEMIA EN MENORES DE 3 A 59 MESES DEL CENTRO DE SALUD I-3 VALLECITO PUNO-2021.

1.2 Nombre de los instrumentos motivo de evaluación: ENCUESTA SOBRE VALIDEZ

II ASPECTOS DE VALIDACION

INDICADORES	CRITERIOS	DEFICIENTE	REGULAR	BUENO	MUY BUENO
		0-5	6-10	11-15	16-20
1. Claridad	Esta formulado con leguaje apropiado				20
2. Objetividad	Esta expresado en conductas observables				19
3. Organización	Existe una organización lógica				19
4. Intencionalidad	Adecuado para valorar los instrumentos de investigación				20
5. Consistencia	Basado en aspectos teóricos científicos				20
6. Coherencia	Entre lo Items e Indicadores				20
7. Metodología	La estrategia responde al propósito del diagnóstico				19
8. Pertinencia	Es útil y adecuado para la investigación				20

157

OPINION DE APLICABILIDAD

a) Deficiente b) Regular c) Bueno d) Muy bueno

Nombres y apellidos	Olga Velásquez Mamani
Título profesional	Enfermera
Grado académico	Licenciada en Enfermería
Mención	Enfermería En el Cuidado Materno Infantil
Lugar y fecha	Puno, enero 2021
DNI 01322489	TLF/CEL: 973254197

Olga Velásquez Mamani
LIC. EN ENFERMERÍA
C.E.P. 35635

SELLO Y FIRMA DE VALIDADOR



ANEXOS 03. CUESTIONARIO REALIZADO A MADRES QUE ASISTIERON AL
EESS VALLECITO.



UNIVERSIDAD NACIONAL DEL ALTIPLANO

MODELO LOGIT PARA IDENTIFICAR FACTORES
ASOCIADOS EN LA PREVALENCIA DE ANEMIA EN
MENORES DE 3 A 59 MESES DEL CENTRO DE SALUD I-3
VALLECITO-PUNO

FICHA DE RECOLECCION DE DATOS

El presente instrumento es parte del estudio de mi investigación.
Yo Rosa Angélica Payehuanca Marca, bachiller de la
Universidad Nacional del Altiplano, solicito de su colaboración
sincera y veraz para responder las siguientes preguntas
expresándole que es de carácter anónimo y confidencial.
Responda por favor con mucha sinceridad y veracidad los
siguientes ítems, ya que esto permitirá cumplir los objetivos de
mi investigación.

Agradezco anticipadamente su colaboración.

DATOS DEL NIÑO:

- 1.- NUMERO DE HISTORIA CLINICA
.....
- 2.- NUMERO DE DNI
.....
- 3.- SEXO
.....



DATOS DE LA MADRE:

FACTORES SOCIODEMOGRAFICOS

1.- EDAD.....

2.- CUENTA CON SIS

- 1. SI
- 2. NO

1.- CONDICION CIVIL:

- 1. SOLTERA
- 2. CONVIVIENTE
- 3. CASADA
- 4. DIVORCIADA
- 5. SEPARADA

2.- NIVEL DE INSTRUCCIÓN:

- 1. SIN INSTRUCCIÓN
- 2. PRIMARIA
- 3. SECUNDARIA
- 4. SUPERIOR NO UNIVERSITARIA
- 5. SUPERIOR UNIVERSITARIA

3.- INGRESO PROMEDIO MENSUAL FAMILIAR

- 1. MENOS DE 1000.00 SOLES
- 2. ENTRE 1000 A 3000.00 SOLES
- 3. ENTRE 3000 A 5000.00 SOLES
- 4. MAS DE 5000.00 SOLES

4.- NUMERO DE HIJOS:

- 1. UNO
- 2. DOS
- 3. TRES
- 4. MAS DE TRES

FACTORES GINECO-OBSTETRICOS

1.- TIPO DE PARTO:

- 1. NORMAL
- 2. CESAREA

2.- ANEMIA EN ETAPA GESTACIONAL:

- 1. CON ANEMIA
- 2. SIN ANEMIA

3.- PATOLOGIA MATERNA:

