COMILLAS
UNIVERSIDAD PONTIFICIA

ICAI

GRADO EN INGENIERÍA MATEMÁTICA E INTELIGENCIA ARTIFICIAL

TRABAJO FIN DE GRADO

SMART NUTRITION AND COOKING. APPLICATIONS TO SPECIFIC DIETARY RESTRICTIONS.

Author: Elena Ardura Carnicero

Director: Pablo Dueñas Martínez

Co-Director: Rocío Teresa Jiménez de la Peña Armada

Madrid, junio de 2025

Declaro, bajo mi responsabilidad, que el Proyecto presentado con el título

“Smart nutrition and cooking. Applications to specific dietary restrictions.”

en la ETS de Ingeniería - ICAI de la Universidad Pontificia Comillas en el

curso académico 2024/25 es de mi autoría, original e inédito y

no ha sido presentado con anterioridad a otros efectos.

El Proyecto no es plagio de otro, ni total ni parcialmente y la información que ha sido

tomada de otros documentos está debidamente referenciada.

Fdo.: Elena Ardura Carnicero Fecha: 11/06/2025

Autorizada la entrega del proyecto

EL DIRECTOR DEL PROYECTO

Fdo.: Pablo Dueñas Martínez Fecha: 11/06/2025

ACKNOWLEDGEMENTS

This Final Degree Project would not have been possible without the support, guidance, and company of many people who have been by my side throughout the entire process. To all of them, thank you.

I would like to begin by especially thanking my advisor, Pablo Dueñas, for his commitment and approachability throughout the development of this project. Thank you as well for being an exceptional tutor, always looking out for each of us. It has truly been a privilege to count on your support.

I also extend my thanks to Rocío Jiménez de la Peña, co-director of the project, for her expert insight and for always being willing to help with anything we asked for.

I want to thank my family, especially my parents, my siblings, and my aunt Chus, for their constant support, for putting up with me when I was stressed, and for always giving me the confidence I needed.

Finally, I would like to thank all my classmates for the support and the environment we created together, making our graduating class truly unforgettable. A very special thank you to my friends, to the Pochunas and the Spicy, for giving me four incredible years, for everything you've taught me, and for helping me become a better version of myself.

SMART NUTRITION AND COOKING. APPLICATIONS TO SPECIFIC DIETARY RESTRICTIONS.

Author: Ardura Carnicero, Elena
Director: Dueñas Martínez, Pablo
Co-Director: Rocío Teresa Jiménez de la Peña Armada
Entidad Colaboradora: ICAI – Universidad Pontificia Comillas

Abstract

This project presents a complete system capable of generating weekly ingredient plans that meet official nutritional recommendations for different demographic profiles. Using only age and gender as input, the model optimizes ingredient selection, ensuring macro- and micronutrient balance. To evaluate the culinary viability, daily ingredient groups are clustered to simulate meals, and their similarity with real recipes is assessed using Jaccard similarity. Results confirm that the model creates nutritionally sound combinations, though challenges remain in translating them into realistic recipes. The system runs efficiently on standard hardware, offering a solid foundation for future developments in personalized nutrition.

Keywords: optimization, mixed-integer linear programming, nutrition, dietetic , clustering, recipe similarity.

1. Introduction

A balanced and personalized diet is one of the main pillars of health and well-being. However, designing weekly meal plans that are both nutritionally adequate and appealing is a complex task, especially when considering the variability of nutritional needs across different age and gender groups. This project addresses this challenge by developing a system that automatically generates ingredient-based weekly plans that comply with official nutritional guidelines and assesses their culinary feasibility through clustering and recipe suggestion methods.

2. Project Statement

The objective of this project is to build an end-to-end system that generates personalized ingredient combinations for a full week, ensuring compliance with dietary recommendations for specific population profiles (e.g., age and gender). The system solves a large-scale optimization problem to satisfy both nutrient targets and food group frequencies. To evaluate the gastronomic plausibility of the results, the daily ingredients are grouped into three clusters (representing meals), which are then compared with real recipes using a similarity metric.

