



Tipo de artículo: Artículos originales  
Temática: Inteligencia artificial  
Recibido: 29/10/2023 | Aceptado: 04/01/2024 | Publicado: 30/03/2024

Identificadores persistentes:  
DOI: [10.48168/innosoft.s15.a156](https://doi.org/10.48168/innosoft.s15.a156)  
ARK: [ark:/42411/s15/a156](https://nbn-resolving.org/urn:ark:/42411/s15/a156)  
PURL: [42411/s15/a156](https://purl.org/urn:42411/s15/a156)

## Concientización sobre la obesidad en Latinoamérica en los centros de salud utilizando un árbol de decisión

### *Obesity awareness in Latin America in health centers using a decision tree*

Diego Moises Chuctaya Ruiz <sup>1</sup> [0000-0002-7296-4592]\*, Luis Pablo Condori Villalba <sup>2</sup> [0000-0002-7119-4856], Gilbert Wil Ramos Ticona <sup>3</sup> [0000-0001-2200-3333], Esteba Cruz Santos Adilson <sup>4</sup> [0000-0002-2094-5760]

<sup>1</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [dchuctayar@unsa.edu.pe](mailto:dchuctayar@unsa.edu.pe)

<sup>2</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [lcondorivill@unsa.edu.pe](mailto:lcondorivill@unsa.edu.pe)

<sup>3</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [gramosti@unsa.edu.pe](mailto:gramosti@unsa.edu.pe)

<sup>4</sup> Universidad Nacional de San Agustín. Arequipa, Perú. [sestebac@unsa.edu.pe](mailto:sestebac@unsa.edu.pe)

\* Autor para correspondencia: [dchuctayar@unsa.edu.pe](mailto:dchuctayar@unsa.edu.pe)

---

#### Resumen

El presente documento busca concientizar acerca de la obesidad localizada en Latinoamérica y sus centros de salud, incitando a reducir la obesidad en la población, tomando como herramienta un software que a partir de los datos tomados en la investigación pueda determinar los niveles de obesidad en un paciente utilizando técnicas relacionadas a la Inteligencia Artificial donde el modelo aplicado a los datos se puede observar cuales son las personas con un mayor grado de obesidad y analizar sus respectivas causas.

**Palabras clave:** Inteligencia Artificial, obesidad, árbol de decisión, entropía

#### Abstract

*This paper seeks to raise awareness about localized obesity in Latin America and its health centers, inciting to reduce obesity in the population, taking as a tool a software that from the data taken in the research can determine the levels of obesity in a patient and recommend some ways to improve health using techniques related to Artificial Intelligence where the model applied to the data can be observed which are the people with a higher degree of obesity and analyze their respective causes.*

**Keywords:** Artificial Intelligence, obesity, decision tree, entropy

---

## Introducción

En la actualidad la obesidad se le puede denominar “la Epidemia del Siglo 21” ya que abarca el problema que se presenta en forma reciente mundialmente [7]. El país que puede representar esta epidemia a grandes rasgos son los Estados Unidos donde el porcentaje de adultos que son obesos (IMC >30) subió de 15, 3% en 1995 a 23,9% en 2005 de los cuales el 4,8 % posee un IMC >40. La obesidad viene acompañada con riesgo aumentado de hipertensión, diabetes, hiperlipidemia y enfermedad coronaria; por lo cual se estima que la obesidad podría llevar a una disminución en la esperanza de vida en USA en el Siglo 21, siendo el indicador de cómo probablemente se muestren el los países latinoamericanos que son influenciados por este mismo [5].

En América tiene la prevalencia más alta de todas las regiones de la OMS, con 62,5% de los adultos con sobrepeso u obesidad (64.1% de los hombres y 60.9% de las mujeres) [8]. Donde la obesidad afecta tanto a adultos como niños; y en América Latina los centros de salud buscan alternativas para afrontar y detener el aumento de las tasas de obesidad, promoviendo y apoyando ideas políticas que permitan a las personas mejorar la alimentación, la actividad física y la salud en la Región de las Américas [6].

La justificación por la cual se realiza esta investigación es mejorar la forma de vida del paciente. Además que la obesidad no es un problema singular, sino que también puede traer otros problemas de salud como enfermedades del corazón, diabetes, enfermedades pulmonares, entre otros [1].

