

# Modelo predictivo para la detección de anemia en mujeres embarazadas.

*por* Alexander Daniel Lizana Yauri

---

**Fecha de entrega:** 26-jun-2023 11:49a.m. (UTC-0500)

**Identificador de la entrega:** 2123025737

**Nombre del archivo:** DRAFT\_PAPER\_G5.pdf (1.19M)

**Total de palabras:** 4279

**Total de caracteres:** 22500

**UNIVERSIDAD NACIONAL MAYOR DE SAN MARCOS**

**FACULTAD DE INGENIERÍA DE SISTEMAS E INFORMÁTICA**

**E.A.P DE INGENIERÍA DE SOFTWARE**



**Modelo predictivo para la detección de anemia en mujeres embarazadas.**

**Curso:** Analítica de Datos

**Grupo 5**

**Profesor:** Pedro Martín Lezama Gonzales

**Integrantes:**

- Lizana Yauri Alexander Daniel
- Muñoz Silva Brayan
- Ginocchio Salcedo Gino

**Lima - 2023**

## Modelo predictivo del nivel <sup>25</sup> de anemia de gestantes en la provincia de Cutervo

### Introducción:

<sup>61</sup> En el Perú la anemia en gestantes es un problema latente, <sup>25</sup> en el primer semestre del 2022, se evaluaron a 122 202 gestantes, de las cuales 24 809 gestantes se registraron como pacientes anémicas, por lo que un 20,3% de mujeres gestantes son anémicas, esto representa un 1% más que lo registrado en el primer semestre 2021. Por ende la anemia es una enfermedad latente dentro del país y debe ser tratada adecuadamente para reducir el peligro en el que se encuentra la gestante como el feto.

<sup>41</sup> El siguiente trabajo tiene como objetivo predecir el nivel de anemia de las gestantes de la provincia de Cutervo, a través <sup>25</sup> de algoritmos de aprendizaje automático. A través de algoritmos de aprendizaje automático, se podrá predecir el nivel de anemia de gestantes en las provincia de Cutervo.

### Justificación de la investigación

#### 1. Justificación teórica:

<sup>46</sup> Como se ha explicado anteriormente la anemia en gestantes es un problema de salud <sup>38</sup> importante que requiere atención debido a sus efectos negativos tanto para la madre como para el feto. Por lo que <sup>44</sup> predecir la anemia es una tarea importante en la medicina, ya que puede ayudar a identificar a los pacientes que están en riesgo de desarrollar anemia y, por lo tanto, permitir una intervención temprana y efectiva. Además, el uso de modelos predictivos también puede ayudar a reducir los costos asociados con la detección y el tratamiento de la anemia, ya que permite una identificación temprana y un tratamiento efectivo antes de que la anemia se desarrolle completamente.

#### 2. Justificación práctica:

La predicción de la anemia en gestantes es importante por varias razones prácticas:

- Prevención: Al predecir la anemia en una gestante tempranamente, se pueden tomar medidas preventivas para evitar que la anemia se desarrolle o empeore. Esto incluye la recomendación de una dieta equilibrada, suplementos de hierro y otros nutrientes clave.

- Tratamiento temprano: La predicción temprana de la anemia permite que se inicie el tratamiento de manera oportuna, lo que puede ayudar a prevenir complicaciones graves y mejorar la salud de la madre y el feto.
- Mejora de la eficacia del tratamiento: Al conocer la gravedad de la anemia, se pueden personalizar los tratamientos para asegurar que sean lo más efectivos posible.
- Reducción de costos: La predicción temprana de la anemia puede ayudar a reducir los costos asociados con la atención médica y el tratamiento de las complicaciones relacionadas con la anemia.

### Marco Teórico:

La anemia en las gestantes es una condición en la que la cantidad de hemoglobina en la sangre es menor de lo normal. La hemoglobina es una proteína presente en los glóbulos rojos que transporta el oxígeno desde los pulmones hasta los tejidos del cuerpo. Durante el embarazo, el cuerpo necesita más oxígeno para el crecimiento y desarrollo del feto, lo que aumenta la demanda de hemoglobina. Si no hay suficiente hemoglobina, la madre y el feto pueden sufrir consecuencias graves.

Las causas comunes de anemia en las gestantes incluyen:

- **Deficiencia de hierro:** El hierro es esencial para la producción de hemoglobina, y la demanda de hierro aumenta durante el embarazo debido al crecimiento del feto y la placenta.
- **Deficiencia de ácido fólico:** El ácido fólico es esencial para la producción de glóbulos rojos y también es necesario para el desarrollo del sistema nervioso del feto.
- **Pérdida de sangre:** Puede ocurrir durante el embarazo debido a complicaciones como el sangrado de las encías o una placenta previa.

Los síntomas de anemia pueden incluir fatiga, debilidad, dificultad para respirar, palidez de la piel y del iris, y dolores de cabeza. Si se sospecha de anemia en una mujer embarazada, se realizará un análisis de sangre para medir los niveles de hemoglobina. Si se diagnostica anemia, el tratamiento dependerá de la causa subyacente y puede incluir suplementos de hierro y ácido fólico, así como cambios en la dieta. Es importante que las mujeres

embarazadas reciban atención prenatal adecuada para detectar y tratar la anemia tempranamente.

Estado del Arte

2 Para la investigación de trabajos relacionados con el estado del arte, se utilizó la metodología de revisión sistemática. Bajo este enfoque, se utilizó las siguientes preguntas de investigación:

- P1: ¿Qué algoritmos de aprendizaje automático se han aplicado para la predicción de anemia?
- P2: ¿Qué estudios de aprendizaje automático se han realizado para la predicción de anemia?
- P3: 56 ¿En qué tipos de pacientes se han aplicado algoritmos de aprendizaje automático para la predicción de anemia ?
- P4: ¿Qué métodos utilizaron para la recolección de datos y el preprocesamiento de los datos (construcción de dataset)?