- 1. PRE ECLAMP SIA
- 2. ENFERMEDAD INFECCIOSA
- 3. OTRA
- 4. NINGUNA



ANEXO 04. DESCRIPCIÓN DE VARIABLES DEL SIEN EESS VALLECITO (I-3)-
DIRESA PUNO 2021

<i>Variable</i>	<i>Descripción</i>
EESS	Establecimiento de salud: Corresponde a la identificación de los 12 establecimientos de salud que cuenta el Ministerio de Salud para las atenciones en niños y gestantes en la MICRORED de Yauri.
Fecha	Se identifica la fecha en dd/mm/aaaa de la atención recibida por el niño menor de cinco años.
Hist_Clinica	Historia Clínica: Corresponde al número de historia clínica del niño que fue atendido en el establecimiento de salud.
Provincia	Corresponde a la identificación de la provincia, para nuestro caso ESPINAR, donde reside el niño, de acuerdo a lo señalado por el INEI.
Distrito	Corresponde a la identificación del distrito, de los 8 distritos, donde reside el niño, de acuerdo a lo señalado por el INEI.
Localidad	Corresponde a la identificación de la localidad donde reside el niño.
Hemoglobina	Corresponde al resultado de la hemoglobina (Hb) obtenido del análisis sanguíneo respectivo en el niño, expresado en g/dl.
Altitud_Loc	Corresponde a la altitud de la localidad o distrito donde reside el niño, en metros sobre el nivel del mar (msnm) identificados por INEI.
Hbc	Hemoglobina corregida: Corresponde a la cifra de hemoglobina



	aplicando los factores de corrección de acuerdo a la altitud, y se expresa en g/dl.
Sexo	Se identifica el sexo del niño, puede ser masculino o femenino.
Fec_Nacimiento	Fecha de nacimiento: Corresponde a la fecha del nacimiento del niño en dd/mm/aaaa.
Edad	Corresponde a la edad del niño en meses y fracción de mes.
Peso	Corresponde al peso del niño en kilogramos.
Talla	Corresponde a la talla del niño en centímetros.
Dx_P/T	Peso para la talla (P/T): Este índice compara el peso del niño con el peso esperado para su talla y permite establecer si ha ocurrido una pérdida reciente de peso corporal. Identifica desnutrición aguda y en el patrón OMS son los niños menores de cinco años que se encuentran entre los z-Score P/T entre menor a -2 hasta -5. En caso de sobrepeso los z-Score P/T son mayor a +2 hasta +3 y en caso de obesidad son z-Score mayor a +3 (hasta +5).
Dx_T/E	Talla para la edad (T/E): Este índice compara la longitud o talla del niño con la longitud o talla esperada para su edad y permite establecer si está ocurriendo un retraso en el crecimiento. Identifica el retardo de crecimiento (desnutrición crónica) y en el patrón OMS son los niños menores de cinco años que se encuentran entre los z-Score T/E menor a -2 a -3 y es severo si el z-Score es menor de -3.



Dx_P/E

Peso para la edad (P/E):

Este índice compara el peso del niño con el peso esperado para su edad y permite establecer si está ocurriendo desnutrición. Identifica la desnutrición global y en el patrón OMS son los niños menores de cinco años que se encuentran entre los z-Score P/E menor a -2. En caso de sobrepeso los z-Score son mayores a +2.

Dx_Anemia

Diagnóstico de Anemia: Es

la clasificación en función de la cifra de hemoglobina, de acuerdo a lo establecido por la OMS, que considera anemia si la hemoglobina es menor de 11.0 g/dl, para los niños a partir de los 6 meses de edad; se clasifica en Normal ($Hb \geq 11,0$ g/dl), anemia leve (Hb entre 10,0 – 10,9 g/dl), anemia moderada (Hb entre 7,0 – 9,9 g/dl) y anemia severa ($Hb < 7,0$ g/dl).

N_CCred

Número de Controles

CRED: Es el número de controles de Crecimiento y Desarrollo (CRED) realizados al niño menor de cinco años, que tiene a la fecha de la atención, incluyendo el control que se realiza en ese momento.

N_Sachets

Número de Sachets

consumidos: Es el número de sachets de multimicronutrientes consumidos por el niño menor de cinco años desde la consulta anterior a la fecha de la atención de ese momento.

N_Consejerías

Número de consejerías: Es

el número de consejerías nutricionales recibidas por las madres de los niños menores de cinco años desde la consulta anterior a la fecha de atención de ese momento.