Como otro punto, se sabe que la obesidad afecta al ámbito psicosocial, ya que está asociada con la depresión, la baja autoestima, el bullying, entre otros. Por otro lado, en el ámbito laboral, aparece el absentismo laboral, mayor riesgo de accidentes, dificultad de movilidad, peor desempeño académico, somnolencia, falta de concentración, entre otros, y a la vez no está apto para trabajos que requieran mucha movilidad como podría ser el trabajo de un bombero o militar [2].

En el año 2012, el trabajo “Prevention of Obesity using Artificial Intelligence Techniques” realizado por Bouharati S. et al. nos habla que los desórdenes de obesidad varían de acuerdo a cada persona, hábitos e incluso el ambiente que lo rodea, por lo que una forma para prevenir la obesidad, tomando en cuenta todos estos factores ambiguos, es utilizar el modelo de lógica difusa y con ello poder reconocer patrones causantes de la obesidad [3].

En el año 2017, el trabajo “Artificial Intelligence technologies to manage obesity” realizado por Bruna Marmett et. al. nos habla que la enorme creciente de casos de obesidad los niveles de mortalidad va en aumento y un método eficaz para manejarlo es la Inteligencia Artificial. Estos sistemas son complejos por lo que para realizar uno se requiere de pruebas no tradicionales de métodos de verificación y validación [4].

Además se recalca que estos sistemas no buscan reemplazar al personal médico experto, sino que este sistema debe ser tratado como un apoyo en la decisión de los médicos para tener un diagnóstico más preciso.[3]

El objetivo trazado para la investigación es incitar a los centros de salud de Latinoamérica a reducir la obesidad en la población, tomando como herramienta un software que a partir de los datos tomados en la investigación pueda determinar los niveles de obesidad en un paciente y recomendar algunas formas para mejorar la salud. En esta ocasión los niveles de obesidad se calculan en función de la condición física y hábitos alimenticios.

## Marco referencial

### Trabajos Relacionados

1. Herrera, D. en su trabajo “Hábitos Alimentarios y su Relación con el Sobrepeso y Obesidad en Adolescentes en la Unidad Educativa Julio María Matovelle en el año 2016” sostiene que en la Unidad Educativa Matovelle en la ciudad de Quito, la población se realizó un estudio a adolescentes con edades comprendidas entre 12 y 18 años. El tipo de estudio empleado fue observacional y el enfoque cualicuantitativo analítico y de corte transversal. Para determinar el estado nutricional, se empleó el programa AnthroPlus OMS, en 722 estudiantes, correspondiente a la totalidad de alumnos, posteriormente se seleccionó mediante muestreo aleatorio sistemático a individuos con normopeso (n=40), sobrepeso (n=35) y obesidad (n=32) para evaluar los hábitos alimentarios. Los principales resultados indican que la prevalencia de sobrepeso y obesidad es de 28% (sobrepeso y obesidad), sin diferencias por sexo. Mientras al relacionar hábitos alimentarios de los normopeso con los de sobrepeso muestran diferencia estadísticamente significativa, lo que no demostró con el grupo de obesidad, donde se llegó a la conclusión de que la relación entre la existencia de sobrepeso y obesidad y los hábitos alimenticios en los estudiantes de la Unidad Educativa Matovelle, determina, que la mejor calidad de alimentación la tienen los adolescentes con estado nutricional normal, ya que su índice de calidad promedio es 5,12 que es mayor a los adolescentes con sobrepeso que tienen un índice de calidad promedio de 3,67 y los