The project also explores how parameter choices, such as minimum cluster size and frequency thresholds, affect the balance between nutritional accuracy and culinary realism.

3. Description of the System

The proposed system consists of four main stages. Figure 1 shows the execution pipeline of the system, including inputs and outputs for the first three stages. The final analysis stage is not represented.

Data Preprocessing: Nutritional data and official recommendations from the Spanish Agency for Food Safety and Nutrition and the European Food Safety Authority are loaded and harmonized. Ingredients are filtered, normalized, and simplified to ensure consistency.

Optimization Model: A Pyomo-based mixed-integer linear programming model is built to select the optimal amount of each ingredient per day. The model includes constraints on nutrient intake, food group frequency, and ingredient repetition, and minimizes the deviations from the nutritional recommendations.

Clustering & Recipe Matching: The selected ingredients per day are grouped into three meals using KMeansConstrained. Each cluster is compared to real recipes using the Jaccard similarity index.

Analysis & Evaluation: The results are analyzed in terms of nutritional compliance, ingredient diversity, and similarity to real meals.

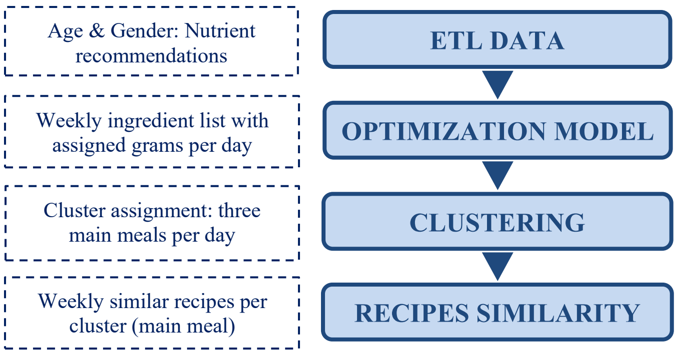

Figure 1: System pipeline

4. Results

The system achieves full feasibility across all constraints for all profiles. The optimization is completed in under 10 seconds, and the additional processing for clustering and recipe matching remains under 3 seconds in total.

From a nutritional perspective, the system successfully meets the targets for approximately 30% (10 out of 33) of nutrients per day, while most deviations fall below minimum thresholds (Figure 2). This indicates that although the optimizer balances the solution across the week, stricter targets remain challenging due to the concentration of some nutrients in specific foods, which is usual in nutrient intake.

The clustering analysis shows that using a minimum cluster size of 4 improves the balance between meals, even though the similarity with real recipes slightly decreases. The trade-off between nutritional compliance and culinary acceptance is clearly visible in the analysis of frequency constraints.