N_Sesiones

Número de Sesiones: Es el

número de sesiones demostrativas recibidas por las madres de los niños menores de cinco años desde la consulta anterior a la fecha de atención de ese momento.

N_Visitas

Número de visitas

domiciliarias: Es el número de visitas domiciliarias recibidas por las familias con niños menores de cinco años



desde la consulta anterior a la fecha de atención de ese momento.

JUNTOS

Se identifica si la familia del niño, está afiliado al programa JUNTOS.

SIS

Se identifica si el niño, está afiliado al Seguro Integral de Salud - SIS.

PIN

Se identifica si el niño, está afiliado al Subprograma Infantil del Programa Integral de Nutrición - PIN.

Fuente: (SIEN, 2021)



ANEXO 05. SOLICITUD DIRIGIDA A DIRECTOR DE RED DE SALUD PUNO PARA FACILIDAD DE ACCESO A INFORMACIÓN



UNIVERSIDAD NACIONAL DEL ALTIPLANO

SOLICITA: Solicito datos estadísticos del HIS MINSA (CRED) y gestantes de la Unidad de Estadística, Red de Salud Puno.

SEÑOR DIRECTOR DE LA RED DE SALUD PUNO, M.C. JOSELIX YURI CANO ZIRENA



Yo, ROSA ANGELICA PAYEHUANCA MARCA,
Bachiller en Ciencias Estadística e Informática de la Universidad Nacional del Altiplano Puno, Identificado(a) con DNI. N° 77165031 y con domicilio(a) en Jr. Bellavista N° 129 de la ciudad de Puno.
Ante usted con el debido respeto me presento y expongo:

Tengo la complacencia de dirigirme a Ud. con el fin de solicitar el acceso a la información disponible del HIS MINSA de la Unidad de Estadística obrante en su base de datos y registros. A tal fin, le hago saber que el uso de la información recolectada será para la ejecución de mi proyecto de investigación, por lo tanto, con metas meramente científicas y cumplirá a cabalidad con el principio de confidencialidad manteniéndome en reserva la identidad de los sujetos de estudio. La solicitud comprende de registros HIS del control de crecimiento y desarrollo de niños menores de 3 años con y sin anemia y gestantes de la Red de Salud Puno durante el año 2019-2020-2021.

POR LO EXPUESTO:

Señor Director; ruego a usted acceder a mi petición para realizar el trabajo de investigación, por ser justa y necesaria.