Literatura Revisada

Nº	Título	Referencia	Pregunta de Investigación	Año
1	34 Predicting the level of anemia among Ethiopian pregnant women using homogeneous ensemble machine learning algorithm	(Dejene et al., 2022)	P1, P2, P3, P4	2022
2	32 Building a prediction model for iron deficiency anemia among infants in Shanghai, China	(Zhang & Tang, 2020)	P1, P2, P3	2020
3	64 Prediction of anemia and estimation of hemoglobin concentration using a smartphone camera	(Suner, 2021)	P1, P2, P3	2021
4	65 A new machine learning approach for predicting the response to anemia treatment in a large cohort of End Stage Renal Disease patients undergoing dialysis	(Barbieri et al., 2015)	P2	2015

5	<sup>35</sup> Diagnosis of iron deficiency anemia using density-based fractionation of red blood cells	(Hennek, 2016)	P1, P2	2016
6	Classifying anemia types using artificial learning methods	(Yıldız <sup>39</sup> et al., 2021)	P1, P2	2021
7	<sup>33</sup> Machine Learning Algorithms To Predict The Childhood Anemia In Bangladesh	(Khan et al., 2021)	P1, <sup>20</sup> P2	2021
8	<sup>33</sup> Predicting child anaemia in the North-Eastern states of India: a machine learning approach	(Meitei <sup>54</sup> et al., 2022)	P1, P3, P4	2022
9	<sup>16</sup> Smartphone-based photo analysis for the evaluation of anemia, jaundice and COVID-19	(Mazzu et al., 2021)	P1, P4	2021
10	<sup>1</sup> ThalPred: a web-based prediction tool for discriminating thalassemia trait and iron deficiency anemia	(Lal <sup>43</sup> Sri et al., 2019)	P1, P4	2019
11	Detection of anaemia from retinal fundus images via deep learning	(Mitani et al., 2020)	P1, P3, P4	2020
12	<sup>10</sup> Machine Learning Algorithms for Anemia Disease Prediction. Recent Trends in Communication, Computing, and Electronics	(Jaiswal et al., 2018)	P1, P3, P4	2018
13	<sup>34</sup> MACHINE LEARNING ALGORITHMS FOR ANEMIA DISEASE PREDICTION - A REVIEW	(Verma & Chopra, 2022, )	P1, P3	2022
14	<sup>5</sup> A new machine learning approach for predicting the response to anemia treatment in a large cohort of End Stage Renal Disease patients undergoing dialysis	(Barbieri, 2015)	P1, P3	2015
15	<sup>29</sup> Modelo de minería de datos para la predicción de anemia en gestantes de la provincia de Ilo	(Sadan, 2019)	P1, P3, P4	2019

### Resultados de la revisión

De los artículos revisados podemos extraer los algoritmos más utilizados para la predicción de anemia.

Modelos	Referencias
Naive Bayes	(Dejene et al., 2022), (Yıldız et al., 2021) (Hennek, 2016) (Verma & Chopra, 2022) (Sadan, 2019)
KNN	(Khan et al., 2021) (Laengsri et al., 2019)
Regresión Logística	(Khan et al., 2021)
Random Forest	(Yıldız et al., 2021) (Khan et al., 2021) (Laengsri et al., 2019) (Sadan, 2019)
Cat Boost	(Dejene et al., 2022)
Redes Neuronales	(Zhang & Tang, 2020) (Yıldız et al., 2021) (Khan et al., 2021) (Laengsri et al., 2019)

### Aporte Teórico

Para el aporte teórico de este trabajo se busca plantear una metodología propuesta para el análisis y procesamiento del modelo predictivo del nivel de anemia.

### Metodología a propuesta:

La metodología que utilizará para desarrollar este trabajo de investigación constará de los siguientes pasos:

- Origen del Dataset.
- Limpieza y preprocesamiento de datos
- Selección de las características
- Construcción final del conjunto de datos final
- Modelado predictivo
- Resultados y discusiones

### Origen del Dataset

El origen de un conjunto de datos es el lugar o la fuente desde donde se obtuvieron los datos. Puede ser una encuesta, una base de datos, un archivo de registro, una colección de imágenes, entre otros. Es importante conocer el origen de un conjunto de

datos para entender su calidad, su relevancia y su validez para un determinado propósito.

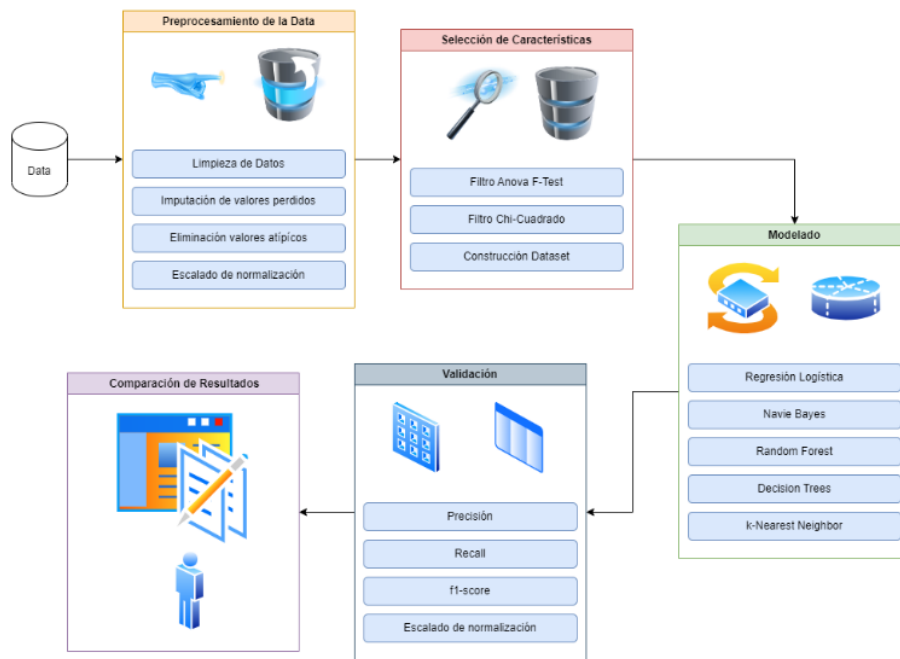
Para este trabajo el dataset fue recogido del Repositorio de datos abiertos, el cual recoge el estado nutricional en gestantes en la provincia de Cutervo en el año de 2017, la finalidad de este dataset es brindar información del estado nutricional de la gestante que acceden a los establecimientos de salud, para la toma de decisiones y la planificación de intervenciones que contribuyan a mejorar la calidad de vida de la población.