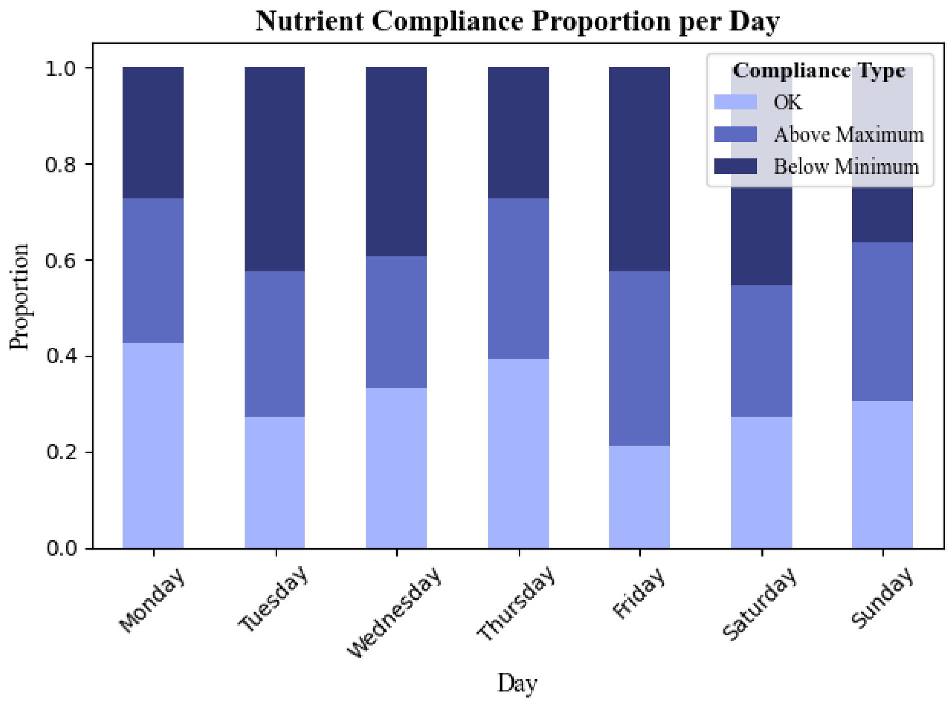

Figure 2: Daily nutrient compliance by deviation type

5. Conclusions

This project demonstrates the feasibility of generating personalized ingredient combinations that are nutritionally sound and reasonably structured for meal preparation. While full menu generation with concrete recipes is not yet implemented, the foundation for such a system has been successfully built and validated.

Future work includes improving data sources for both ingredients and recipes, incorporating intelligent learning systems for iterative refinement, and developing a user-friendly visual interface to make the system accessible to the general public.

6. References

[1] Agencia Española de Seguridad Alimentaria y Nutrición (AESAN), *Informe del Comité Científico: Recomendaciones dietéticas para la población española*, AESAN, Madrid, 2022.

[2] European Food Safety Authority (EFSA), *Dietary Reference Values for Nutrients: Summary Report*, EFSA, Parma, Italy, 2019.

[3] Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P., *Tabela da Composição de Alimentos – v6.0*, 2023. [Online]. Available: https://portfir-insa.min-saude.pt/

[4] Spoonacular API, *Spoonacular – The Food API*, [Online]. Available: https://spoonacular.com/food-api

NUTRICIÓN INTELIGENTE Y COCINA. APLICACIONES A RESTRICCIONES DIETÉTICAS ESPECÍFICAS.

Autor: Ardura Carnicero, Elena
Director: Dueñas Martínez, Pablo
Co-Director: Rocío Teresa Jiménez de la Peña Armada
Entidad Colaboradora: ICAI – Universidad Pontificia Comillas

Abstract

Este proyecto presenta un sistema completo capaz de generar planes semanales de ingredientes que cumplen con las recomendaciones nutricionales oficiales para distintos perfiles demográficos. Usando únicamente la edad y el sexo como entrada, el modelo optimiza la selección de ingredientes, garantizando el equilibrio de macro y micronutrientes. Para evaluar la viabilidad culinaria, los ingredientes diarios se agrupan mediante técnicas de clustering para simular comidas, y se evalúa su similitud con recetas reales utilizando el índice de Jaccard. Los resultados confirman que el modelo genera combinaciones nutricionalmente válidas, aunque persisten desafíos al traducirlas en recetas realistas. El sistema se ejecuta eficientemente en hardware estándar, ofreciendo una base sólida para futuros desarrollos en nutrición personalizada.

Keywords: optimización, programación lineal entera mixta, nutrición, dietética, clustering, similitud de recetas.

1. Introducción

Una dieta equilibrada y personalizada es uno de los pilares fundamentales de la salud y el bienestar. Sin embargo, diseñar menús semanales que sean tanto nutricionalmente adecuados como atractivos resulta una tarea compleja, especialmente al considerar la variabilidad de necesidades nutricionales entre distintos grupos de edad y sexo. Este proyecto aborda dicho desafío desarrollando un sistema que genera automáticamente planes semanales basados en ingredientes, que cumplen con las guías nutricionales oficiales y evalúa su viabilidad culinaria mediante métodos de agrupamiento y sugerencia de recetas.

2. Definición del proyecto

El objetivo de este proyecto es construir un sistema integral que genere combinaciones personalizadas de ingredientes para una semana completa, garantizando el cumplimiento de las recomendaciones dietéticas para perfiles poblacionales específicos (por ejemplo, edad y sexo). El sistema resuelve un problema de optimización a gran escala para satisfacer tanto objetivos nutricionales como frecuencias recomendadas por grupo de alimentos.
Para evaluar la plausibilidad gastronómica de los resultados, los ingredientes diarios se agrupan en tres clústeres (representando comidas), que posteriormente se comparan con recetas reales mediante una métrica de similitud.

