UNIVERSIDAD POLITÉCNICA DE MADRID

E.T.S. DE INGENIERÍA DE SISTEMAS INFORMÁTICOS

PROYECTO FIN DE GRADO

GRADO EN INGENIERÍA DEL SOFTWARE

## Planificación nutricional mediante algoritmos evolutivos

Autor: Javier Quesada Pajares

**Director:** Cristian Ramírez Atencia

Madrid, 5 de junio de 2024

Planificación nutricional mediante algoritmos evolutivos

Proyecto Fin de Grado, 5 de junio de 2024

**Autor**: Javier Quesada Pajares **Director**: Cristian Ramírez Atencia

#### E.T.S. de Ingeniería de Sistemas Informáticos

Campus Sur UPM, Carretera de Valencia (A-3), km. 7 28031, Madrid, España

Si deseas citar este trabajo, la entrada completa en BiBTEX es la siguiente:

```
@mastersthesis{citekey,
title = {Planificación nutricional mediante algoritmos
evolutivos},
author = { \& }
school = {E.T.S. de Ingeniería de Sistemas Informáticos},
year = {2024},
month = {6},
type = {Proyecto Fin de Grado};
```

Esta obra está bajo una licencia Creative Commons «Atribución-NoComercial-Compartirlgual 4.0 Internacional». Obra derivada de https://github.com/blazaid/UPM-Report-Template.



Todo cambio respecto a la obra original es responsabilidad exclusiva del presente autor.

## **Agradecimientos**

Aquí los agradecimientos que quieras dar. Y si no quieres, borras la entrada \acknowledgements de report.tex y ya está.

## Resumen

Para el final.

Condensado no quiere decir incompleto. Debe contener la información más destacable. Lo ideal es que ocupe entre media y una cara de un folio A4. Comenzará por el propósito y principales objetivos de la memoria. Luego hablaremos sobre los aspectos más destacables de la metodología empleada, seguido de los resultados obtenidos. Por último se presentarán las conclusiones de forma condensada.

Debe tener un estilo claro y conciso, sin ambigüedades de ningún tipo. Además, al ser un resumen de todo el contenido, ni que decir tiene que deberá ser lo último que elaboraremos, y deberá mantener una absoluta fidelidad con el contenido de la memoria.

**Palabras clave**: Cuatro o cinco; Expresiones clave; Que resuman; Nuestro proyecto o; Investigación

## **Abstract**

This section must contain the summary that we have written before in Spanish, but in English, as well as the keywords.

**Keywords**: Four or five; Key Expressions; Summarising; Our Project or; Research

# Índice general

1 Introducción						
	1.1	Contexto	1			
	1.2	Motivación	2			
	1.3	Justificación	3			
	1.4	Objetivos	4			
	1.5	Estructura de la memoria	4			
2	Estado del arte					
	2.1	Inteligencia artificial	5			
	2.2	Algoritmos evolutivos	7			
	2.3	Planificación nutricional mediante algoritmos evolutivos	10			
3	Marco teórico					
	3.1	Población	13			
	3.2	Evaluación	14			
	3.3	Operadores	16			
4	Des	sarrollo	21			
	4.1	Base de datos	21			
	42	Nutrición	22			

# Índice de figuras

1.1	Prevalencia (%) de obesidad y exceso de peso por grupos de sexo y edad. Fuente [8]
2.1	Imagen generada con DALL-E 3
3.1	Algoritmo genético simple
3.2	Estructura de un cromosoma
3.3	Operadores
3.4	Cruce de un punto
3.5	Mutación uniforme

# Índice de listados

2.1	Algoritmo genéti	CO	 8

## Introducción

## 1.1. Contexto

En los últimos años, la inteligencia artificial (IA) se ha posicionado como una de las herramientas más útiles e interesantes de las que disponemos. El término «inteligencia artificial» fue acuñado por primera vez por John McCarthy en la Conferencia de Darmouth de 1956 [1]. Años después de que Alan Turing formulase la pregunta sobre si las máquinas podían pensar y plantease el famoso Test de Turing [2], varios científicos se reunieron con el objetivo de discutir acerca de la posiblidad de un artefacto de comportarse de manera inteligente. Se llegó a la conclusión de que todo aspecto del aprendizaje se puede describir con tanta precisión que resulte factible construir una máquina que los simule.

La IA se enfoca en crear sistemas que puedan realizar tareas que normalmente requerirían de inteligencia humana. Aunque ha tenido un reciente auge debido a los chatbots o los asistentes personales, presenta una gran cantidad de finalidades. Dentro de los usos que se le dan a la IA destacan la creación y el análisis de productos o la automatización de servicios, además de la optimización de procesos. Esta optimización se enfoca en encontrar la mejor solución posible a un problema dado dentro de un conjunto de opciones factibles. Los algoritmos de optimización y de personalización permiten, por ejemplo, crear aplicaciones que permiten organizar tareas de manera inteligente [3] o termostatos que ajustan la calefacción automaticamente [4].

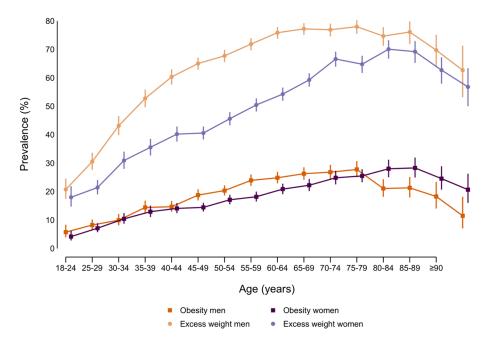
Este Proyecto Fin de Grado (PFG) se centrará en la resolución de un problema de optimización, la creación de un menú semanal de comidas personalizado que cumpla distintos objetivos, como el número de calorías diarias o la cantidad de macronutrientes ingeridos. Se hará uso de la computación evolutiva que, mediante algoritmos genéticos, diseñará una dieta equilibrada a partir de la selección, cruce y mutación de los distintos alimentos.

1.2. Motivación 2

## 1.2. Motivación

La motivación para realizar este proyecto viene dada por el interés creciente en el área de la inteligencia artificial y cómo ha cambiado la manera de plantear los problemas respecto al pasado. En España, el sector TIC (Tecnologías de la Información y la Comunicación) cuenta con más del 40 % de empresas que usan estas herramientas para la automatización de flujos de trabajo, análisis de datos o para la gestión de la cadena de suministro [5]. Buscan mejorar la precisión y la eficiencia a la hora de diseñar soluciones.