- adolescentes con obesidad con un índice de calidad promedio de 4,40.[12]
2. Gardi, P., Gonzalo, L. y Medina, J. en su trabajo “Hábitos alimentarios y su relación con la obesidad en adolescentes” sostienen que los hábitos alimentarios no saludables constituyen factores de riesgo modificables, por lo que la identificación es importante para tomar medidas por lo que el objetivo del estudio fue estimar la prevalencia de obesidad en adolescentes de la Institución Educativa de tipo Experimental de la Universidad Nacional de Educación Enrique Guzmán y Valle e identificar su asociación con los hábitos alimentarios no saludables. El estudio se realizó en adolescentes de 14 a 16 años, en el periodo de septiembre 2016 a diciembre de 2016 tomando en cuenta los hábitos alimentarios y la obesidad. Se llegó a la conclusión de que el consumo de alimentos es una práctica que contribuye significativamente en los adolescentes de la Institución Educativa[9].
  3. Álvarez, N. en su trabajo “Alimentación Y Salud: La Obesidad Como Factor De Riesgo” tuvo como objetivo realizar una revisión y puesta al día del problema de la obesidad, desde un punto de vista global, por su frecuencia e importancia tanto a nivel mundial como en nuestro país, así como de los diversos factores que contribuyen a su desarrollo y mantenimiento y las consecuencias que produce en la salud y enfermedades asociadas. Este trabajo se realizó mediante el análisis de trabajos, reportes, informes y datos estadísticos relacionados a este tema. Se llegó a la conclusión de que la obesidad es una situación muy frecuente en los países desarrollados, que va aumentando progresivamente en las últimas décadas. Además, constituye un problema de Salud Pública a nivel mundial, pues influye negativamente en la calidad de vida de las personas que lo presentan[10].
  4. Chew, H., Ang, W., & Lau, Y. en su trabajo “The potential of artificial intelligence in enhancing adult weight loss: a scoping review” tiene como objetivo en su trabajo presentar una visión general de cómo podría utilizarse la inteligencia artificial (IA) para regular los comportamientos alimentarios y dietéticos, las conductas de ejercicio y la pérdida de peso. Este trabajo se realizó mediante búsqueda literaria en ocho bases de datos (CINAHL, Cochrane-Central, Embase, IEEE Xplore, PsycINFO, PubMed, Scopus y Web of Science). Se llegó a la conclusión de que el uso de la IA para la pérdida de peso aún no se ha desarrollado. Un punto importante es que se propone un marco sobre la aplicabilidad de la IA para la pérdida de peso, pero se advierte de su dependencia del compromiso y la contextualización[11].

## Marco teórico

1. **Metodología ágil:** Las metodologías ágiles son un formato más contemporáneo para el desarrollo de software donde funciona el compromiso activo del cliente, la comunicación recurrente entre el equipo de desarrollo y la entrega rápida del valor del cliente, de acuerdo con el Manifiesto Ágil [5]. Al tener un modelo de comunicación interna y externa muy fácil, ya que las reuniones de lanzamiento y planificación se llevan a cabo al comienzo de cada sprint, el cliente siempre tiene fácil acceso al prototipo desarrollado y le permite iterar entre lo que se está haciendo. desarrollado y el usuario final.

## Materiales y métodos

1. **Google Colab:** También conocido como "Collaboratory", te permite programar y ejecutar Python en tu navegador con las siguientes ventajas:
  - No requiere configuración
  - Da acceso gratuito a GPUs
  - Permite compartir contenido fácilmente

Esta herramienta fue utilizada por todos los integrantes de manera síncrona y asíncrona para la realización de nuestro árbol de decisión para el dataset respectivo a la obesidad.

2. **Lenguaje de programación Python:**

La principal implementación de árboles de decisión en Python está disponible en la librería scikit-learn a través de la clase DecisionTreeClassifier. Una característica importante para aquellos que han utilizado otras implementaciones es que, en scikit-learn, es necesario convertir las variables categóricas en variables dummy (one-hot-encoding).

3. **Árbol de decisión:** Estos proveen de una herramienta de clasificación muy potente. Su uso en el manejo de datos la hace ganar en popularidad dadas las posibilidades que brinda y la facilidad con que son comprendidos sus resultados por cualquier usuario. El árbol en sí mismo, al ser obtenido, determina una regla de decisión.

Esta técnica permite:

- Segmentación: establecer qué grupos son importantes para clasificar un cierto ítem.
- Clasificación: asignar ítems a uno de los grupos en que está particionada una población.
- Predicción: establecer reglas para hacer predicciones de ciertos eventos.
- Reducción de la dimensión de los datos: Identificar qué datos son los importantes para hacer modelos de un fenómeno.

- **Identificación-interrelación:** identificar qué variables y relaciones son importantes para ciertos grupos identificados a partir de analizar los datos.
- **Recodificación:** discretizar variables o establecer criterios cualitativos perdiendo la menor cantidad posible de información relevante.