### **Limpieza y preprocesamiento de datos**

La imputación de valores perdidos se refiere al proceso de reemplazar los valores faltantes en un conjunto de datos con valores aproximados. Esto se hace para evitar que los datos faltantes afecten negativamente el análisis o la modelación de los datos. Existen varios métodos comunes de imputación de valores perdidos, como la imputación por media, mediana o valor más frecuente, así como métodos más avanzados como la imputación por regresión o la imputación basada en modelos estadísticos.

Por otro lado, la eliminación de valores atípicos se refiere al proceso de detectar y eliminar los valores que son muy diferentes de los demás en un conjunto de datos. Estos valores se conocen como valores atípicos o outliers y pueden afectar negativamente el análisis o la modelación de los datos. Existen varios métodos comunes para detectar y eliminar valores atípicos, como el uso de gráficos de caja y bigote, así como el uso de límites estadísticos, como los valores que están fuera de los límites de la desviación estándar.





## Selección de características

La selección de características es un proceso en el aprendizaje automático que consiste en elegir un subconjunto de características o variables de un conjunto de datos para utilizar en un modelo.

La selección de características tiene varios objetivos:

- Mejorar la precisión del modelo al reducir el ruido o las características irrelevantes
- Reducir el sobreajuste (overfitting) al eliminar características que no aportan información relevante
- Reducir la complejidad del modelo al utilizar solo las características más relevantes.

Existen varios métodos para la selección de características, algunos de los cuales son:

- **Filtro F-Clasif:** El filtro F-Clasif es un algoritmo utilizado en el campo del procesamiento de datos y la minería de datos para realizar clasificación y filtrado de información. Este filtro se basa en un enfoque estadístico que

utiliza el análisis de la varianza (ANOVA) para identificar las características más relevantes o discriminantes de un conjunto de datos. A través de la comparación de las medias y las varianzas de diferentes grupos de datos, el filtro F-Clasif determina qué características tienen un impacto significativo en la clasificación de los datos. Al identificar estas características clave, el filtro F-Clasif permite seleccionar y filtrar las variables más importantes para mejorar la precisión y eficiencia de los modelos de clasificación y análisis de datos.

- **Filtro K-beast:** El Filtro K-beast es un algoritmo utilizado en el campo de la predicción y el análisis de datos para mejorar la precisión en el modelo predictivo. Este filtro se basa en técnicas avanzadas de aprendizaje automático y análisis de datos para identificar y analizar patrones específicos en los datos recopilados. Al aplicar el Filtro K-beast, es posible realizar un filtrado adaptativo de las variables relevantes relacionadas con la anemia, lo que ayuda a identificar las características clave que influyen en la detección temprana de esta condición. Este enfoque inteligente permite construir un modelo predictivo más preciso y confiable para identificar a las mujeres embarazadas en riesgo de desarrollar anemia, lo que facilita la toma de decisiones y la implementación de intervenciones preventivas oportunas.

### Construcción del conjunto de datos final

La construcción de un conjunto de datos es el proceso mediante el cual se recolectan, limpian, integran y transforman diferentes fuentes de datos para crear un conjunto único y coherente que pueda ser utilizado para una variedad de tareas, como el aprendizaje automático o la toma de decisiones. Este proceso puede ser complejo y requerir habilidades de programación, estadísticas y conocimiento del negocio para asegurar que el conjunto de datos final sea preciso, relevante y útil.

### Modelado predictivo

En esta sección se busca emplear los algoritmos de Regresión logística, Navie Bayes, Random Forest, KNN y Red Neuronal MLP

- Regresión logística
- Navie Bayes

- Random Forest
- KNN
- Red Neuronal MLP

### Resultados y discusiones

En este estudio se busca evaluar el desempeño de cinco algoritmos de aprendizaje automático en el modelo predictivo <sup>23</sup> para la detección de anemia en mujeres embarazadas. Estos algoritmos son: Regresión logística, Navie Bayes, Random Forest, KNN y Red Neuronal MLP.

Categoría	Tipo de dato
DIRESA	Categórica
RED	Categórica
MICRORED	Categórica
EESS	Categórica
RENIPRESS	Numérica
Fecha	Categórica
Edad	Numérica
<sup>3</sup> Edad Gestacional	Numérica
Peso	Numérica
Talla	Numérica
PPG	Numérica
Tipo Embarazo	Categórica
UBIGEO	Numérica
Provincia	Categórica
Distrito	Categórica
Altitud_Dist	Numérica
Altitud_Loc	Numérica
Hematocrito	Numérica

Hemoglobina	Numérica
FechaHemoglobina	Categórica
Hbc	Numérica
Dx_Anemia	Categórica
Dx_CLAP	Categórica
Dxx_IOM	Categórica

**Tabla 1:** Variables del dataset Gestantes Cutervo Estado Nutricional 2017.csv

La “Tabla 2” muestra el resumen general del dataset antes de la etapa de preprocesamiento, donde se aprecian 12 variables numéricas y 12 variables categóricas, por lo que será necesaria convertir a variables numéricas a las variables categóricas, de tal manera que el conjunto de datos sirva como input para el modelo empleado.