No obstante, no solo en sectores tecnológicos se hace uso de la IA. El médico o el alimentario también están incorporándola de manera gradual. Los diagnósticos de imágenes médicas [6] o la agricultura de precisión [7] son cada vez más comunes. También en la nutrición, relacionada con esto ámbitos, estas tecnologías permiten nuevas posibilidades.



**Figura 1.1**. Prevalencia (%) de obesidad y exceso de peso por grupos de sexo y edad. Fuente [8]

Según un estudio llevado a cabo por la Agencia Española de Seguridad Alimentaria y Nutrición (AESAN) y por el Instituto de Salud Carlos III (ISCIII), un 55,8 % de la población adulta española tiene exceso de peso y un 18,7 % padece obesidad, tal y como se muestra en la figura 1.1. Esto trae consigo múltiples

1.3. Justificación 3

problemas de salud, como la aparición de enfermedades crónicas o cardiovasculares, complicaciones respiratorias o dificultades al moverse.

Estos datos demuestran la importancia de una buena nutrición y hábitos alimenticios saludables. La planificación nutricional mediante algoritmos genéticos ayuda a comprender cómo se pueden mejorar estos hábitos a través de estas técnicas innovadoras, además de promover una vida sana y bienestar para todos, lo que se relaciona con los Objetivos de Desarrollo Sostenible (ODS) [9], por lo que es un gran aliciente a la hora de realizar este PFG.

## 1.3. Justificación

En esta sección se deben explicar y argumentar las razones por las cuales se eligió el tema del proyecto, así como su importancia y relevancia. Algunos elementos clave que se pueden abordar en esta sección son:

- Relevancia del tema: ¿Existe alguna necesidad o problema específico que tu proyecto pueda abordar?
- 2. **Justificación teórica**: Mención sobre qué teorías, enfoques o modelos existentes en la literatura respalden la importancia de abordar este tema.
- 3. **Brecha en el conocimiento**: ¿Qué aspectos no se han explorado lo suficiente o no han sido abordados en estudios previos? ¿Cómo puede el proyecto contribuir a cerrar esa brecha en el conocimiento?
- 4. **Contribución práctica**: Aplicaciones del proyecto y cómo pueden beneficiar a la comunidad académica, profesional o a la sociedad en general.

La sección no tiene por qué ser demasiado extensa, ni tiene por qué incluir (o limitarse) a los puntos anteriores, pero debe ser lo suficientemente clara y convincente para que los lectores comprendan por qué el proyecto es relevante y necesario.

1.4. Objetivos 4

## 1.4. Objetivos

El objetivo general de este proyecto es la creación de un planificación semanal de comidas mediante algoritmos genéticos. Experimentando con diferentes algoritmos se busca crear un menú que cumpla con distintos objetivos nutricionales, como la ingesta calórica diaria, a la vez que limitarse según las restricciones dadas. Los objetivos específicos de este PFG son:

- Desarrollar un algoritmo evolutivo capaz de generar menús que cumplan con las restricciones y los objetivos nutricionales establecidos.
- Personalizar el algoritmo para la variación de las comidas según las necesidades específicas de los individuos.
- Experimentar con distintas configuraciones y variantes del algoritmo genético en busca de encontrar la mejor solución posible.
- Evaluar la sensibilidad y eficacia del algoritmo.
- Ejecutar pruebas que validen el algoritmo.
- Documentar los resultados obtenidos.

## 1.5. Estructura de la memoria

En este subapartado se explicará la estructura del documento.

En el capítulo 2, Estado del arte, se centra en dar un enfoque general al proyecto. Se da contexto sobre las áreas en las que se basa el proyecto, la inteligencia artificial y los algoritmos evolutivos, además de sobre el tema central del PFG, la planificación nutricional mediante algoritmos evolutivos.

En el capítulo 3, Marco teórico, se aporta la base teórica que en la que se desarrolla el proyecto. Se da una explicación general sobre la inteligencia artificial y se indaga en el algoritmo evolutivo y su funcionamiento.

En el capítulo 4, Desarrollo,

## 2.1. Inteligencia artificial

La inteligencia artificial se puede definir como el estudio y diseño de agentes inteligentes, es decir, de sistemas que perciben su entorno y toman decisiones o acciones que maximizan sus posibilidades de éxito. Este campo se basa en disciplinas como la informática, la lógica o la neurociencia, que contribuyen a la simulación de capacidades cognitivas humanas en máquinas.

Los agentes inteligentes, que son la base de este campo, se clasifican según la capacidad de reconocer y actuar en su entorno. Podemos encontrar desde agentes simples, que responden directamente a estímulos, hasta agentes basados en utilidad, que evalúan si los resultados de sus acciones son satisfactorios. Es fundamental la racionalidad, la habilidad de realizar elecciones óptimas que maximicen la posibilidad de alcanzar objetivos. Utilizando distintas herramientas de lógica, los agentes inteligentes pueden formular y modificar conocimientos, deduciendo nueva información.

El razonamiento lógico es también fundamental para el desarrollo de algoritmos evolutivos, donde las decisiones sobre selección, cruzamiento o reproducción se basan en decisiones lógicas. Subclase de los métodos de aprendizaje automático inspirados en los procesos biológicos de evolución, aplican estos principios para desarrollar soluciones a complejos problemas.

El campo de la IA en la actualidad se encuentra en un estado de rápida evolución, con una infinidad de usos que se extiende por una gran variedad de sectores económicos y sociales. Su capacidad para analizar y procesar grandes cantidades de datos o para mejorar la eficiencia en distintas industrias han hecho que se trate de una tecnología en auge. El estudio de McKinsey & Company (2022) [10] estima que el 50 % de las empresas ya usan IA en sus tareas diarias, donde destacan la optimización de servicios, la creación de pro-

ductos y el análisis del servicio al cliente.

El informe Al Index Report 2024 [11] muestra que la IA ya supera a las habilidades humanas en ciertas tareas, como en la clasificación de imágenes, con una precisión del 97 % respecto al 95 % de los humanos, o en juegos, como el ajedrez, donde supera consistentemente a los jugadores humanos.