.....
FIRMA

Puno, 06 de abril del 2022



ANEXO 06 CODIGO UTILIZADO EN COLAB - PYTHON

```
from google.colab import drive

drive.mount('/content/drive')

# Test it

!ls /content/drive/My Drive'

!ls /content/drive/My Drive'

##egresión Logística para predecir Anemia

from sklearn.metrics import accuracy_score

# importando bibliotecas

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn import linear_model

from sklearn.model_selection import train_test_split

from IPython.display import display, Math, Latex

#Instalar GGplot

!pip install 'plotnine[all]'

from sklearn import metrics

from pandas import Timestamp

from plotnine import *

import statsmodels.api as sm

import statsmodels.formula.api as smf

import math

import seaborn as sns

from sklearn import datasets

from sklearn.feature_selection import RFE

from sklearn.linear_model import LogisticRegression

from time import time

from sklearn.linear_model import RidgeCV

from sklearn.metrics import accuracy_score
```



```
data_children = pd.read_csv("/content/drive/My
Drive/Tesis_Anglk/data_niño_vallecito_2021.csv", sep=";", encoding='latin-1')

data_children.head()

data_children.columns.values

#Creando una columna llamada Anemia para clasificar de acuerdo al nivel de anemia si es
mayor a 11 "NO", si es menor "SI"

valores=[-math.inf,10.9,math.inf]

categorias=["Si", "No"]

data_children["Anemia"]=pd.cut(x=data_children['Hbc'],bins=valores,labels=categorias)

data_children.info()

#descriptivos de una variable

(unique, counts) = np.unique(data_children['Anemia'], return_counts=True)

frecuencias=pd.DataFrame(counts,unique)

frecuencias

#descriptivos de una variable

data_children['DIRESA'].describe()

#Para corroborar que niveles tiene o si es vacio u otro

(unique, counts) = np.unique(data_children['DIRESA'], return_counts=True)

frecuencias=pd.DataFrame(counts,unique)

frecuencias

#obteniendo solo valores de data_children las variables de interes a estudio

data=data_children.loc[:,['DNI','Sexo','Peso','Edad','Talla','N_CCred', 'N_Sachets',
'N_Consejerias', 'N_Sesiones',
'N_Visitas','Anemia']]

data.head()

#Abriendo base de datos 2 que comprende datos de la madre

data_mother=pd.read_excel("/content/drive/My Drive/Tesis_Anglk/Data_madre.xlsx")
```



```
data_mother.head()

#Creando una columna llamada Anemia_mother para clasificar de acuerdo al nivel de anemia
si es mayor a 11 "NO", si es menor "SI"

valores=[-math.inf, 10.9,math.inf]

categorias=["Si", "No"]

data_mother["Anemia_mother"]=pd.cut(x=data_mother['Hbc'],bins=valores,labels=categori
as)

#Cambiando nombre variable de DNI_Hijo a DNI para concatenar
data_mother1 = data_mother.rename(columns={'DNI_Hijo':'DNI'})

data_mother1.columns

# Juntando ambos dataframes en uno solo que estara en data
data=pd.merge(data, data_mother1, on='DNI', how='outer')

data.info()

data["Anemia"].value_counts()

##GRAFICOS

data.shape

data.info()

x = data["Anemia"].value_counts()

labels = ['Si', 'No']

#colors = ['tab:blue', 'tab:cyan']

explode = [0, 0.1]

fig, ax = plt.subplots()

ax.pie(x, labels = labels,

      startangle=90,
```



```
# colors = colors,  
autopct='% 1.1f%%',  
explode = explode,  
shadow = True)
```

```
ax.set_title('Nivel de Anemia')
```

```
plt.show()
```