**4. Dataset:** Es un conjunto de datos tabulados en cualquier sistema de almacenamiento de datos estructurados.

Nuestro conjunto de datos está distribuido de la siguiente manera:

- a. Cantidad de variables de entrada: 16
- b. Cantidad de instancias: 2111

Tabla 1. Variables de entrada del dataset

Variable	Definición	Tipo de variable	Dominio
Gender	Género	nominal	Female, Male
Age	Edad	discreta	numérico
Height	Tamaño	continua	numérico
Weight	Peso	continua	numérico
FHWO	antecedentes familiares con sobrepeso	nominal	si,no
FAVC	Consumo frecuente de alimentos con alto contenido calórico	nominal	si, no
FCVC	Frecuencia de consumo de verduras	continua	numérico
NCP	Número de comidas principales	discreta	numérico
CAEC	Consumo de alimentos entre horas	nominal	no, a veces, con frecuencia, siempre
SMOKE	fuma	nominal	si, no
CH2O	Consumo de agua diaria	continua	numérico

SCC	Monitoreo del consumo de calorías	nominal	si, no
FAF	Frecuencia de la actividad física	continua	numérico
TUE	Tiempo de uso de dispositivos tecnológicos	continua	numérico
CALC	Consumo de alcohol	nominal	no, a veces, con frecuencia, siempre
MTRANS	Transporte utilizado	nominal	Automóvil, Motocicleta, Bicicleta, Transporte Público, Caminata

Tabla 2. Variables de salida del dataset

Variable	Definición	Tipo de variable	Dominio
NObeyesdad	Nivel de obesidad	ordinal	Insufficient_Weight,Normal_Weight,Overweight_Level_I,Overweight_Level_II,Obesity_Type_I,Obesity_Type_II,Obesity_Type_III

## Resultados y discusión

Dentro de los resultados obtenidos se muestra la matriz de confusión el cual muestra que del 20% de datos de prueba, que son 423 casos, 402 generaron resultados correctos mientras que 21 generaron resultados incorrectos

```

Matriz de Confusión:
[[55  1  0  0  0  0  0]
 [ 2 36  0  0  0 10  0]
 [ 0  0 79  1  0  0  3]
 [ 0  0  1 51  0  0  0]
 [ 0  0  0  0 65  0  0]
 [ 0  2  0  0  0 63  1]
 [ 0  0  0  0  0  0 53]]
    
```

Figura 1: Matriz de confusión

Por lo tanto al momento de comparar los resultados bien clasificados con los que no fueron bien clasificados, nos damos cuenta que la exactitud del modelo es del 95.03%, tal como nos muestra los resultados del código al calcular dicha exactitud.

```
Exactitud del modelo:  
0.950354609929078
```

Figura 2: Exactitud del modelo

Para asegurarnos de que estos datos se pueden considerar fiables y ver que tan cerca está el resultado de una predicción del valor verdadero se realizó la medición de la precisión del modelo, el cual nos arroja una precisión del 95.16%

```
Precisión del modelo:  
0.9516427562434174
```

Figura 3: Precisión del modelo

En cuanto a la tasa de verdaderos positivos del modelo y por lo tanto para determinar la habilidad del modelo de detectar los casos relevantes se calculó la exhaustividad del modelo el cual nos da un valor del 95.03%. Este es un valor alto por lo que podemos decir que nuestro modelo es sensible, es decir, que no se le escapan muchos positivos.

```
Exhaustividad del modelo:  
0.950354609929078
```

Figura 4: Exhaustividad del modelo

Para combinar las medidas de precisión y exhaustividad y poder comparar este rendimiento combinado se calcula el valor F1 del modelo el cual es de 94.93%.

```
F1 del modelo:  
0.9493888070912593
```

Figura 5: Valor F1 del modelo

En la generación del árbol de decisión se tuvo en cuenta como nodos todas las variables predictoras ya que así conseguimos un árbol de decisión completo y que sea más preciso en el cual podemos ver que el atributo que se seleccionó como nodo raíz es el peso ya que este según su entropía es el que mayor ganancia de información tiene por lo que es más adecuado para ser el nodo inicial.