Aunque sin duda, el mayor crecimiento en los dos últimos años corresponde a las herramientas relacionas con la *Inteligencia Artificial Generativa (IAG)*. Esta rama se centra en crear modelos capaces de generar contenido, como conversaciones, imágenes, videos o música. Aprende de datos ya existentes y produce nuevos con características similares. En este ámbito destaca la empresa OpenAI, que cuenta con ChatGPT, capaz de generar texto lógico y actuar como chatbot, o con Dall-E, que puede generar imágenes en base a la descripción que se entregue, como la que se muestra en la figura 2.1. The state of AI in 2023: Generative AI's breakout year [12] muestra que el 79 % de los encuestados dice haber tenido al menos alguna exposición a la IAG, lo que indica que es la herramienta de IA que más rápido está captando el interés del público general.



Figura 2.1. Imagen generada con DALL-E 3.

Si bien existen diversas formas de clasificar la IA, una de las más interesantes respecto a este PFG es la basada en la representación y el procesamiento del conocimiento. Según este tipo de clasificación, existen dos tipos: la simbólica y la subsimbólica. La IA simbólica utiliza representaciones explícitas y legibles para los humanos, como símbolos y reglas lógicas, para resolver problemas. Consiste en incorporar conocimientos humanos y reglas de comportamiento a programas informáticos. Por otro lado, la subsimbólica se centra en métodos que no dependen de representaciones explícitas y se inspiran en

procesos naturales.

En las últimas décadas, la IA subsimbólica ha sido la más desarrollada, destacándose en tres campos principales.

Primero, el aprendizaje automático, particularmente a través de redes neuronales, que permite aprender y procesar grandes volúmenes de datos. Un ejemplo de esto AlphaFold de DeepMind, que utiliza IA para predecir las estructuras de proteínas, acelerando la investigación científica para la creación de nuevos medicamentos [13].

El segundo es la lógica borrosa, que simula la forma en la que los humanos toman decisiones en situaciones de incertidumbre. Un ejemplo interesante de esto es su uso en la industria automotriz para mejorar la estabilidad del vehículo mediante la regulación adaptativa de los amortiguadores en respuesta a las condiciones del camino y las dinámicas del vehículo [14].

Por último, la computación evolutiva, que emplea procesos evolutivos biológicos para la resolución de problemas. Dentro de esta rama destacan principalmente dos tipos de algoritmos. Uno de ellos son los algoritmos de enjambre. Inspirados en el comportamiento colectivo de sistemas naturales como colonias de hormigas o bandadas de aves, son métodos de optimización que solucionan problemas complejos mediante la colaboración de múltiples agentes simples. Un uso actual destacado de estos algoritmos es en la optimización de granjas eólicas, donde se emplean para mejorar la ubicación de turbinas y maximizar la eficiencia energética [15]. La otra son los algoritmos evolutivos, la que se va a tratar en más detalle a lo largo de este proyecto.

## 2.2. Algoritmos evolutivos

Un algoritmo se puede definir como un procedimiento computacional bien definido que toma un valor, o un conjunto de valores, como entrada y produce un valor, o un conjunto de valores, como salida. Se trata de un conjunto de instrucciones que permiten realizar una actividad mediante sucesivos pasos.

Un algoritmo evolutivo es un tipo de algoritmo que proviene de la computación evolutiva. Esta rama de la IA emplea principios inspirados en la evolución biológica para la resolución de problemas. Teoría propuesta por Charles Darwin en 1859, expone que en la naturaleza, en un entorno dado que puede albergar un número limitado de individuos, la selección natural favorecería a aquellos que poseyeran características que les permitiesen adaptarse mejor al medio, teniendo una mayor posibilidad de sobrevivir y reproducirse. A lo largo del tiempo, las características ventajosas se propagan a través de las generaciones, mientras que aquellas menos favorables tienden a desaparecer. Por lo tanto, estas poblaciones irán evolucionado y adaptándose gradualmente al entorno.

Los algoritmos evolutivos han tenido varios proyectos que han marcado el desarrollo de esta tecnología. Los primeros acercamientos tuvieron lugar en las década de 1960 y 1970. Por un lado, Lawrence J. Fogel (1966) [16] comenzó a explorar la programación evolutiva, centrándose en la evolución de autómatas finitos. Uno de los primeros intentos de aplicar principios evolutivos a la informática. Por otro lado, Rechenberg (1973) [17] desarrolló estrategias de evolución para problemas de optimización. Centradas en la selección y en la mutación, resultaron ser muy útiles para aplicarlas en la realidad.

Tras estos primeros pasos, John Holland fue fundamental para el desarrollo de los algoritmos genéticos. Su libro "Adaptation in Natural and Artificial Systems" del año 1975 [18] es básico para entender el funcionamiento de los mismos. Su trabajo se centró en la idea de que la evolución biológica podía ser simulada y utilizada para resolver problemas complejos en computación. Hacen uso de los operadores como la selección, el cruce o la mutación para evolucionar una población candidata de individuos hacia una mejor solución. Cada solución es evaluada según una función de fitness, que mide qué tan buena es la solución al problema en cuestión. En el listado 2.1 se puede observar el funcionamiento de un algoritmo genético básico.

Listado 2.1. Algoritmo genético.

#### **INICIAR**

INICIALIZAR población EVALUAR fitness REPETIR HASTA CUMPLIR condición de parada: SELECCIONAR individuos CRUZAR padres MUTAR hijos
EVALUAR individuos
FORMAR nueva generación
RETORNAR solución
FINALIZAR

Para que un AG se vaya reconduciendo hacia soluciones más favorables, se hace uso de las restricciones. Se usan diversas técnicas, como penalizar las soluciones que violan las restricciones, emplear métodos de reparación para modificar las soluciones no válidas o descartar directamente las soluciones que no sean óptimas. Estas estrategias permiten que los algoritmos genéticos mantengan la diversidad de la población a la vez que mejora la convergncia hacia una solución válida.

Después de que Holland sentara las bases de los algoritmos genéticos, en los años siguientes fueron surgiendo estudios que hicieron evolucionar la rama. Varios de los más importantes fueron de John Koza [19] [20], quien introdujo la programación genética, técnica usada para desarrollar automáticamente programas que realicen una tarea definida por el usuario. Se optimiza una población de individuos (programas) respecto a una función de aptitud. Koza probó su viabilidad para resolver problemas de robótica o de optimización.

Tras él han seguido surgiendo nuevas técnicas dentro de los algoritmos evolutivos. Entre ellas se puede destacar el desarrollo de la *Programación genética cartesiana (CGP)*, que sustituye los árboles de búsqueda usados en la programación genética tradicional por grafos dirigidos acíclicos, lo que es muy útil, por ejemplo, en el diseño de circuitos electrónicos [21].