El proyecto también explora cómo afectan las elecciones de parámetros - como el tamaño mínimo de clúster o los umbrales de frecuencia - al equilibrio entre precisión nutricional y realismo culinario.

3. Descripción del Sistema

El sistema propuesto consta de cuatro etapas principales. La Figure 3 muestra el flujo de ejecución del sistema, incluyendo las entradas y salidas de las tres primeras etapas. La fase final de análisis no está representada.

Preprocesado de datos: Se cargan y armonizan datos nutricionales y recomendaciones oficiales de la Agencia Española de Seguridad Alimentaria y Nutrición (AESAN) y de la Autoridad Europea de Seguridad Alimentaria (EFSA). Los ingredientes se filtran, normalizan y simplifican para garantizar la coherencia.

Modelo de optimización: Se construye un modelo de programación lineal entera mixta (MILP) basado en Pyomo, que selecciona la cantidad óptima de cada ingrediente por día. El modelo incluye restricciones sobre la ingesta de nutrientes, frecuencia de grupos de alimentos y repetición de ingredientes, y minimiza las desviaciones respecto a las recomendaciones nutricionales.

Clustering y comparación con recetas: Los ingredientes seleccionados por día se agrupan en tres comidas utilizando el algoritmo KMeansConstrained. Cada clúster se compara con recetas reales utilizando el índice de similitud de Jaccard.

Análisis y evaluación: Se analizan los resultados en términos de cumplimiento nutricional, diversidad de ingredientes y similitud con comidas reales.

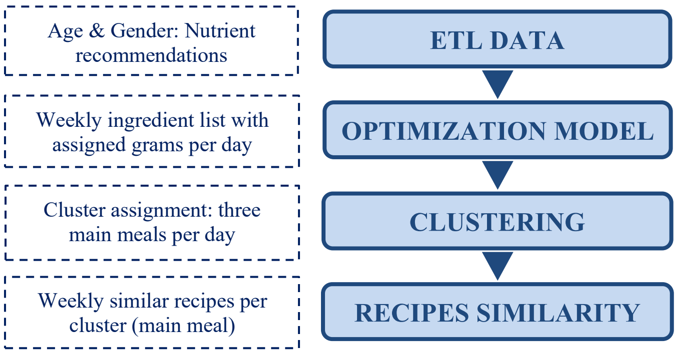

Figure 3: System pipeline

4. Resultados

El sistema alcanza la viabilidad completa respecto a todas las restricciones para todos los perfiles. La optimización se completa en menos de 10 segundos, y el procesamiento adicional para clustering y comparación con recetas permanece por debajo de los 3 segundos en total.

Desde el punto de vista nutricional, el sistema logra cumplir con los objetivos diarios para aproximadamente el 30 % de los nutrientes (10 de 33), mientras que la mayoría de las desviaciones se sitúan por debajo de los umbrales mínimos (Figure 4). Esto indica que, aunque el optimizador equilibra la solución a lo largo de la semana, los objetivos más estrictos siguen siendo difíciles de cumplir debido a la concentración de ciertos nutrientes en alimentos muy específicos, algo habitual en el patrón de ingesta de nutrientes.

El análisis de clustering muestra que usar un tamaño mínimo de clúster de 4 mejora el equilibrio entre comidas, aunque la similitud con recetas reales se reduce ligeramente. El compromiso entre cumplimiento nutricional y aceptabilidad culinaria se hace evidente en el análisis de restricciones de frecuencia.