Nombre	Valor
N° Filas	8331
N° Columnas	24
N° Variables numéricas	12
N° Variables categóricas	12

**Tabla 2:** Resumen general del dataset Gestantes Cutervo Estado Nutricional 2017.csv

### Limpieza y preprocesamiento de datos

Analizando el dataset podemos identificar y eliminar datos que se pueden considerar un ruido debido a que no aportan información relevante a la data. Estos variables se muestran en la siguiente tabla:

Categoría	Tipo de dato
DIRESA	Categórica
RED	Categórica
MICRORED	Categórica
EESS	Categórica
RENIPRESS	Numérica

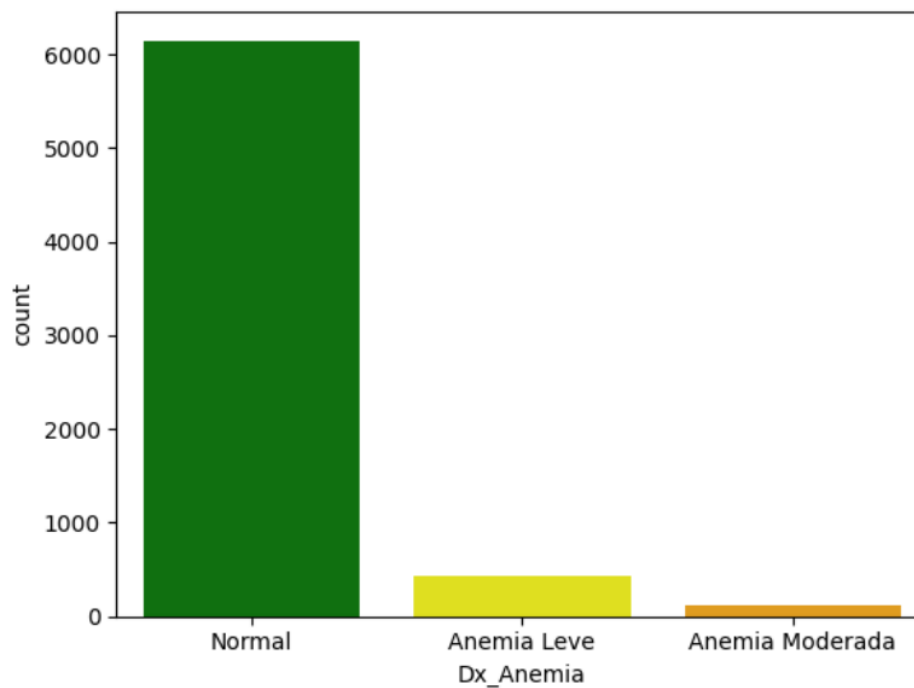
Fecha	Categórica
UBIGEO	Numérica
Hematocrito	Numérica
Fecha Hemoglobina	Categórica
Dx_CLAP	Categórica
Dx_IOM	Categórica

**Tabla 3:** Variables que serán removidas del dataset Gestantes Cutervo Estado Nutricional 2017.csv

Analizando el dataset, se muestra que existen valores vacíos en algunas columnas, por lo que se removerán las columnas con datos vacíos. Tras analizar el dataset notamos que la columna Dx\_Anemia cuenta con valores “No Evaluados” el cual es considerado como datos nulos para la investigación por lo que serán removidos del dataset. A continuación mostramos la cantidad de datos existentes dentro de la columna Dx\_Anemia:

Variables	Cantidad
Normal	6151
Anemia Leve	436
Anemia Moderada	120

**Tabla 4:** Se muestran la cantidad de valores de la categoría Dx\_Anemia del dataset utilizado.



**Img 1:** Se muestran la cantidad de valores de la categoría Dx\_Anemia del dataset utilizado.

### **Selección de características**

Para esta fase se empleará la clase SelectKBest de la librería de aprendizaje automático Scikit-learn. Esta implementación permitió la selección de las mejores características basados en pruebas estadísticas univariadas. Entre ellas tenemos:

#### **Filtro F-Clasif**

Para la selección de características se estableció como parámetro Score = “f\_classif” y como parámetro Variables = “Todas las variables no nulas”. Las entradas para el Algoritmo F-Clasif fue el conjunto de datos numéricos “Dx\_Anemia”, y el resto de la data como variable objetivo. Esta configuración permite obtener las puntuaciones que se observan en la siguiente figura:

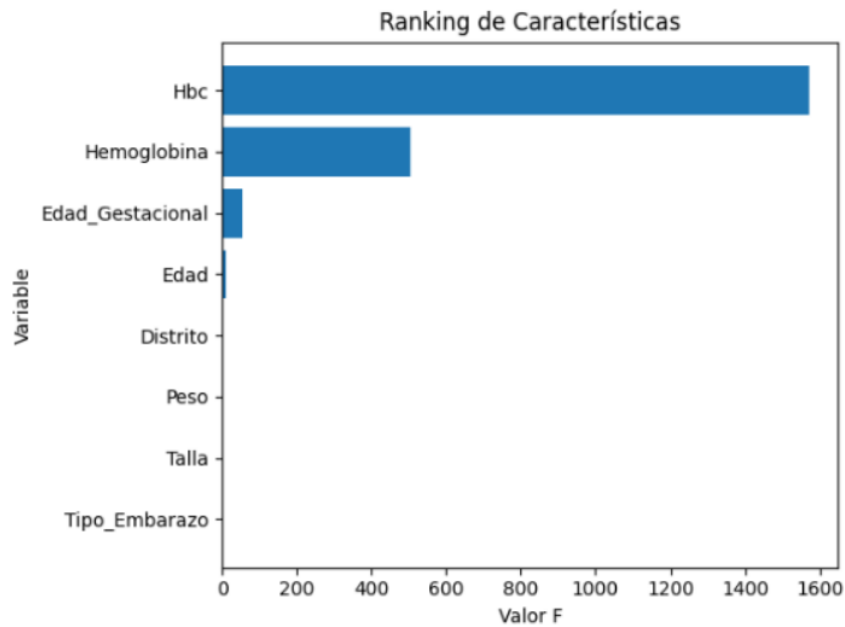
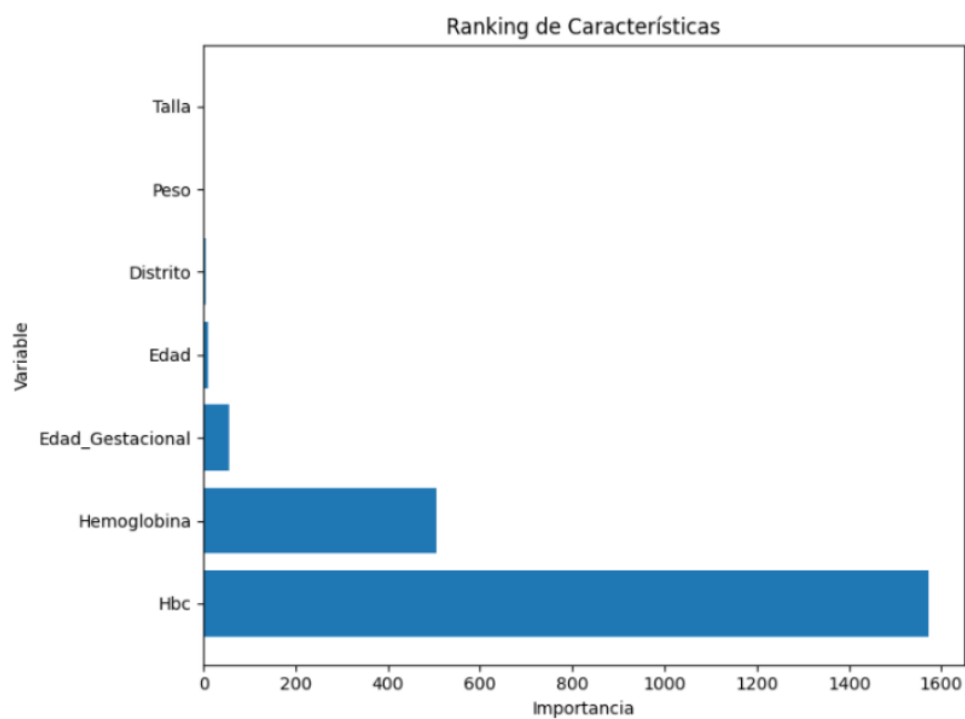