Otro avance que se ha popularizado es el de los *Algoritmos genéticos híbridos (AGH)*, que combinan los algoritmos genéticos con otras técnicas de optimización para mejorar las soluciones a los problemas complejos. En esta subrama destacan los algoritmos meméticos. Estos combinan algoritmos evolutivos con búsqueda local para mejorar la calidad y velocidad de convergencia de las soluciones. Después del cruce y de la mutación, cada nueva solución se refina explorando soluciones cercanas para encontrar mejoras. [22]

A lo largo de los años se han desarrollado distintos proyectos pioneros que han hecho uso de los algoritmos genéticos. Se puede destacar la misión de la

NASA Space Technology 5 (ST5), donde ayudaron a la creación de una antena ultracompacta, con resultados extraordinarios [23] [24]. Además, también han sido empleados en áreas tan diversas como la economía o la bioingeniería, donde puede ayudar a tareas tan diversas como predecir movimientos del mercado [25] o a modelar secuencias genéticas [26], respectivamente.

## 2.3. Planificación nutricional mediante algoritmos evolutivos

Los primeros acercamientos para intentar resolver problemas de optimización nutricional se dan en la década de 1940. George Stigler (1945) [27] planteó el problema de encontrar la dieta de menor coste que cumpliese con unos objetivos nutricionales. Al no poseer ordenadores, utilizó técnicas manuales.

El problema fue formalmente resuelto dos años después por Jack Laderman usando programación lineal. Fue capaz de calcular la combinación de alimentos que cumpliese con los requisitos económicos y nutricionales. Con esto se demostró la utilidad de técnicas computacionales en tareas de optimización.

Tras los estudios de Holland, en la década de los 70 y 80 empezaron a aparecer distintos trabajos que hacían uso de los algoritmos genéticos para resolver problemas de optimización. Si bien estos no estaban relacionados directamente con la planificación nutricional, sí que sentaron las bases para la resolución de este tipo de problemas. Se puede destacar los trabajos de Scott Kirkpatrick et al. (1983) [28] y de David E. Goldberg (1989) [29], donde hacen uso de estos algoritmos para optimizar problemas con múltiples restricciones.

Fue a partir de la década de los 2000 cuando aparecieron muchos de los trabajos y estudios relacionados con la planificación nutricional.

EL trabajo de Kahraman y Seven (2005) [30] desarrolla un algoritmo genético bi-objetivo que propone comidas diarias saludables basados en parámetros especificados por el usuario a través de la interfaz gráfica, como la edad o el género. Optimiza la selección de platos para cumplir con restricciones nutricionales, minimizar costos y maximizar las preferencias del usuario.

Kaldirim y Köse (2006) [31] hacen una continuación de este trabajo. Hace uso del algoritmo multiobjetivo NSGA-II para gestionar de manera separada los objetivos de coste mínimo y máxima satisfacción, mejorando el manejo de restricciones nutricionales con la inclusión de límites. También presenta una mejora a la interfaz gráfica del anterior proyecto.

El estudio de Kashima et al. (2009) [32] presenta una interesante diferencia respecto a trabajos anteriores. Busca crear una aplicación web para ofrecer un servicio para compartir menús, creados para cada individuo haciendo uso de algoritmos genéticos, entre una comunidad de usuarios con el objetivo de fomentar los hábitos saludables.

Heinonen y Juuso (2016) [33] presentan el uso que le dan a los algoritmos genéticos en su aplicación Nutri-Flow, un software que proporciona guías dietéticas personalizadas. Emplea un sistema experto difuso (Fuzzy Expert System, FES) junto con algoritmos genéticos para optimizar las recomendaciones. El FES evalúa los alimentos según sus características nutricionales y su contribución a las necesidades dietéticas individuales, mientras que los AG buscan la combinación óptima de estos alimentos para ajustar la ingesta diaria.

Ya en esta década han seguido surgiendo proyectos pioneros que hacen evolucionar la rama de los algoritmos genéticos.

En el trabajo realizado por Kilicarslan et al. (2021) [34] se propone un modelo híbrido que combina algoritmos genéticos con aprendizaje profundo para la predicción y clasificación de anemias nutricionales. Optimiza los parámetros de los algoritmos de aprendizaje mediante computación evolutiva, lo que mejora la precisión de la planificación nutricional del paciente.

Joanne B. Cole y Rosita Gabbiannelli (2022) [35] muestran cómo se puede integrar la IA con el análisis genético para la nutrición personalizada del paciente. Basándose en el perfil genético y otros datos biométricos, es posible, junto a los algoritmos evolutivos, generar planes de comida y predicciones sobre la salud del usuario.

## Marco teórico

En 1975 John Holland propone imitar los procesos biológicos naturales que rigen la selección natural usando ordenadores, lo que sería el principio de los algoritmos genéticos (AGs) [18]. En un AG, las soluciones son modeladas como individuos o cromosomas, que generalmente se representan como cadenas de bits, aunque también se puede usar otro tipo de cadenas. Estas soluciones son evaluadas por una función de aptitud o fitness, y las más adecuadas son selecciondas para reproducirse mediante cruzamiento y mutación, procesos que mezclan y alteran aleatoriamente los cromosomas para generar diversidad. Los descendientes resultantes forman nuevas generaciones que vuelven a ser evaluadas, creando un ciclo que se repite hasta cumplir alguna condición de parada, como pudiera ser un número limitado de generaciones o que se alcance la solución deseada.

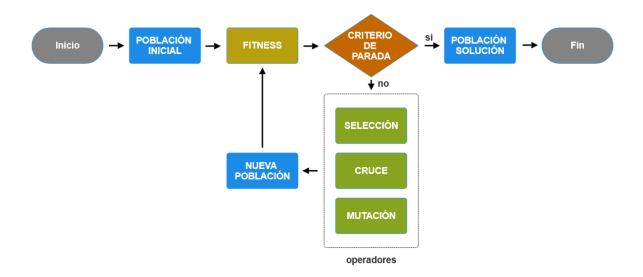


Figura 3.1. Algoritmo genético simple.

Explicado el concepto general, se va a desglosar cada una de las fases de la figura 3.1, que muestra el esquema básico de un algoritmo genético.

3.1. Población

## 3.1. Población

Antes de explicar la generación de una población inicial, es necesario conocer algunos conceptos que son usados para representar y entender las soluciones. Son términos que son usados en la biología y en el campo de la genética.