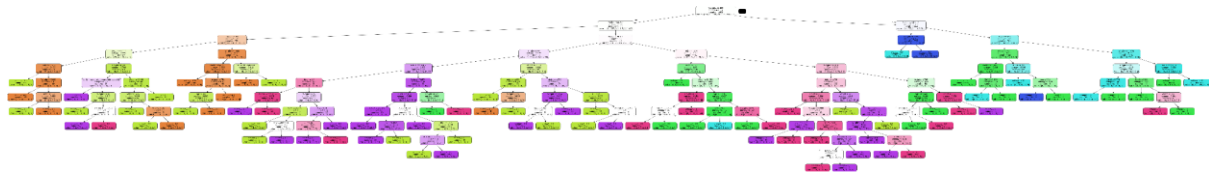


Figura 6: Árbol de decisión

Ya con el árbol de decisión generado podemos predecir el nivel de obesidad de una personas tomando en función de sus hábitos alimenticios y su condición física. Para interpretar este modelo se vio por conveniente generar un par de histogramas. Primeramente se está tomando en cuenta el nivel de obesidad de las personas en función de su edad. Podemos ver como el nivel de obesidad tipo 3, el más alto, es más predominante en las personas con edades entre los 20 y 30 años.

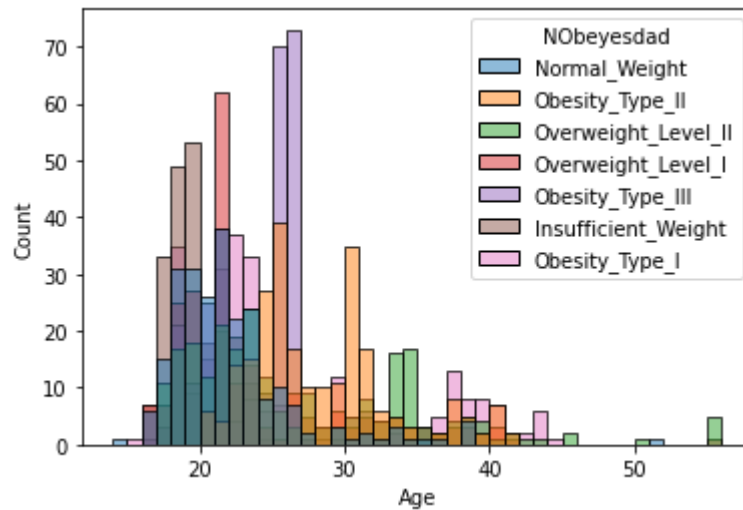


Figura 7: Histograma del nivel de obesidad en función de la edad

En el histograma generado con el nivel de obesidad respecto a la frecuencia de la actividad física de una persona, se visualiza que las personas de obesidad tipo 3, son las que menos se ejercitan y las que más se ejercitan son las personas de peso normal.

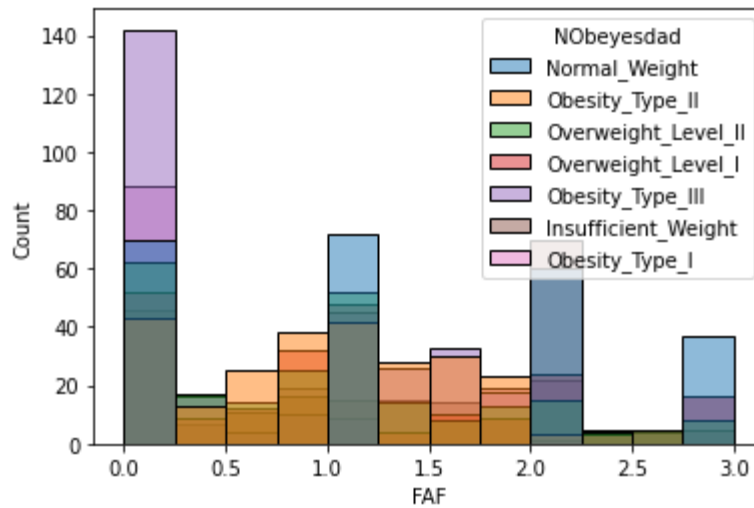


Figura 8: Histograma del nivel de obesidad en función de la Frecuencia de la actividad física(FAF)

En el histograma generado con el nivel de obesidad respecto al tiempo del uso de dispositivos de una persona, se visualiza que la gran parte de las personas que participaron en las encuestas no utilizan mucho los dispositivos.