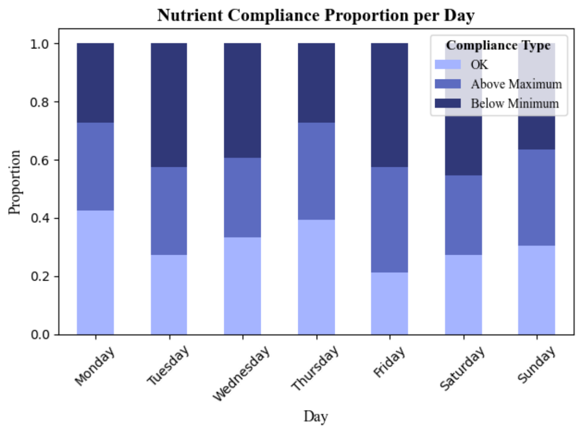

Figure 4: Daily nutrient compliance by deviation type

5. Conclusiones

Este proyecto demuestra la viabilidad de generar combinaciones personalizadas de ingredientes que sean nutricionalmente equilibradas y razonablemente estructuradas para su uso en la preparación de comidas. Aunque aún no se implementa la generación de menús completos con recetas concretas, se ha construido y validado una base sólida para un futuro sistema integral.

Entre las líneas futuras se incluye la mejora de las fuentes de datos tanto para ingredientes como para recetas, la incorporación de sistemas inteligentes de aprendizaje para refinamiento iterativo y el desarrollo de una interfaz visual accesible para el público general.

6. References

[1] Agencia Española de Seguridad Alimentaria y Nutrición (AESAN), *Informe del Comité Científico: Recomendaciones dietéticas para la población española*, AESAN, Madrid, 2022.

[2] European Food Safety Authority (EFSA), *Dietary Reference Values for Nutrients: Summary Report*, EFSA, Parma, Italy, 2019.

[3] Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P., *Tabela da Composição de Alimentos – v6.0*, 2023. [Online]. Available: https://portfir-insa.min-saude.pt/

[4] Spoonacular API, *Spoonacular – The Food API*, [Online]. Available: https://spoonacular.com/food-api

Table of Contents

List of Figures

1. INTRODUCTION

CONTEXT AND MOTIVATION

Over the past decades, the prevalence of chronic diseases and food allergies such as diabetes, hypertension, gluten intolerance, and lactose allergies, have increased at an alarming rate. These conditions are strongly linked to poor dietary quality, characterized by the overconsumption of processed foods and insufficient intake of essential nutrients. Among populations that are increasingly adopting vegetarian or vegan diets, inadequate nutritional planning can lead to deficiencies in key nutrients like iron, vitamin B12, and omega-3 fatty acids.

Despite growing awareness and the existence of nutritional guidelines, many people still struggle to meet basic dietary requirements, especially when additional restrictions are present. This is compounded by a general lack of knowledge about food interactions, which affect the absorption and effectiveness of nutrients such as the inhibitory effect of calcium and tannins on iron absorption.

Technology offers a promising avenue to address these issues, but most existing tools focus on calorie counting and generic advice, lacking the personalization needed to support users with complex nutritional needs. There is a clear opportunity to leverage mathematical optimization and machine learning to create personalized nutritional recommendations that adapt dynamically to individual preferences, restrictions, and external factors like ingredient availability.

This project aims to address these challenges by developing a system that integrates nutrition science, artificial intelligence, and culinary knowledge to help users with dietary restrictions plan balanced meals more easily and effectively, ultimately contributing to improved health outcomes and reduced pressure on healthcare systems.

OBJECTIVES

The main objective of this project is to develop a system that effectively integrates the principles of nutrition, mathematics, and gastronomy to provide personalized recommendations for recipes that meet users' nutritional needs throughout the week. Using mathematical optimization techniques, the system will seek to maximize compliance with nutritional recommendations, considering gender and age**.** To this end, various databases must be collected and processed, including official nutritional compositions, national dietary guidelines, and a wide range of actual recipes, so that they can be integrated into a common model.

PLANNING AND ECONOMIC FEASIBILITY

The development of the project has followed a structured and modular approach, beginning with data preparation and continuing through system implementation and testing. Publicly available nutritional data from official health agencies and online sources was curated and processed to support the development of a functional model.

All components of the project have been developed using Python and open-source tools, supported by an academic license for the optimization solver. The system has been designed to operate efficiently on standard computing resources, allowing all development and testing to be carried out on a personal computer. This ensures a low-cost, accessible, and reproducible solution, suitable for academic use and potential future applications.