- **Gen.** Es la unidad básica de información en un cromosoma. En una cadena de bits que representa una solución, cada bit puede considerarse un gen. En el caso que se trata en este PFG, el menú semanal, cada tipo de comida se podría considerar un gen. Por ejemplo, un gen sería bebida, plato principal o postre.
- **Alelo.** Es la forma específica o valor que puede tomar un gen. Siguiendo el ejemplo de bebida, posibles alelos serían agua, té o cerveza.
- **Cromosoma.** Es una colección de genes y representa una solución completa al problema de optimización. El cromosoma es la cadena de bits completa. En nuestro caso, la lista completa de alimentos seleccionados para una semana es el cromosoma.
- Fenotipo. Es la manifestación real de la solución codificada. En el PFG sería cómo se preparan y sirven estos alimentos en la realidad.

En la figura 3.2 se expone la estructura de un cromosoma con el ejemplo del menú semanal de comida.

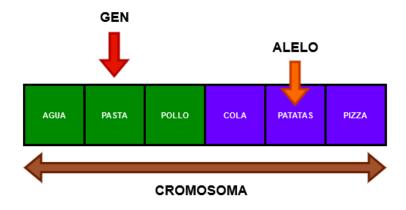


Figura 3.2. Estructura de un cromosoma.

3.2. Evaluación

La generación de una población inicial implica crear un conjunto de soluciones candidatas. Generalmente, los individuos son seleccionados aleatoriamente dentro de los límites definidos en cada problema, asegurando que todas las áreas del espacio de búsqueda puedan ser exploradas. También existen métodos alternativos de generación que aplican ciertas restricciones para formar soluciones iniciales más prometedoras. Tomando de ejemplo el menú semanal, el espacio de búsqueda comprendería la base de datos en la que aparecen todos los alimentos con sus respectivas calorias y macronutrientes.

### 3.2. Evaluación

Se mide lo bueno que es un individuo para nuestros propósitos (la calidad del individuo). Lo más importante es definir una función de fitness correcta y representativa del problema a evaluar. Si se trata de un problema multiobjetivo, se tendrá que diseñar una función de fitness diferente por cada objetivo.

Según vayan pasando generaciones, la población inicial irá evolucionando hacia poblaciones candidatas que presentarán una mejor aptitud. Si la población ha alcanzado el objetivo, la condición de parada se activará, convirtiendo el conjunto de individuos candidatos en la población solución. En caso de no alcanzarlo, seguirá evolucionando hacia otra población candidata distinta.

Para diseñar una función de fitness acorde al problema y al objetivo (u objetivos) se pueden seguir los siguientes pasos:

- 1. Determinar si el objetivo es maximizar o minimizar un valor específico. En el caso que se trata en este proyecto, se busca que la diferencia entre las calorias necesarias y la suma de las calorias de los alimentos sea la menor posible, por lo que se trataría de un problema de minimización.
- 2. Identificar las restricciones. En el ejemplo, una podría ser los límites superior e inferior de calorías permitidas.
- 3. Definir la función matemática que represente el objetivo del problema. Si existen múltiples objetivos, se pueden ponderar.
- 4. Incorporar las penalizaciones para soluciones que no cumplan con las restricciones.

3.2. Evaluación

Por lo tanto, la mejor solución debería recibir la mejor evaluación. Esto ayudará al algoritmo a entender el camino a seguir, ya que penalizará las soluciones infactibles.

#### 3.2.1. Restricciones

Uno de los apartados a destacar cuando se diseña la función de calidad es entender cómo se manejan las restricciones para asegurar que las soluciones generadas sean viables y de alta calidad.

Existen tres tipos de restricciones que limitan el espacio de soluciones válidas. Sea x las variables del problema:

- Restricciones de igualdad.  $h_i(x) = 0$
- Restricciones de desigualdad.  $g_k(x) \ge 0$  ó  $g_k(x) \le 0$
- Restricciones de contorno o de caja (box-constraints).  $x^L \le x \le x^U$

Estos tipos de restricciones se pueden tratar mediante distintos métodos.

- Funciones de penalización. Se penalizan las soluciones infactibles. Estas penalizaciones representan cómo de lejos está la solución de la región factible.
- **Separatista.** Dividen el problema en subproblemas más manejables que se pueden resolver por separado y luego combinar.
- Preservación de la factibilidad. Se usan representaciones especiales (codificación binaria, entera, real, etc.) que simplifican la forma del espacio de búsqueda y operadores que mantienen la factibilidad de las soluciones, lo que elimina las soluciones inválidas en cada etapa del algoritmo.
- **Reparación.** Se modifican las soluciones candidatas infactibles de forma que no violen las restricciones.

#### 3.2.2. Condición de parada

Como se ha mencionado, tras la evaluación de calidad la condición de parada determina cuándo el algoritmo debe detenerse y presentar la solución. Si bien existen una gran variedad de condiciones de parada, las más comunes son:

- Aptitud del individuo. El algoritmo se detiene si se alcanza una solución, ya sea alcanzando el valor objetivo o estando dentro de unos umbrales predefinidos.
- Número máximo de generaciones. Se para si la población ya ha evolucionado un número de veces predefinido.
- Tiempo máximo de ejecución. Parecida a la anterior, pero con un tiempo máximo.
- Convergencia. Se puede detener la ejecución si la población muestra poca o ninguna mejora en la aptitud durante un número determinado de generaciones consecutivas.

## 3.3. Operadores

Son los procesos que se aplican a poblaciones de individuos para desarrollar generaciones futuras. Sirven para la exploración del espacio de soluciones y para la mejora de las poblaciones a través de las generaciones. Se busca que el conjunto de individuos presente una alta diversidad, ya que si es baja se corre el riego de caer en mínimos locales, lo que no permitiría explorar el espacio de soluciones ampliamente. En la figura 3.3 aparecen los operadores básicos en los algoritmos evolutivos, que son los que se van a explicar en más detalle.

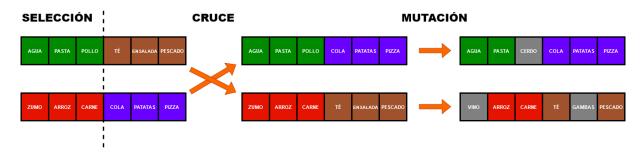


Figura 3.3. Operadores.

#### 3.3.1. Selección

El primero de los operadores básicos. Elige un individuo para reproducirlo. Comúnmente, se busca un equilibrio entre la diversidad y la convegencia. El algoritmo debe explorar el espacio de búsqueda en amplitud, ya que podría caer en mínimos locales si la población es muy parecida entre sí, lo que no permitiría encontrar buenas soluciones. Pero también es de interés que las poblaciones candidatas vayan evolucionando hacia una mayor aptitud, reconduciendo el algoritmo por zonas del espacio de búsqueda más probables de encontrar una solución.