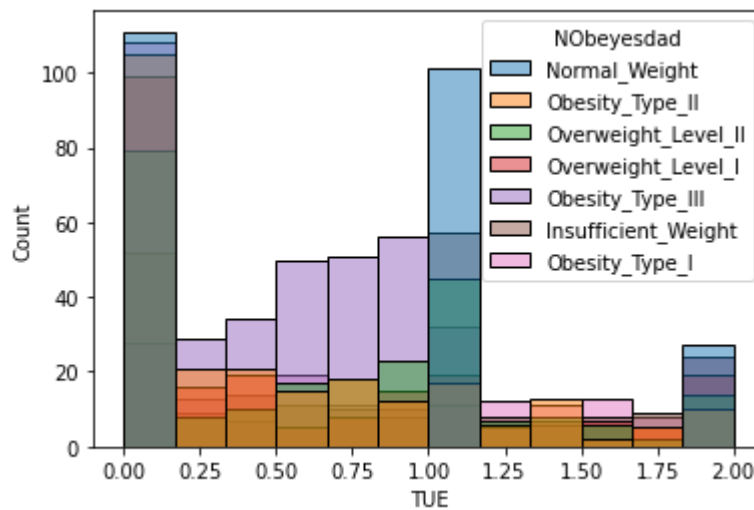


Figura 9: Histograma del nivel de obesidad en función del Tiempo de uso de dispositivos tecnológicos(TUE)

En el histograma generado con el nivel de obesidad respecto al consumo de agua diaria de una persona, se visualiza que la mayoría de las personas encuestadas consumen 2 litros de agua diaria. Las personas que toman de uno a dos litros de

agua diaria son mayormente las personas de peso normal. Los que beben más agua son las personas de obesidad tipo 1 y 3.

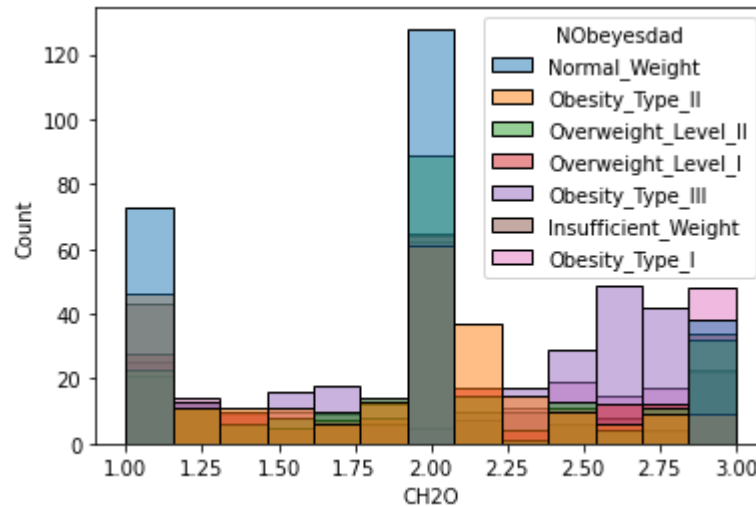


Figura 10: Histograma del nivel de obesidad en función del Consumo de agua diaria(CH2O)

## Conclusiones

En el modelo aplicado a los datos se puede observar que las personas con un grado mayor de obesidad se encuentran entre los 20 a 30 años. También se pudo ver que una de las causas por el cual el grado de obesidad sube es porque las personas no se ejercitan, siendo una de las razones el sedentarismo causado en los últimos años por la pandemia y la evolución de la tecnología. Por esta razón es que es importante que los centros de salud deben de actuar ante estos sucesos, para que las personas que residen en Latinoamérica tengan un mejor estilo de vida.

El proyecto contribuye en el campo de la salud, específicamente en el ámbito de una vida saludable tratando de concientizar sobre un problema como lo es la obesidad. Por otro lado, el modelo de predicción aplicado puede ser muy útil para calcular el grado de obesidad de una persona respondiendo un par de preguntas simples.

Como trabajo futuro sería empezar a visitar diversas instituciones para concientizar a las personas de Latinoamérica sobre la obesidad, para esto se puede comenzar por los centros educativos, ya que los niños suelen ser los más vulnerables a sufrir bullying y otros tipos de discriminación.

## Contribución de Autoría

**Diego Moises Chuctaya Ruiz:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Software](#), [Validación](#), [Redacción - borrador original](#). **Luis Pablo Condori Villalba:** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Recursos](#), [Visualización](#), [Supervisión](#), [Administración de proyectos](#), [Adquisición de fondos](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Gilbert Wil Ramos Ticona** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Recursos](#), [Visualización](#), [Supervisión](#), [Administración de proyectos](#), [Adquisición de fondos](#), [Curación de datos](#), [Escritura, revisión y edición](#). **Esteba Cruz Santos Adilson** [Conceptualización](#), [Investigación](#), [Metodología](#), [Análisis formal](#), [Recursos](#), [Visualización](#), [Supervisión](#), [Administración de proyectos](#), [Adquisición de fondos](#), [Curación de datos](#), [Escritura, revisión y edición](#).