Fig 1. Resultados del filtro Anova F Test

#### Filtro K-beast:

Para usar este filtro se realiza un análisis de características para identificar las más relevantes en un conjunto de datos representado por un DataFrame. Primero, se obtienen los índices de las características seleccionadas y se accede a sus nombres. Luego, se obtienen los valores de importancia de estas características. Luego, se crea un Data Frame que muestra el ranking de características, ordenado de forma descendente según su importancia. Finalmente, se genera un gráfico de barras que visualiza este ranking.



**Fig. 2** Resultados del filtro chi-cuadrado con respecto Dx\_Anemia y los Campos seleccionados

Debido a los resultados se decidió elegir 8 características del conjunto de datos categóricos.

Los cuales se detallan en la siguiente tabla:

Atributo	Descripción
Edad	Edad de la Persona
Edad Gestacional	Edad Gestacional de la persona
Peso	Peso de la Persona
Talla	Talla de la persona
Distrito	Distrito de las personas
Hemoglobina	valor de hemoglobina
Dx_Anemia	Valor de Detector de anemia
IMC	Índice de masa corporal
Valor_IMC	IMC en texto e interpretación

**Tabla 5:** Atributos seleccionados



### Construcción del conjunto de datos final

En la Tabla 6 se pueden apreciar la estructura de atributos usada en el conjunto de datos final.

Edad	Edad_Gestacional	Peso	Talla	Distrito	Hemoglobina	Hbc	Dx_Anemia	IMC	Valor_IMC
19	27	61.0	156.0	7	14.3	13.5	0	25.06574621959237	Sobrepeso
19	27	67.0	160.0	7	11.6	10.8	1	26.171874999999996	Sobrepeso
25	14	61.5	151.0	7	14.0	13.2	0	26.97250120608745	Sobrepeso
35	18	62.0	158.0	7	12.0	11.2	0	24.83576349943919	Normal
14	28	59.0	152.0	7	13.3	12.5	0	25.536703601108034	Sobrepeso

Tabla 6: Estructura del conjunto de datos

### Modelado predictivo

#### Regresión Logística:

Para el primer modelo dividimos nuestros <sup>17</sup> datos en un conjunto de entrenamiento y un conjunto de pruebas. Luego creamos los modelos de entrenamiento con la librería sklearn y para luego generar los resultados del modelo.

El reporte de la clasificación:

	<sup>15</sup> precisión	recall	f1-score	support
Anemia leve	0.83	0.79	0.81	127
Anemia moderada	0.84	0.79	0.81	33
Normal	0.99	0.99	0.99	1853
<sup>6</sup> accuracy			0.98	2013
macro avg	0.89	0.86	0.87	2013
weighted avg	0.98	0.98	0.98	2013

### Naive Bayes:

Para este modelo dividimos nuestros datos en un conjunto de entrenamiento y un conjunto de pruebas. Luego creamos los modelos de entrenamiento con la librería sklearn. Después del proceso se puede observar la matriz de confusión obtenida.

El reporte de la clasificación:

	precisión	recall	f1-score	support
Anemia leve	0.85	0.88	0.86	127
Anemia moderada	0.83	0.76	0.79	33
Normal	0.99	0.99	0.99	1853
accuracy			0.98	2013
macro avg	0.89	0.88	0.88	2013
weighted avg	0.98	0.98	0.98	2013

### KNN:

Primero se transformó la data para posteriormente crear los modelos de entrenamiento con la librería sklearn, para así obtener la siguiente matriz de confusión para un K=5.

El reporte de clasificación:

	precisión	recall	f1-score	support
Anemia leve	0.38	0.06	0.11	127
Anemia moderada	0.25	0.03	0.05	33
Normal	0.93	1.00	0.96	1853
accuracy			0.95	2013
macro avg	0.52	0.36	0.37	2013
weighted avg	0.88	0.92	0.89	2013

En los resultados obtenidos de los autores muy pocos usaron K-Vecinos por los resultados bajos. En nuestra matriz obtuvo una precisión de 95% predictivo. A comparación de otras es menor.