STRUCTURE OF THE REPORT

This report is organized into five additional chapters. Chapter 2 presents a review of the current state of the art in the areas of nutrition, gastronomy, medicine, and mathematical modeling. Chapter 3 describes the methodology followed, including the technical approach, algorithmic development, and system design. Chapter 4 details the experimental setup, data processing, and implementation environment. Chapter 5 presents the main results and their analysis, and Chapter 6 concludes the report with a summary of findings and suggestions for future development.

2. STATE OF THE ART

Greater interest in personalized nutrition has encouraged experts from mathematics, gastronomy, medicine, and computer science to come together. Through mathematical modeling, it is now possible to make balanced diets that meet all the complex nutritional standards. Using linear and mixed-integer programming, one can find ideal choices for food that fit into designated nutrient and calorie limits. These techniques have been used to design menus in controlled feeding trials [1] with high efficiency, managing complex constraints such as multiple energy levels and inter-group nutritional variability.

Similarly, machine learning is now also used to look at huge amounts of data on food and health. Projects like Latam Israel apply segmentation algorithms to adapt diets to individual needs [2] and other models detect hidden patterns in dietary habits that relate to health conditions, improving the relevance and effectiveness of personalized plans [3]. This helps healthcare professionals design adaptive recommendations, reducing the risk of chronic diseases.

There has been an obvious effort by many chefs to add nutritional benefits to the dishes they make. In culinary medicine, professionals combine cooking with medical knowledge to create meals that promote health without sacrificing taste. Chefs frequently use substitutions such as olive oil instead of saturated fats, or yogurt instead of cream. Functional ingredients like turmeric, ginger, and kale-known for their anti-inflammatory and antioxidant properties, are also increasingly present in modern recipes, along with polyphenol rich foods such as berries and cocoa, which improve cardiovascular health [4].

At the same time, nutritional apps have appeared, allowing users to keep track of their calories, get advice for different diets and prepare meals. Apps like MyFitnessPal or Yazio are popular but often rely on fixed databases and offer limited flexibility. Many lack scientific validation or fail to adapt to evolving health conditions. Some support specific diets like ketogenic or gluten-free, but do not offer deep integration of nutrient optimization, interaction tracking, or contextual recommendations.

Because of these limitations, it becomes clear that we need optimization systems that also include the flexibility and learning power of artificial intelligence. With these technologies, new solutions can offer users real, varied, and nutritious meal plans. On top of that, following the official health recommendations and real cooking practices helps these systems stay accurate and relevant.

Our aim is at identifying edible meals that match nutritional goals and transfer them into recipes with the help of data sources. Combining the structure of models with our day-to-day habits, we expect to offer support for a healthier life through automation, personalization, and evidence-based meal plans.

3. METHODOLOGY

DATA COLLECTION AND PREPROCESSING

The foundation of the system lies in a structured data preprocessing pipeline that transforms raw nutritional information and recipe data into a format suitable for mathematical modeling. This preprocessing stage comprises several key tasks: translation and simplification of ingredient names, filtering and normalization of nutritional information, and alignment between datasets to ensure consistency.

The initial nutritional database used in the project was obtained from the official Portuguese Food Composition Table (PorFir [5]). Since the original dataset was in Portuguese, ingredient names were translated into English using online tools. In addition, a simplification step was performed to generalize complex ingredient names and facilitate their matching with recipe ingredients later on. Nutrients whose values were below 5% of the mean across all ingredients were set to zero to reduce noise and eliminate negligible contributions.

To define serving size and frequency constraints, official recommendations from the Spanish Agency for Food Safety and Nutrition (AESAN) were consulted, specifically the report issued by its Scientific Committee for the revision and update of national dietary guidelines [6]. These documents provided: (1) the acceptable gram range per serving for each food subgroup, (2) whether serving recommendations should be met on a daily or weekly basis, and (3) the number of servings required per food group. This information was structured to allow integration into the optimization model.