## Referencias

- [1]. Miguel Soca, P. y Niño Peña, A., 2009. Consecuencias de la obesidad . [en línea] Scielo.sld.cu. Disponible en: <[http://scielo.sld.cu/scielo.php?script=sci\\_arttext&pid=S1024-94352009001000006](http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1024-94352009001000006)> [Consultado el 25 de junio de 2022].
- [2]. Martínez Gárate, I., Valdés del Olmo, L., Bayona González, A. y Martínez Castellanos, J., 2021. Relación entre la obesidad y el estrés laboral: una revisión sistemática . [en línea] scielo.isciii. Disponible en: <[https://scielo.isciii.es/scielo.php?pid=S0465-546X2021000200112&script=sci\\_arttext&tlng=pt](https://scielo.isciii.es/scielo.php?pid=S0465-546X2021000200112&script=sci_arttext&tlng=pt)> [Consultado el 25 de junio de 2022].
- [3]. Bouharati S. et. al., 2012, Prevention of Obesity using Artificial Intelligence Techniques. International Journal of Science and Engineering Investigations. vol. 1, num. 9.
- [4]. Marmett, B., Böek Carvalho, R., Santos Fortes, M., Cazellab, S., 2017. Artificial Intelligence technologies to manage obesity. vol. 30, num. 2. (2018). pag. 73-79. DOI: <https://doi.org/10.14295/vittalle.v30i2.7654>
- [5]. Parvez Hossain, Bisher Kavar, Meguid El Nahas. Obesity and Diabetes in the Developing World. A Growing Challenge. New Engl J Med 2007; 356: 213- 215
- [6]. Global Strategy on Diet, Physical Activity and Health [Internet]. OMS Organización Mundial de la Salud. 2017. [Citado 12 noviembre 2017]. Recuperado a partir de: <http://www.who.int/dietphysicalactivity/goals/en/>
- [7]. Juan A Rivera, Simón Barquera, Fabricio Campirano, Ismael Campos, Margarita Safdie, Víctor Tovar. Epidemiological and nutritional transition in Mexico: rapid increase of non-communicable chronic diseases and obesity. Public Health Nutrition 2002; 5(1A): 113-122

- [8]. Aguirre B. H., García T. J. F., Vázquez H. M. C., Alvarado A. M., Romero Z. H. Panorama general y programas de protección de seguridad alimentaria en México. Rev. Méd Electrón [Internet]. 2017 [citado:12-noviembre-2017]; 39 Supl 1: S741-749. Recuperado a partir de: <http://www.revmedicaelectronica.sld.cu/index.php/rme/article/view/2124/3525>
- [9]. Gardi, P., Gonzalo, L. y Medina, J. 2019. Hábitos alimentarios y su relación con la obesidad en adolescentes. [Tesis de Licenciado en Nutrición Humana]. Universidad Nacional de Educación Enrique Guzmán y Valle
- [10]. Álvarez, N. 2019. ALIMENTACIÓN Y SALUD: LA OBESIDAD COMO FACTOR DE RIESGO, Volumen II. Número 17. Recuperado a partir de: <https://www.npunto.es/revista/17/alimentacion-y-salud-la-obesidad-como-factor-de-riesgo>
- [11]. Chew, H., Ang, W., & Lau, Y. 2021. The potential of artificial intelligence in enhancing adult weight loss: a scoping review. Public health nutrition, 24(8), 1993–2020. <https://doi.org/10.1017/S1368980021000598>
- [12]. Herrera, D. 2016. Hábitos Alimentarios y su Relación con el Sobrepeso y Obesidad en Adolescentes en la Unidad Educativa Julio María Matovelle en el año 2016. Pontificia Universidad Católica del Ecuador. Recuperado a partir de: <https://core.ac.uk/download/pdf/143442581.pdf>

## Anexos

- Estimation of obesity levels based on eating habits and physical condition:  
<https://archive.ics.uci.edu/ml/datasets/Estimation+of+obesity+levels+based+on+eating+habits+and+physical+condition+>