### Red Neuronal MLP:

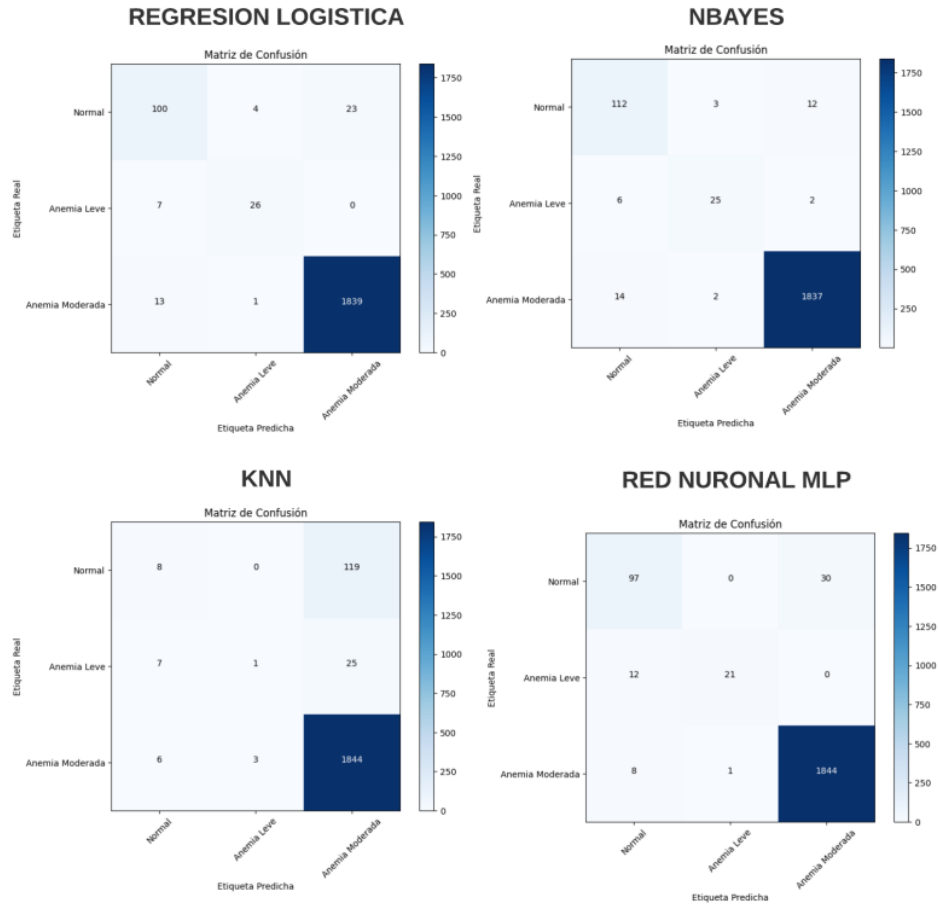
Se dividió en un conjunto de entrenamiento y de prueba, para posteriormente entrenar el modelo, después de este proceso se logró obtener el siguiente reporte:

	precisión	recall	f1-score	support
Anemia leve	0.38	0.06	0.11	127
Anemia moderada	0.25	0.03	0.05	33
Normal	0.93	1.00	0.96	1853
accuracy			0.92	2013
macro avg	0.52	0.36	0.37	2013
weighted avg	0.88	0.92	0.89	2013

En un estudio se realizó la predicción usando el algoritmo de Red Neuronal MLP, que logró la mejor precisión de clasificación con un 68,53% y una sensibilidad del 70,73%. El modelo propuesto obtuvo una precisión de 100% que se estimaría a un 99.99% de precisión.

### Resultados de los modelos

Las matrices obtenidas por los distintos algoritmos de predicción se detallan en la figura, para el modelo de regresión logística obtuvo como resultado de la predicción 0.83 para anemia leve, 0.84 para detectar gestantes anémicas moderadas y se obtuvo un 0.99 para gestante normales, para el modelo naive bayes se obtuvo como resultado de la predicción 0.85 para anemia leve, 0.83 para detectar gestantes anémicas moderadas y se obtuvo un 0.99 para gestante normales, el modelo KNN se obtuvo como resultado de la predicción 0.38 para anemia leve, 0.25 para detectar gestantes anémicas moderadas y se obtuvo un 0.93 para gestante normales, para el modelo árboles de decisión, para el modelo final Red neuronal MLP se obtuvo como resultado de la predicción 0.81 para anemia leve, 1.00 para detectar gestantes anémicas moderadas y se obtuvo un 0.97 para gestante normales.



**Fig. 3** Resultados de la matriz de confusión de los modelos planteados

En la tabla 8 se analizaron 4 algoritmos con respecto a la literatura revisada estos algoritmos fueron: Regresión Logística, Naive Bayes, KNN y Random Forest.

Donde se observa que en el algoritmo con la mayor precisión es el de los Árboles de Decisión del trabajo de Meena, K., Tayal con un 97.65% , seguido además de los algoritmos de Random Forest y Naive Bayes del trabajo de Jaiswal, M. con unas precisiones de 95.3241 y 96.0909 respectivamente, mientras que en el algoritmo trabajado por nosotros este resultado es de un 98% para Regresión Logística y Naive Bayes, demostrando así que nuestro algoritmo es más preciso para la predicción de anemia.

Mientras que los algoritmos menos precisos fueron los de Regresión Logística y KNN del trabajo de T. Karagül Yıldız donde apenas se alcanzó una precisión de 76.1% y 76.4 respectivamente, mientras que en el algoritmo trabajado por nosotros

este resultado es de un 98% y 95% respectivamente, donde también nuestro algoritmo menos preciso fue el de Redes Neuronales MLP con una precisión de 92%, demostrando así que aun con los algoritmo menos preciso de la literatura, el nuestro logra tener una precisión más que aceptable en la predicción de la anemia.

## References

- 4 Barbieri, C. (2015). *A new machine learning approach for predicting the response to anemia treatment in a large cohort of End Stage Renal Disease patients undergoing dialysis*. *Comput Biol Med*. 10.1016/j.combiomed.2015.03.019
- 1 Barbieri, C., Mari, F., Stopper, A., Gatti, E., Escandell-Montero, P., Martínez-Martínez, J. M., & Martín-Guerrero, J. D. (2015). *A new machine learning approach for predicting the response to anemia treatment in a large cohort of End Stage Renal Disease patients undergoing dialysis*. *Comput Biol Med*. 10.1016/j.combiomed. 2015.03.019
- 26 Cari, G. (2021). *Factores de riesgo asociados a la anemia en gestantes que acuden al CAP II*. Universidad Nacional Jorge Basadre Grohmann.
- 4 Dejene, B. E., Abuhay, T. M., & Bogale, D. S. (2022). *Predicting the level of anemia among Ethiopian pregnant women using homogeneous ensemble machine learning algorithm* (nro 1 ed., Vol. vol 22). *BMC Med Inform Decis Mak*. 10.1186/s12911-022-01992-6
- 55 Gerget, O., Devyatykh, D., & Devyatykh, M. (2017). *Data-driven approach for modeling of control action impact on anemia dynamics based on energy-informational health state criteria*. *Commun. Comput. Inf. Sci*.
- 28 Gonzales, G. F., & Olavegoya, P. (2019). *Fisiopatología de la anemia durante el embarazo: ¿anemia o hemodilución?* Instituto de Investigaciones de la Altura y Departamento de Ciencias Biológicas y Fisiológicas, Facultad de Ciencias y Filosofía, Universidad Peruana Cayetano Heredia. <https://doi.org/10.31403/rpgo.v65i2210>
- 14 Hennek, J. W. (2016). *Diagnosis of iron deficiency anemia using density-based fractionation of red blood cells* (nro. 20 ed., Vol. vol. 16). *Lab Chip*. 10.1039/C6LC00875E