Gráfico para cada variable cualitativa

```
=====
```

```
=====
```

```
# Ajustar número de subplots en función del número de columnas
```

```
fig, axes = plt.subplots(nrows=3, ncols=3, figsize=(12, 8))
```

```
axes = axes.flat
```

```
columnas_object = data.select_dtypes(include=['object']).columns
```

```
for i, column in enumerate(columnas_object):
```

```
    data[column].value_counts().plot.barh(ax = axes[i])
```

```
# pd.crosstab(data.column, data.Anemia).plot(kind="bar")
```

```
    axes[i].set_title(column, fontsize = 7, fontweight = "bold")
```

```
    axes[i].tick_params(labelsize = 6)
```

```
    axes[i].set_xlabel("")
```

```
fig.tight_layout()
```

```
plt.subplots_adjust(top=0.9)
```

```
fig.suptitle('Distribución variables cualitativas',
```

```
            fontsize = 10, fontweight = "bold");
```

```
%matplotlib inline
```

```
pd.crosstab(data.Sexo, data.Anemia).plot(kind="bar")
```

```
plt.title("Frecuencia de anemia en función de sexo de recién nacido")
```

```
plt.xlabel("Sexo")
```



```
plt.ylabel("Frecuencia de Anemia")

sns.boxplot(x=data['Anemia'],y=data['Edad'],data=data, hue="Anemia").set(title='Anemia
segun edad(Meses) de Niño')

sns.boxplot(x=data['Anemia'],y=data['Peso'],data=data, hue="Anemia").set(title='Anemia
segun Peso(Kg) de Niño')

sns.boxplot(x=data['Anemia'],y=data['Talla'],data=data, hue="Anemia").set(title='Anemia
segun Talla(Cm) de Niño')

%matplotlib inline

pd.crosstab(data.N_CCred, data.Anemia).plot(kind="bar")

plt.title("Frecuencia de anemia en función del Numero de controles CRED de recién nacido")

plt.xlabel("N_CCred")

plt.ylabel("Frecuencia de Anemia")

%matplotlib inline

pd.crosstab(data.N_Sachets, data.Anemia).plot(kind="bar")

plt.title("Frecuencia de anemia en función del consumo de Sachets")

plt.xlabel("Sachets")

plt.ylabel("Frecuencia de Anemia")

%matplotlib inline

pd.crosstab(data.N_Consejerias, data.Anemia).plot(kind="bar")

plt.title("Frecuencia de anemia en función de Consejería recibida por madre")

plt.xlabel("Consejería a madre")

plt.ylabel("Frecuencia de Anemia")

##Conversión de las variables categóricas a dummies
#convirtiendo a 0 y 1 la variable Y

data["Anemia"] = (data["Anemia"]=="Si").astype(int)

dat['Anemia'].values

#Convirtiendo a dummy variables categóricas

=====

=====

data.columns.values
```



```
categories = ['Sexo','N_Sachets','N_Consejerias', 'N_Sesiones',
'N_Visitas','Condicion_Civil', 'Nivel_Instrucción','SIS', 'Ingreso',
'Tipo_parto', 'Patologia_Materna','Anemia_mother']
#categories = ['Patologia_Materna','Anemia_mother']
for category in categories:
    cat_list = "cat"+ "_" +category
    cat_dummies = pd.get_dummies(data[category], prefix=cat_list)
    data_new = data.join(cat_dummies)
    data = data_new

data.columns.values

data_vars = data.columns.values.tolist()

categories = ['Sexo','N_Sachets','N_Consejerias', 'N_Sesiones',
'N_Visitas','Condicion_Civil', 'Nivel_Instrucción','SIS', 'Ingreso',
'Tipo_parto', 'Patologia_Materna','Anemia_mother']

to_keep = [v for v in data_vars if v not in categories]
to_keep = [v for v in to_keep if v not in ["default"]]

bank_data = data[to_keep]
bank_data.columns.values
#Eliminando variables que ya no se van a utilizar
bank_data=bank_data.drop(['DNI', 'NRO','cat_N_Visitas_0'], axis=1)
bank_data.columns.values
bank_data=bank_data.drop(['Hemoglobina', 'Hbc'], axis=1)
bank_data.columns.values

bank_data_vars = bank_data.columns.values.tolist()
```



```
Y = ['Anemia']

X = [v for v in bank_data_vars if v not in Y]

bank_data3=bank_data

bank_data3=bank_data3.insert(len(bank_data3.columns.values),      'newColMean6',
bank_data3['Anemia'])

##Selección de rasgos para el modelo

bank_data_vars = bank_data.columns.values.tolist()

Y = ['Anemia']

X = [v for v in bank_data_vars if v not in Y]

n=19