Daily nutrient intake recommendations were derived primarily from merging official reports published by AESAN and the European Food Safety Authority (EFSA) [7][8][9][10]. These include the updated nutritional objectives, and EFSA's dietary reference values for fats and sodium. The recommended ranges were adapted to the user's gender and age group. All nutrient values were normalized to a consistent unit (grams) and rescaled to the [0, 1] interval by dividing by the maximum of each range, enabling stable optimization.

The recipe dataset was constructed using the Spoonacular API [11], which returned approximately 975 recipes. These recipes were processed using AI-assisted tools to extract and clean ingredient lists, transforming them into standardized arrays of simplified ingredients. While the original data had inconsistencies, automated processing helped reduce variability and enhance structure.

To match the ingredients present in the nutritional dataset with those found in recipes, a similarity analysis was carried out based on string distance metrics. Ingredients from recipes were mapped to the closest available nutritional ingredient using a Levenshtein similarity threshold of at least 75%. The Levenshtein distance between two strings measures the number of single-character edits-insertions, deletions, or substitutions-required to transform one string into the other. Formally, for two strings $a = a_1 a_2 \dots a_m$ and $b = b_1 b_2 \dots b_n$, the recursive definition is:

$$\mathrm{Lev}(i,j) = \begin{cases} \max(i,j) & \text{if} \min(i,j) = 0 \\ \min \begin{cases} \mathrm{Lev}(i-1,j)+1 \\ \mathrm{Lev}(i,j-1)+1 \\ \mathrm{Lev}(i-1,j-1) + \delta(a_i,b_j) \end{cases} & \text{otherwise} \end{cases}$$

E. 1: Recursive definition of the Levenshtein distance

where $\delta(a_i, b_j) = 0$ if $a_i = b_j$, and 1 otherwise.

To account for different stings lengths, the normalized Levenshtein similarity is computed as:

$$sim(a,b) = 1 - \frac{Lev(a,b)}{max(|a|,|b|)}$$

E. 2: Normalized Levenshtein similarity between two strings

A similarity of 0.75 or higher was considered acceptable to ensure semantic alignment while allowing for linguistic or typographic variations.

After identifying the matched ingredients, the nutritional database was filtered to retain only those that appeared in the recipe dataset. This ensured that only usable and relevant ingredients were passed to the optimization stage.

Finally, to support the modeling of co-occurrence patterns among ingredients, a co-occurrence matrix was built. This matrix counts how often pairs of ingredients appear together in the same recipe and is later used to guide clustering and recipe reconstruction in the postprocessing stage.

In parallel, an expert-based penalty system was defined to prioritize nutrients that are more critical for human health. Each nutrient was assigned to a penalty coefficient ranging from 1 (least important) to 5 (most important), according to its relative significance in a balanced diet.

OPTIMIZATION MODEL

The core of the system is a mathematical optimization model formulated using the Pyomo modeling language and solved with the Gurobi solver under an academic license. The objective is to determine, for each day of the week, the quantity (in grams) of each ingredient to be consumed such that nutritional recommendations and food group constraints are satisfied, while deviations from the ideal nutritional targets are minimized.

The problem is modeled as a mixed-integer linear program (MILP). The sets used in the formulation are:

- *I*: the set of available ingredients
- *D*: the set of days in the week
- *N*: the set of nutrients
- *G*: the set of food groups
- *SG*: the set of food subgroups

The decision variables are:

- $g_{i,d} \in \mathbb{R}_{\geq 0}$: quantity (in grams) of ingredient *i* consumed on day *d*.
- $k_{i,d} \in \{0,1\}$: binary variable indicating whether ingredient *i* is served on day *d*.
- $u_i \in \{0,1\}$: whether ingredient *i* is used at least once during the week.
- $\delta^+_{n,d}, \delta^-_{n,d} \in \mathbb{R}_{\geq 0}$: positive and negative deviations from the daily nutritional targets for nutrient *n*.