- 10 Jaiswal, M., Srivastava, A., & Siddiqui, T. (2018). *Machine Learning Algorithms for Anemia Disease Prediction. Recent Trends in Communication, Computing, and Electronics*.  
10.1007/978-981-13-2685-1\_44
- 1 Khan, J. R., Chowdhury, S., Islam, H., & Raheem, E. (2021). *Machine Learning Algorithms To Predict The Childhood Anemia In Bangladesh* (nro. 1 ed., Vol. vol. 17). Journal of Data Science. doi: 10.6339/JDS.201901\_17(1).0009
- 7 Laengsri, V., Shoombuatong, S., Adirojananon, W., Nantasenamat, C., Prachayasittikul, V., & Nuchnoi, P. (2019). *ThalPred: a web-based prediction tool for discriminating thalassemia trait and iron deficiency anemia*. 10.1186/s12911-019-0929-2
- Mazzu, T., Nogueira, D., Abubakar, O., Soares, A., Cândido, L., Florencio, S., Luporini, R., Vinícius, L., & Furtado, D. (2021). *Smartphone-based photo analysis for the evaluation of anemia, jaundice and COVID-19*. 10.1055/s-0041-1734014
- 16
- 19 Meitei, A., Saini, A., Mohapatra, B., & Singh, K. (2022). *Predicting child anaemia in the North-Eastern states of India: a machine learning approach*. 10.1007/s13198-022-01765-4
- 11 Ministerio de Salud. (2022). *Informe: Estado Nutricional de niños menores de cinco años y gestantes que acceden a los establecimientos de salud del Ministerio de Salud. Informe Gerencial Nacional. 2022 – I Semestre*. Ministerio de Salud.
- 8 Mitani, A., Huang, A., Venugopalan, S., Corrado, G., Peng, L., Webster, D., Hammel, N., Liu, Y., & Varadarajan, A. (2020). *Author Correction: Detection of anaemia from retinal fundus images via deep learning*. *Nature Biomedical Engineering*. 10.1038/s41551-020-0530-0
- 2 Sadan, C. (2019). *MODELO DE MINERÍA DE DATOS PARA LA PREDICCIÓN DE CASOS DE ANEMIA EN GESTANTES DE LA PROVINCIA DE ILO. UNIVERSIDAD NACIONAL DE MOQUEGUA*.
- 2 Sadan, C. (2019). *MODELO DE MINERÍA DE DATOS PARA LA PREDICCIÓN DE CASOS DE ANEMIA EN GESTANTES DE LA PROVINCIA DE ILO. UNIVERSIDAD NACIONAL DE MOQUEGUA*.
- 36 Suner, S. (2021). *Prediction of anemia and estimation of hemoglobin concentration using a smartphone camera* (nro 7 ed., Vol. vol. 16). PLoS One. 10.1371/journal.pone.0253495

30

Yıldız, T. K., Yurtay, N., & Öneç, B. (2021). *Classifying anemia types using artificial learning methods* (nro. 1 ed., Vol. vol. 24). Engineering Science and Technology, an

International Journal. 10.1016/j.jestch.2020.12.003

22

Zhang, J., & Tang, W. (2020). *Building a prediction model for iron deficiency anemia among infants in Shanghai, China* (nro 1 ed., Vol. vol. 8). Food Sci Nutr. doi: 10.1002/fsn3.1301

# Modelo predictivo para la detección de anemia en mujeres embarazadas.

## INFORME DE ORIGINALIDAD

31 %  
INDICE DE SIMILITUD

29 %  
FUENTES DE INTERNET

16 %  
PUBLICACIONES

16 %  
TRABAJS DEL  
ESTUDIANTE

## FUENTES PRIMARIAS

1 [www.mdpi.com](http://www.mdpi.com) 2 %  
Fuente de Internet

2 [cybertesis.unmsm.edu.pe](http://cybertesis.unmsm.edu.pe) 2 %  
Fuente de Internet

3 [datos.ins.gob.pe](http://datos.ins.gob.pe) 2 %  
Fuente de Internet

4 [ouci.dntb.gov.ua](http://ouci.dntb.gov.ua) 1 %  
Fuente de Internet

5 [repositorio.usmp.edu.pe](http://repositorio.usmp.edu.pe) 1 %  
Fuente de Internet

6 [github.com](https://github.com) 1 %  
Fuente de Internet

7 [link.springer.com](http://link.springer.com) 1 %  
Fuente de Internet

8 Submitted to National College of Ireland 1 %  
Trabajo del estudiante

9 [docplayer.es](http://docplayer.es)  
Fuente de Internet



1 %

10

[icame.balikesir.edu.tr](http://icame.balikesir.edu.tr)

Fuente de Internet

1 %

11

[web.ins.gob.pe](http://web.ins.gob.pe)

Fuente de Internet

1 %

12

[mdpi-res.com](http://mdpi-res.com)

Fuente de Internet

1 %

13

Submitted to National University College -  
Online

Trabajo del estudiante

1 %

14

[www.scielo.org.pe](http://www.scielo.org.pe)

Fuente de Internet

1 %

15

[neptune.ai](http://neptune.ai)

Fuente de Internet

1 %

16

[biodatamining.biomedcentral.com](http://biodatamining.biomedcentral.com)

Fuente de Internet

1 %

17

[www.javaer101.com](http://www.javaer101.com)