lr = LogisticRegression(solver='lbfgs',class_weight='balanced', max_iter=10000)

rfe = RFE(lr, n_features_to_select=n)

rfe = rfe.fit(bank_data[X], bank_data[Y].values.ravel())

z=zip(bank_data_vars,rfe.support_, rfe.ranking_)

list(z)

cols = ["Peso","cat_Sexo_M","N_CCred","cat_N_Sachets_0", 'cat_N_Sesiones_0',
'cat_Condicion_Civil_Casada',"cat_Nivel_Instrucción_Secundaria",
"cat_Nivel_Instrucción_Sup No Univ",'cat_Nivel_Instrucción_Sup Univ',
'cat_Ingreso_De 1000 a 3000','cat_Ingreso_De 3000 a 5000','cat_Ingreso_Mas de 5000',
"cat_Patologia_Materna_Otra",'cat_Patologia_Materna_Ninguna',
"cat_Anemia_mother_Si"]

len(cols)
```



```
X = bank_data[cols]

Y = bank_data["Anemia"]

## IMPLEMENTACION DE MODELO

=====

=====

bank_data_vars = bank_data.columns.values.tolist()

Y1 = ['Anemia']

X1 = [v for v in bank_data_vars if v not in Y1]

# aJUSTE ajuste de variables

n=15

lr = LogisticRegression(solver='lbfgs',class_weight='balanced', max_iter=1000)

rfe = RFE(lr, n_features_to_select=n)

rfe = rfe.fit(bank_data[X1], bank_data[Y1].values.ravel())

z=zip(bank_data_vars,rfe.support_, rfe.ranking_)

list(z)

cols = ["Peso","cat_Sexo_F",'N_CCred', "cat_N_Sachets_0",'cat_Condicion_Civil_Casada',
"cat_Nivel_Instrucción_Secundaria",
"cat_Nivel_Instrucción_Sup Univ", "cat_Nivel_Instrucción_Sup No Univ",
'cat_Ingreso_De 1000 a 3000','cat_Ingreso_De 3000 a 5000','cat_Ingreso_Mas de 5000',
"cat_Patologia_Materna_Otra", "cat_Anemia_mother_Si"]

len(cols)

X1 = bank_data[cols]

Y1 = bank_data["Anemia"]

X_train1, X_test1, Y_train1, Y_test1 = train_test_split(X1,Y1, test_size = 0.3,
random_state=0)

X_train1 = sm.add_constant(X_train1, prepend=True)
```



```
logit_model1 = sm.Logit(Y_train1, X_train1)

result1 = logit_model1.fit()

print(result1.summary())

# Accuracy de test del modelo

#

=====

==

umbral=0.75

X_test1 = sm.add_constant(X_test1, prepend=True)

predicciones1 = result1.predict(exog = X_test1)

prob_df1 = pd.DataFrame({"Probabilidad1":predicciones1})

prob_df1["prediction1"] = np.where(prob_df1['Probabilidad1']>=umbral, 1, 0)

prob_df1["actual1"] = list(Y_test1)

prob_df1.head()

# Matriz de confusión de las predicciones de test

#

=====

=====

confusion_matrix = pd.crosstab(

    prob_df1.actual1,

    prob_df1.prediction1,

    rownames=['Real'],

    colnames=['Predicción']

)

confusion_matrix
```



```
TN=confusion_matrix[0][0]

TP=confusion_matrix[1][1]

FN=confusion_matrix[0][1]

FP=confusion_matrix[1][0]

#SENSIBILIDAD

sens = TP/(TP+FN)

#ESPECIFICIDAD

espc_1 = 1-TN/(TN+FP)

#PRECISION

prec=TP/(FP+TP)

#ACCURACY o Exactitud

acc=(TN+TP)/(TN+TP+FN+FP)

##-----CURVA ROC-----

----

thresholds = [0.000001, 0.001,0.01,0.02,0.03,0.05, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40,
0.45,                                0.50,                                0.55,
0.60,0.65,0.70,0.75,0.80,0.85,0.9,0.95,0.955,0.9555,0.96,0.965,0.97,0.975,0.98,0.985,0.990,0.995,0.9
99]

sensitivities = [1]

especificities_1 = [1]

for t in thresholds:

    prob_df5["prediction"] = np.where(prob_df5['Probability']>=t, 1, 0)

    prob_df5["actual"] = list(Y_test4)

    prob_df5.head()

    confusion_matrix = pd.crosstab(prob_df5.prediction, prob_df5.actual)

    TN=confusion_matrix[0][0]

    TP=confusion_matrix[1][1]

    FP=confusion_matrix[0][1]
```



```
FN=confusion_matrix[1][0]

sens = TP/(TP+FN)
sensitivities.append(sens)

espc_1 = 1-TN/(TN+FP)
especificities_1.append(espc_1)

sensitivities.append(0)
especificities_1.append(0)

auc = metrics.auc(especificities_1, sensitivities)
auc

%matplotlib inline
plt.plot(especificities_1, sensitivities, marker="o", linestyle="--", color="r")
x=[i*0.01 for i in range(100)]
y=[i*0.01 for i in range(100)]
plt.plot(x,y)
plt.xlabel("1-Especificidad")
plt.ylabel("Sensibilidad")
plt.title("Curva ROC Y AUC= "+ str(auc*100)+"%")
```