Si bien se puede seleccionar de manera equiprobable, existen métodos basados en la aptitud del individuo que buscan respetar este equilibrio. Estos son algunos:

- Método estándar (rueda de la fortuna). Asigna a cada individuo una probabilidad proporcional a su fitness, por lo que los más aptos tienen mayor probabilidad de reproducirse. Existe un derivado de este método en el que se normalizan las probabilidades, lo que favorece aún más a los adaptados.
- **Método del rango.** Se ordenan los individuos según su aptitud, y la probabilidad de selección se asigna según este ránking.
- Selección por torneo. Se escoge aleatoriamente un subconjunto de individuos de la población, y el mejor de este es elegido.

#### Selección ambiental (Remplazo)

A parte de los métodos basados en la aptitud, se puede usar usa la técnica de la selección ambiental para seleccionar aquellos individuos que puedan influir significativamente en la eficacia del algoritmo genético. Existen varios métodos de selección ambiental:

- Reemplazo generacional completo. Método tradicional en la selección y explicado en el apartado 3.3.1. Toda la población actual es reemplazada por una nueva. Promueve altas tasas de diversidad, pero conlleva una aptitud de los individuos irregular.
- Reemplazo generacional parcial (Steady-State). Solo una parte de la población es reemplazada en cada generación. Se promueve una mejora constante en la calidad ya que las soluciones con un fitness elevado pueden mantenerse.
- Reemplazo elitista. Escoge los mejores individuos de una generación para que se mantengan en la siguiente. Esto ayuda a que la calidad del mejor individuo de una generación sea siempre igual o superior a su equivalente de la generación anterior, consiguiendo una progresión constante hacia la solución óptima, es decir, una mayor convergencia.
- **Selección**  $(\mu, \lambda)$ . De  $\mu$  padres se generan  $\lambda$  hijos. Los  $\mu$  mejores de los  $\lambda$  hijos forman la siguiente generación de padres.
- **Selección**  $(\mu + \lambda)$ . Igual que en la anterior, de  $\mu$  padres se generan  $\lambda$  hijos, pero en este caso los  $\mu$  mejores de la combinación entre los  $\mu$  padres y los  $\lambda$  hijos forman la siguiente generación.

#### 3.3.2. Cruce

Se combina el material genético de dos individuos, padres, para producir descendencia, hijos. Es decir, se utiliza para intercambiar características de dos soluciones parentales con el objetivo de generar nuevas soluciones. Se aplica con probabilidad  $P_c$ . Al promover la mezcla de genes, ayuda a aumentar la diversidad y la convergencia, debido a la generación de nuevos descendientes con características deseables. Se usan distintos métodos:

• Cruce de un punto. Como se muestra en la figura 3.4, se selecciona aleatoriamente una posición en el cromosoma. Todos los genes que están antes de este punto se intercambian con los que están después en la otra cadena, y viceversa, para producir dos nuevos descendientes.

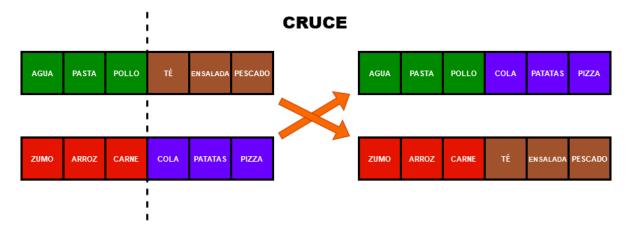


Figura 3.4. Cruce de un punto.

- Cruce de dos puntos. Similar al cruce de un punto, pero se seleccionan dos puntos de corte. Las cadenas de genes entre estos dos puntos se intercambian entre los dos padres.
- Cruce uniforme. Cada gen tiene una probabilidad igual de ser elegido de uno de los dos padres.

#### 3.3.3. Mutación

Último de los operadores básicos. Se recorre toda la cadena, mutando cada gen con probabilidad  $P_m$ , es decir, eligiendo un nuevo valor mediante una elección equiprobable sobre el alfabeto. La mutación aumenta la diversidad y ayuda a no caer en mínimos locales. Algunos de los métodos usados son:

■ Mutación uniforme. Cada gen puede cambiar a otro valor con probabilidad  $P_m$ , como se expone en la figura 3.5.

**MUTACIÓN** 

# PASTA POLLO COLA PATATAS PIZZA AGUA PASTA CERDO COLA PATATAS PIZZA

Figura 3.5. Mutación uniforme.

■ Mutación por intercambio. Dos genes aleatorios en el cromosoma son seleccionados y sus posiciones se intercambian.

## 4.1. Base de datos

La base de datos de alimentos y platos utilizada para el desarrollo del algoritmo es la perteneciente al gobierno de Reino Unido [36]. El proyecto proporciona información detallada sobre la composición nutricional de los alimentos consumidos comúnmente en Reino Unido.

La base de datos "Composition of foods integrated dataset (CoFID)" se originó a partir del trabajo de Widdowson y McCance en la década de 1940. Ha sido periódicamente actualizada desde entonces para reflejar los últimos cambios en la dieta y las nuevas investigaciones científicas. La última actualización se publicó en marzo de 2021.

La elección de esta base de datos se basa en la gran cantidad de datos nutricionales se aportan sobre los alimentos, junto con una amplia documentación que explica en profundidad cada uno de los detalles necesarios para entender el trabajo. A todo esto hay que sumar la facilidad para la exportación y trabajo de la misma.

CoFID incluye más de 2500 alimentos, cada uno categorizado según el tipo producto que sea. Cada alimento incorpora información sobre una amplia gama de nutrientes, como vitaminas, minerales, azúcares y otros componentes, además de los macronutrientes (carbohidratos, proteínas y grasas), que van a ser los principalmente usados en este PFG.

En la página web se pueden encontrar principalmente dos archivos. Una guía de usuario en formato PDF que explica en detalle en qué consiste la base de datos y la información que se va a encontrar en ella. El segundo archivo trata de ua hoja de cáculo de Microsoft Excel donde se encuentra todos la información nutricional relativa a 2.887 alimentos. Existen diferentes casos en que se presenta un mismo alimento con diversos tipos de cocción o de

4.2. Nutrición

aliño, lo que cambia los valores nutricionales, razón por la que se considera como dos alimentos o platos distintos.

