

**Proyecto Fin de Máster**

Máster en Big Data, Inteligencia Artificial y Data Science

**HealthBite:**

Aplicación con Visión Artificial y Lenguaje Natural para la nutrición personalizada a partir de ingredientes y estado del usuario

Septiembre 2025

**Autor:** Óscar Xu Zhou

Guión:

* Resumen
* 1) Introducción
  + 1.1 Contexto y justificación del proyecto
  + 1.2 Objetivos del proyecto
* 2) Flujo y estructura de la aplicación inteligente
* 3) Revisión del estado del arte y análisis del mercado de inteligencia artificial en la nutrición
* 4) Análisis exploratorio de datos (EDA)
* 5) Creación de modelo YOLO
  + 5.1) Construcción de datasets
  + 5.2) Entrenamiento del modelo
  + 5.3) Evaluación del modelo y ejemplos de predicción
* 6) Creación de modelo NLP
  + 6.1) Construcción de datasets
  + 6.2) Entrenamiento del modelo
  + 6.3) Evaluación del modelo y ejemplos de predicción
* 7) Algoritmo de recomendación y RAG LLM como juez
  + 7.1) Combinación de los modelos YOLO + NLP
  + 7.2) Traducción del dataset utilizado de recetas
  + 7.3) Algoritmo de recomendación por reglas y puntuación
  + 7.4) RAG LLM como juez final y prompt engineering
* 8) Conclusiones y resultados
  + 8.1) Limitaciones actuales
  + 8.2) Trabajo futuro

# **Introducción**

## Contexto y justificación del proyecto

El desperdicio alimenticio es uno de los grandes problemas que afronta la sociedad actual. Se estima que en la Unión Europea los hogares generan un 54% del desperdicio total, lo que equivale a aproximadamente 72 kilos de desperdicio por persona. (Eurostat, 2024). Este fenómeno no solo implica pérdidas económicas, sino también un fuerte impacto ambiental, dado que se desaprovecha toda la energía invertida en la producción y la descomposición de los alimentos en vertederos genera gases de efecto invernadero.

Diversos estudios señalan que el uso cotidiano de la nevera está relacionado con desperdicio alimenticio. Un estudio sociológico descubrió que, en muchos casos, los productos quedan olvidados o quedan ocultos por la posición, causando que caduquen sin ser consumidos (Heidenstrøm & Hebrok, 2020) Por otra parte, otro estudio reveló que el uso de la nevera es asociado a un aumento del 24.35% de desperdicio alimenticio (Longqiang Zhao, Min, Wang, & Yu, 2024). Estos datos evidencian la necesidad de soluciones que ayuden a la optimización del consumo de los alimentos almacenados en la nevera.

Adicionalmente al contexto de desperdicio, la nutrición personalizada se ha consolidado como un campo emergente alcanzando grandes mejoras sobre todo con los últimos avances en inteligencia artificial. Aunque existen aplicaciones móviles que ofrecen recomendaciones dietéticas, una gran mayoría se han enfocado a proporcionar información tras un escaneo códigos de barras o proponer recetas aleatorias o genéricas, mayoritariamente orientadas a dietas saludables o fitness. Estas soluciones rara vez tienen en cuenta los ingredientes realmente disponibles en el hogar ni cubren las necesidades dietéticas específicas del usuario.

Este proyecto propone el desarrollo de HealthBite, una aplicación inteligente basada en visión artificial, algoritmos de lenguaje natural y modelos de lenguaje de gran escala (LLM). Su objetivo es recomendar recetas que, por un lado, prioricen el uso de los ingredientes disponibles en la nevera para reducir el desperdicio alimenticio, y por otro, contribuyan a mejorar la salud dietética del usuario al cubrir carencias nutricionales potenciales inferidas a partir de descripciones de estado anímico/físico. De este modo, la propuesta combina sostenibilidad, innovación tecnológica y bienestar personal en una única solución escalable.

## Objetivos del proyecto

**Objetivo principal:**

Desarrollar una aplicación inteligente capaz de recomendar recetas considerando tanto los ingredientes disponibles en la nevera del usuario como sus posibles deficiencias nutricionales, inferidas a partir de descripciones de su estado anímico o físico. Adicionalmente, será capaz de explicar las razones de cada recomendación.

**Objetivos específicos:**

* Realizar un análisis exploratorio de datos (EDA) sobre resultados de encuestas de consumición alimentario en hogares de España, con el fin de obtener mejor entendimiento sobre patrones y productos más consumidos
* Creación de un modelo de visión artificial (YOLO) para identificar automáticamente los ingredientes disponibles en la nevera a partir de imágenes
* Entrenar un modelo de procesamiento del lenguaje natural (NLP) capaz de clasificar descripciones del estado anímico/físico del usuario en síntomas y asociarlas con posibles deficiencias nutricionales
* Implementar un sistema de traducción automática a gran escala mediante modelos NLP para adaptar el dataset de recetas del inglés al español
* Construir un sistema de puntuación que permita ordenar las recetas en función de su afinidad al contexto específico del usuario (ingredientes disponibles y necesidades nutricionales).
* Desarrollar e integrar un LLM con RAG, ajustado mediante técnicas de prompt engineering, que actúe como juez final para validar que las condiciones de recomendación se cumplen y genere una explicación clara para el usuario.
* Integrar todos los modelos en una aplicación interactiva con Streamlit, lista para ser utilizada en un entorno real por usuarios finales.

# **Flujo y estructura de la aplicación inteligente**