AUTORIZACIÓN PARA EL DEPÓSITO DE TESIS O TRABAJO DE INVESTIGACIÓN EN EL REPOSITORIO INSTITUCIONAL

Por el presente documento, Yo ROSA ANGELICA PAYEHUANCA MARCA,
identificado con DNI 77165031 en mi condición de egresado de:
FACULTAD DE INGENIERIA ESTADISTICA E INFORMATICA

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado

INGENIERIA ESTADISTICA E INFORMATICA,

informo que he elaborado el/la Tesis o Trabajo de Investigación denominada:

“MODELO LOGIT PARA IDENTIFICAR LOS FACTORES ASOCIADOS A LA PREVALENCIA DE
ANEMIA EN MENORES DE 3 A 59 MESES DEL CENTRO DE SALUD I-3 VALLECITO PUNO-2021”

para la obtención de Grado, Título Profesional o Segunda Especialidad.

Por medio del presente documento, afirmo y garantizo ser el legítimo, único y exclusivo titular de todos los derechos de propiedad intelectual sobre los documentos arriba mencionados, las obras, los contenidos, los productos y/o las creaciones en general (en adelante, los “Contenidos”) que serán incluidos en el repositorio institucional de la Universidad Nacional del Altiplano de Puno.

También, doy seguridad de que los contenidos entregados se encuentran libres de toda contraseña, restricción o medida tecnológica de protección, con la finalidad de permitir que se puedan leer, descargar, reproducir, distribuir, imprimir, buscar y enlazar los textos completos, sin limitación alguna.

Autorizo a la Universidad Nacional del Altiplano de Puno a publicar los Contenidos en el Repositorio Institucional y, en consecuencia, en el Repositorio Nacional Digital de Ciencia, Tecnología e Innovación de Acceso Abierto, sobre la base de lo establecido en la Ley N° 30035, sus normas reglamentarias, modificatorias, sustitutorias y conexas, y de acuerdo con las políticas de acceso abierto que la Universidad aplique en relación con sus Repositorios Institucionales. Autorizo expresamente toda consulta y uso de los Contenidos, por parte de cualquier persona, por el tiempo de duración de los derechos patrimoniales de autor y derechos conexos, a título gratuito y a nivel mundial.

En consecuencia, la Universidad tendrá la posibilidad de divulgar y difundir los Contenidos, de manera total o parcial, sin limitación alguna y sin derecho a pago de contraprestación, remuneración ni regalía alguna a favor mío; en los medios, canales y plataformas que la Universidad y/o el Estado de la República del Perú determinen, a nivel mundial, sin restricción geográfica alguna y de manera indefinida, pudiendo crear y/o extraer los metadatos sobre los Contenidos, e incluir los Contenidos en los índices y buscadores que estimen necesarios para promover su difusión.

Autorizo que los Contenidos sean puestos a disposición del público a través de la siguiente licencia:

Creative Commons Reconocimiento-NoComercial-CompartirIgual 4.0 Internacional. Para ver una copia de esta licencia, visita:

En señal de conformidad, suscribo el presente documento.

Puno, 13 julio del 2023


FIRMA



Huella



DECLARACIÓN JURADA DE AUTENTICIDAD DE TESIS

Por el presente documento, Yo ROSA ANGELICA PAYEHUANCA MARCA,
identificado con DNI 77165031 en mi condición de egresado de:

Escuela Profesional, Programa de Segunda Especialidad, Programa de Maestría o Doctorado
INGENIERIA ESTADISTICA E INFORMATICA,

informo que he elaborado el/la Tesis o Trabajo de Investigación denominada:
“MODELO LOGIT PARA IDENTIFICAR FACTORES ASOCIADOS A LA PREVALENCIA DE
ANEMIA EN MENORES DE 3 A 59 MESES DEL CENTRO DE SALUD I-3 VALLECITO
PUNO-2021”

Es un tema original.

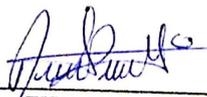
Declaro que el presente trabajo de tesis es elaborado por mi persona y **no existe plagio/copia** de ninguna naturaleza, en especial de otro documento de investigación (tesis, revista, texto, congreso, o similar) presentado por persona natural o jurídica alguna ante instituciones académicas, profesionales, de investigación o similares, en el país o en el extranjero.

Dejo constancia que las citas de otros autores han sido debidamente identificadas en el trabajo de investigación, por lo que no asumiré como tuyas las opiniones vertidas por terceros, ya sea de fuentes encontradas en medios escritos, digitales o Internet.

Asimismo, ratifico que soy plenamente consciente de todo el contenido de la tesis y asumo la responsabilidad de cualquier error u omisión en el documento, así como de las connotaciones éticas y legales involucradas.

En caso de incumplimiento de esta declaración, me someto a las disposiciones legales vigentes y a las sanciones correspondientes de igual forma me someto a las sanciones establecidas en las Directivas y otras normas internas, así como las que me alcancen del Código Civil y Normas Legales conexas por el incumplimiento del presente compromiso

Puno, 13 de julio del 2023



FIRMA



Huella