### 4.2. Nutrición

Se ha hecho uso de distintas fuentes para los valores nutricionales utilizados para el cáculo de los objetivos, restricciones y límites que se representan en el problema.

En el algoritmo se busca crear un menú que equilibrado y personalizado para el usuario, con el propósito de mantener el peso y la masa corporal del individuo.

El primer objetivo es el calórico. Para ello es necesario cuántas kilocalorías diarias necesita el usuario. La Tasa Metabólica Basal (TMB) es la cantidad de energía que el cuerpo necesita para mantener las funciones vitales en reposo. Existen diversas fórmulas de calcularla, pero en este proyecto se hace uso de una las más extendidas actualmente, la fórmula de Harris-Benedict revisada por Mifflin et al. [37]. Esta fórmula toma en cuenta el peso, la altura, la edad y el sexo para estimarla en kilocalorías.

#### Fórmula

Para hombres:

```
\mathsf{TMB} = (10 \times \mathsf{peso} \; \mathsf{en} \; \mathsf{kg}) + (6.25 \times \mathsf{altura} \; \mathsf{en} \; \mathsf{cm}) - (5 \times \mathsf{edad} \; \mathsf{en} \; \mathsf{años}) + 5
```

■ Para mujeres:

```
TMB = (10 \times \text{peso en kg}) + (6.25 \times \text{altura en cm}) - (5 \times \text{edad en años}) - 161
```

Como se explica en la definición de la TMB, esta energía es calculada en reposo. Por lo tanto, para alcanzar una aproximación real será necesario multiplicar el resultado por un factor que depende del nivel de actividad física [38].

■ Sedentario (poco o ningún ejercicio): TMB × 1.2

4.2. Nutrición

 $\blacksquare$  Actividad ligera (ejercicio ligero o deportes 1-3 días a la semana): TMB  $\times$  1.375

- $\blacksquare$  Actividad moderada (ejercicio moderado o deportes 3-5 días a la semana): TMB  $\times$  1.55
- ullet Actividad alta (ejercicio intenso o deportes 6-7 días a la semana): TMB imes 1.725
- Actividad muy alta (ejercicio muy intenso, trabajos físicos o entrenamiento dos veces al día): TMB × 1.9

Por ejemplo, para un hombre de 23 años que pesara 75 kg y midiera 175 cm su TMB sería de 1734 kilocalorías por dia. Pero como también realiza una actividad moderada, habría que multiplicar por 1.55, dando un total de 2687.31 kcal/dia.

- [1] J. McCarthy, M. Minsky, N. Rochester y C. Shannon, «Dartmouth Summer Research Project on Artificial Intelligence,» Dartmouth College, Hanover, New Hampshire, USA, 1956. dirección: https://home.dartmouth.edu/about/artificial-intelligence-ai-coined-dartmouth.
- [2] A. M. TURING, «I.—COMPUTING MACHINERY AND INTELLIGENCE,» *Mind*, vol. LIX, n.º 236, págs. 433-460, oct. de 1950, ISSN: 0026-4423. DOI: 10. 1093/mind/LIX.236.433. eprint: https://academic.oup.com/mind/article-pdf/LIX/236/433/30123314/lix-236-433.pdf. dirección: https://doi.org/10.1093/mind/LIX.236.433.
- [3] Todoist AI Assistant Integrations, https://todoist.com/es/integrations/apps/ai-assistant, Último acceso: 23 de mayo de 2024.
- [4] Google, Behind the scenes with the new Nest Thermostat, Último acceso: 23 de mayo de 2024, 2020. dirección: https://blog.google/products/google-nest/behind-scenes-new-nest-thermostat/.
- [5] O. N. de Tecnología y Sociedad (ONTSI), Uso de inteligencia artificial y big data en las empresas españolas, 2023. dirección: https://www.ontsi.es/es/publicaciones/uso-de-inteligencia-artificial-y-big-data-en-las-empresas-espanolas.
- [6] Philips, Bringing Confident Diagnosis to More Patients at Low Cost: Philips Launches New AI-enabled CT 5300 at ECR 2024, Último acceso: 23 de mayo de 2024, 2024. dirección: https://www.philips.com/a-w/about/news/archive/standard/news/press/2024/bringing-confident-diagnosis-to-more-patients-at-low-cost-philips-launches-new-ai-enabled-ct-5300-at-ecr2024.html.
- [7] Y. Majeed, L. Fu y L. He, «Editorial: Artificial intelligence-of-things (AIoT) in precision agriculture,» Frontiers in Plant Science, vol. 15, 2024, ISSN: 1664-462X. DOI: 10.3389/fpls.2024.1369791. dirección: https://www.frontiersin.org/journals/plant-science/articles/10.3389/fpls.2024.1369791.

[8] E. Gutiérrez-González, M. García-Solano, R. Pastor-Barriuso et al., «Sociogeographical disparities of obesity and excess weight in adults in Spain: insights from the ENE-COVID study,» Frontiers in Public Health, vol. 11, 2023, ISSN: 2296-2565. DOI: 10.3389/fpubh.2023.1195249. dirección: https://www.frontiersin.org/articles/10.3389/fpubh.2023.1195249.

- [9] Naciones Unidas, Objetivos de Desarrollo Sostenible. dirección: https: //www.un.org/sustainabledevelopment/es/objetivos-de-desarrollosostenible/.
- [10] McKinsey & Company, «The state of AI in 2022—and a half decade in review,» McKinsey Global Institute, 2022. dirección: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2022-and-a-half-decade-in-review.
- [11] S. U. Human-Centered Al Institute, «Al Index Report 2024,» Stanford University, 2024. dirección: https://aiindex.stanford.edu/.
- [12] McKinsey & Company, «The state of AI in 2023: Generative AIs breakout year,» McKinsey Global Institute, 2023. dirección: https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai-in-2023-generative-ais-breakout-year.
- [13] DeepMind, AlphaFold: Using Al for scientific discovery, Accedido: 23 de mayo de 2024, 2024. dirección: https://deepmind.com/research/ case-studies/alphafold.
- [14] V. Ivanov, «A review of fuzzy methods in automotive engineering applications,» European Transport Research Review, vol. 7, pág. 29, 2015. DOI: 10.1007/s12544-015-0179-z. dirección: https://doi.org/10.1007/s12544-015-0179-z.
- [15] J. Dong, «Recent Advances in Swarm Intelligence Algorithms and Their Applications,» *Mathematics*, vol. 11, n.º 12, pág. 2624, 2023. DOI: 10.3390/math11122624. dirección: https://www.mdpi.com/journal/mathematics/special\_issues/Y758SX8ZQC.
- [16] L. J. Fogel, A. J. Owens y M. J. Walsh, *Artificial Intelligence through Simulated Evolution*. New York: John Wiley & Sons, Inc., 1966.
- [17] I. Rechenberg, Evolution Strategy: Optimization of Technical Systems According to the Principles of Biological Evolution. Stuttgart, Germany: Frommann-Holzboog, 1973.