Fuente de Internet

1 %

18

[www.springerprofessional.de](http://www.springerprofessional.de)

Fuente de Internet

1 %

19

Submitted to The University of the West of  
Scotland

Trabajo del estudiante

1 %

20	<a href="https://dspace.cvut.cz">dspace.cvut.cz</a> Fuente de Internet	<1 %
21	<a href="http://www.anmat.gov.ar">www.anmat.gov.ar</a> Fuente de Internet	<1 %
22	Submitted to University of Utah Trabajo del estudiante	<1 %
23	<a href="http://bbmundo.com">bbmundo.com</a> Fuente de Internet	<1 %
24	Submitted to University of Stirling Trabajo del estudiante	<1 %
25	<a href="http://hdl.handle.net">hdl.handle.net</a> Fuente de Internet	<1 %
26	<a href="http://renati.sunedu.gob.pe">renati.sunedu.gob.pe</a> Fuente de Internet	<1 %
27	<a href="http://121044061159533.blogspot.com">121044061159533.blogspot.com</a> Fuente de Internet	<1 %
28	Submitted to Universidad Cesar Vallejo Trabajo del estudiante	<1 %
29	<a href="http://unam.edu.pe">unam.edu.pe</a> Fuente de Internet	<1 %
30	Submitted to Adtalem Global Education, Inc. Trabajo del estudiante	<1 %
31	Submitted to Universidad Privada San Juan Bautista	<1 %

32

[doaj.org](https://doaj.org)

Fuente de Internet

<1 %

33

[www.wjgnet.com](https://www.wjgnet.com)

Fuente de Internet

<1 %

34

Dimas Chaerul Ekty Saputra, Khamron Sunat, Tri Ratnaningsih. "A New Artificial Intelligence Approach Using Extreme Learning Machine as the Potentially Effective Model to Predict and Analyze the Diagnosis of Anemia", Healthcare, 2023

Publicación

<1 %

35

[acikerisim.omu.edu.tr](https://acikerisim.omu.edu.tr)

Fuente de Internet

<1 %

36

[cdn.t3kys.com](https://cdn.t3kys.com)

Fuente de Internet

<1 %

37

[www.sciencegate.app](https://www.sciencegate.app)

Fuente de Internet

<1 %

38

[www.slideshare.net](https://www.slideshare.net)

Fuente de Internet

<1 %

39

Sudha, Harkesh Sehrawat, Yudhvir Singh, Vivek Jaglan. "Machine Learning Approaches For Disease Prediction:- A Review", 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), 2022

Publicación

<1 %

40	<a href="https://prezi.com">prezi.com</a> Fuente de Internet	<1 %
41	<a href="https://fs-morente.filos.ucm.es">fs-morente.filos.ucm.es</a> Fuente de Internet	<1 %
42	<a href="https://pesquisa.bvsalud.org">pesquisa.bvsalud.org</a> Fuente de Internet	<1 %
43	<a href="https://www.nature.com">www.nature.com</a> Fuente de Internet	<1 %
44	<a href="https://www.seden.org">www.seden.org</a> Fuente de Internet	<1 %
45	<a href="https://gestiopolis.com">gestiopolis.com</a> Fuente de Internet	<1 %
46	<a href="https://repositorioacademico.upc.edu.pe">repositorioacademico.upc.edu.pe</a> Fuente de Internet	<1 %
47	<a href="https://www.coursehero.com">www.coursehero.com</a> Fuente de Internet	<1 %
48	<a href="https://www.net96.it">www.net96.it</a> Fuente de Internet	<1 %
49	<a href="https://etica.uazuay.edu.ec">etica.uazuay.edu.ec</a> Fuente de Internet	<1 %
50	<a href="https://gynmed.at">gynmed.at</a> Fuente de Internet	<1 %
51	<a href="https://issuu.com">issuu.com</a> Fuente de Internet	<1 %

52	<a href="https://repositorio.upct.es">repositorio.upct.es</a> Fuente de Internet	<1 %
53	<a href="https://www.mdsau.de.com">www.mdsau.de.com</a> Fuente de Internet	<1 %
54	Huimin Li, Jing Cao, Andrzej Grzybowski, Kai Jin, Lixia Lou, Juan Ye. "Diagnosing Systemic Disorders with AI Algorithms Based on Ocular Images", Healthcare, 2023 Publicación	<1 %
55	<a href="https://decisionaid.ohri.ca">decisionaid.ohri.ca</a> Fuente de Internet	<1 %
56	<a href="https://intellectum.unisabana.edu.co">intellectum.unisabana.edu.co</a> Fuente de Internet	<1 %
57	<a href="https://portal.amelica.org">portal.amelica.org</a> Fuente de Internet	<1 %
58	<a href="https://repositorio.puce.edu.ec">repositorio.puce.edu.ec</a> Fuente de Internet	<1 %
59	<a href="https://repositorio.uns.edu.pe">repositorio.uns.edu.pe</a> Fuente de Internet	<1 %
60	<a href="https://sedici.unlp.edu.ar">sedici.unlp.edu.ar</a> Fuente de Internet	<1 %
61	<a href="https://worldwidescience.org">worldwidescience.org</a> Fuente de Internet	<1 %
62	<a href="https://www.dimelo.com">www.dimelo.com</a> Fuente de Internet	<1 %

63	<a href="http://www.emagister.com">www.emagister.com</a> Fuente de Internet	<1 %
----	--	------

64	Justice Williams Asare, Peter Appiahene, Emmanuel Timmy Donkoh. "Detection of anaemia using medical images: A comparative study of machine learning algorithms – A systematic literature review", Informatics in Medicine Unlocked, 2023 Publicación	<1 %
----	---	------

65	S. Rogg, D. H. Fuertinger, S. Volkwein, F. Kappel, P. Kotanko. "Optimal EPO dosing in hemodialysis patients using a non-linear model predictive control approach", Journal of Mathematical Biology, 2019 Publicación	<1 %
----	---	------

66	<a href="http://dokumen.pub">dokumen.pub</a> Fuente de Internet	<1 %
----	--	------

Excluir citas

Apagado

Excluir coincidencias

Apagado

Excluir bibliografía

Apagado