[18] J. H. Holland, Adaptation in Natural and Artificial Systems. Ann Arbor, MI, USA: University of Michigan Press, 1975.

- [19] J. R. Koza, Genetic Programming: On the Programming of Computers by Means of Natural Selection. Cambridge, MA, USA: MIT Press, 1992.
- [20] J. R. Koza, Genetic Programming II: Automatic Discovery of Reusable Programs. Cambridge, MA, USA: MIT Press, 1994.
- [21] J. F. Miller y P. Thomson, «Cartesian Genetic Programming,» en Genetic Programming, R. Poli, W. Banzhaf, W. B. Langdon, J. Miller, P. Nordin y T. C. Fogarty, eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2000, págs. 121-132, ISBN: 978-3-540-46239-2. DOI: 10.1007/978-3-540-46239-2\_9.
- [22] P. Moscato y C. Cotta, «A Gentle Introduction to Memetic Algorithms,» en *Handbook of Metaheuristics*, F. Glover y G. A. Kochenberger, eds., Springer, 2003, págs. 105-144. DOI: 10.1007/978-1-4419-1665-5\_5. dirección: https://link.springer.com/chapter/10.1007/978-1-4419-1665-5\_5.
- [23] NASA Jet Propulsion Laboratory, *Space Technology 5*, Accedido: 23 de mayo de 2024, 2006. dirección: https://www.jpl.nasa.gov/nmp/st5.
- [24] J. D. Lohn, G. S. Hornby y D. S. Linden, An Evolved Antenna for Deployment on NASA's Space Technology 5 Mission. NASA Technical Reports Server (NTRS), 2004, Accedido: 23 de mayo de 2024. dirección: https://ntrs.nasa.gov/citations/20040152147.
- [25] R. Abraham, M. E. Samad, A. M. Bakhach et al., «Forecasting a Stock Trend Using Genetic Algorithm and Random Forest,» *Journal of Risk and Financial Management*, vol. 15, n.º 5, 2022, ISSN: 1911-8074. DOI: 10. 3390/jrfm15050188. dirección: https://www.mdpi.com/1911-8074/15/5/188.
- [26] C. Notredame y D. G. Higgins, «SAGA: Sequence Alignment by Genetic Algorithm,» *Nucleic Acids Research*, vol. 24, n.º 8, págs. 1515-1524, 1996, ISSN: 0305-1048. DOI: 10.1093/nar/24.8.1515. dirección: https://doi.org/10.1093/nar/24.8.1515.
- [27] G. J. Stigler, «The Cost of Subsistence,» *Journal of Farm Economics*, vol. 27, n.° 2, págs. 303-314, 1945.

[28] S. Kirkpatrick, C. D. Gelatt y M. P. Vecchi, «Optimization by Simulated Annealing,» *Science*, vol. 220, n.º 4598, págs. 671-680, 1983. DOI: 10.1126/science.220.4598.671. dirección: https://science.sciencemag.org/content/220/4598/671.

- [29] D. E. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning. Reading, MA, USA: Addison-Wesley, 1989.
- [30] A. Kahraman y H. Seven, «Healthy Daily Meal Planner,» jun. de 2005, págs. 390-393. doi: 10.1145/1102256.1102345.
- [31] E. Kaldirim y Z. Köse, «Application of a Multi-objective Genetic Algorithm to the Modified Diet Problem,» en *Genetic and Evolutionary Computation Congress (GECCO), Undergraduate Student Workshop*, Seattle, USA, 2006.
- [32] T. Kashima, S. Matsumoto y H. Ishii, «Evaluation of Menu Planning Capability Based on Multi-dimensional 0/1 Knapsack Problem of Nutritional Management System,» *IAENG International Journal of Applied Mathematics*, vol. 39, n.º 3, págs. 163-170, 2009. dirección: https://www.iaeng.org/IJAM/issues\_v39/issue\_3/IJAM\_39\_3\_04.pdf.
- [33] P. Heinonen y E. Juuso, «Development of a Genetic Algorithms Optimization Algorithm for a Nutritional Guidance Application,» en *Proceedings of The 9th EUROSIM Congress on Modelling and Simulation, EUROSIM 2016, The 57th SIMS Conference on Simulation and Modelling SIMS 2016*, E. D. E. Juuso y K. Leiviskä, eds., ép. Linköping Electronic Conference Proceedings, Linköping University Electronic Press, 2016, págs. 755-761. DOI: 10.3384/ecp1714255.
- [34] S. KILICARSLAN, M. CELIK y afak SAHIN, «Hybrid models based on genetic algorithm and deep learning algorithms for nutritional Anemia disease classification,» *Biomedical Signal Processing and Control*, vol. 63, pág. 102231, 2021, ISSN: 1746-8094. DOI: 10.1016/j.bspc.2020.102231. dirección: https://www.sciencedirect.com/science/article/pii/S174680942030361X.
- [35] J. B. Cole y R. Gabbianelli, «Recent Advances in Nutrigenomics: Making Strides Towards Precision Nutrition,» *Frontiers in Genetics*, vol. 13, 2022. DOI: 10.3389/fgene.2022.997266. dirección: https://www.frontiersin.org/articles/10.3389/fgene.2022.997266/full.

[36] P. H. England, McCance and Widdowson's Composition of Foods Integrated Dataset (CoFID), Accessed: 2024-06-04, 2021. dirección: https://www.gov.uk/government/publications/composition-of-foods-integrated-dataset-cofid.

- [37] M. Mifflin, S. St Jeor, L. Hill, B. Scott, S. Daugherty e Y. Koh, «A new predictive equation for resting energy expenditure in healthy individuals,» American Journal of Clinical Nutrition, vol. 51, n.º 2, págs. 241-247, 1990.
- [38] L. Mahan y J. Raymond, *Krause's Food & the Nutrition Care Process*, 14th. Elsevier Health Sciences, 2016, págs. 370-372.

# Índice de términos