The model also relies on several input parameters:

- $a_{i,n}$: normalized contribution of ingredient *i* to nutrient *n*.
- l_n, u_n: lower and upper bounds for the intake of nutrient *n*, adapted to the user profile.
- ρ_n: penalty coefficient associated with nutrient *n*.
- $group\_ing_{i,f}$: binary parameter indicating if ingredient *i* belong to food group *f*.
- $subgroup\_ing_{i,sg}$: binary parameter indicating if ingredient *i* belongs to food subgroup *sg*.
- $min\_g\_serving_{sg}$, $max\_g\_serving_{sg}$: valid grams limits for serving of subgroup *sg*.
- $min\_serving_f$, $max\_serving_f$: lower and upper recommended servings of food group *f*.
- $daily\_groups_f$, $weekly\_groups_f$: indicators for whether group *f* is constrained daily or weekly.
- $scale\_factor_n$: normalization factor to recover original nutrient units
- $ing\_frequency_i$: percentage of recipes in which ingredient *i* appears.

The objective function aims to minimize the total weighted deviation from nutritional targets:

$$min \sum_{d \in D} \sum_{n \in N} \rho_n (\delta^+_{n,d} + \delta^-_{n,d})$$

E. 3: Objective function for minimizing weighted nutrient deviations

where ρ_n reflects the clinical priority if nutrient *n*.

The model includes the following constraints:

1. Nutrient compliance constraints

These ensure that for each nutrient and day, total intake lies within the recommended range. Slack variables allow deviations, which are penalized:

$$\sum_{i \in I} a_{i,n} \cdot g_{i,d} \leq u_n + \delta^+_{n,d} \quad \forall d \in D, \forall n \in N$$

$$\sum_{i \in I} a_{i,n} \cdot g_{i,d} \geq l_n - \delta^-_{n,d} \quad \forall d \in D, \forall n \in N$$

E. 4: Valid nutrient intake constraints

2. Food group servings constraints

Each food group has a recommended number of servings, either daily or weekly:

- If $f \in G$ is a daily group:

$$\sum_{i \in I} k_{i,d} \cdot group\_ing_{i,f} \leq max\_serving_f \quad \forall d \in D$$

$$\sum_{i \in I} k_{i,d} \cdot group\_ing_{i,f} \geq min\_serving_f \quad \forall d \in D$$

E. 5: Valid daily servings per food group constraints

- If $f \in$ G is a weekly group:

$$\sum_{d \in D} \sum_{i \in I} k_{i,d} \cdot group\_ing_{i,f} \leq max\_serving_f$$

$$\sum_{d \in D} \sum_{i \in I} k_{i,d} \cdot group\_ing_{i,f} \geq min\_serving_f$$

E. 6: Valid weekly servings per food group constraints

3. Valid serving size constraints

If ingredient i belongs to subgroup s, its quantity in grams on day d must respect the minimum and maximum thresholds defined for a valid serving of that subgroup:

$$g_{i,d} \leq k_{i,d} \cdot max\_g\_serving_s \; \forall d \in D, \forall i \in I$$

$$g_{i,d} \geq k_{i,d} \cdot min\_g\_serving_s \; \forall d \in D, \forall i \in I$$

E. 7: Valid serving size constraints

4. Ingredient repetition constraint

To avoid overly repetitive plans and ensure variety, the model enforces a constraint that limits how many days an ingredient can be included in the weekly plan:

$$\sum_{d \in D} k_{i,d} \leq max\_repeat \quad \forall i \in I$$

E. 8: Weekly upper limit on ingredient repetition

where max_repeat is a tunable parameter (e.g. 2) that defines the maximum number of days in which a given ingredient i can appear across the weekly plan.

5. Ingredient usage logic constraint

This constraint ensures consistency between the binary weekly usage variable u_i and the ingredient's actual use throughout the week. It guarantees that if an ingredient is chosen on any day, the corresponding weekly usage variable must be activated, and conversely, if the ingredient is marked as used, it cannot appear in more than max_repeat days:

$$\sum_{d \in D} k_{i,d} \leq max\_repeat \cdot u_i \quad \forall i \in I$$

$$\sum_{d \in D} k_{i,d} \geq u_i \quad \forall i \in I$$

E. 9: Link between binary usage and actual daily use of an ingredient

6. Ingredient popularity constraint

To prioritize the selection of ingredients that are commonly used in real recipes, the model enforces a minimum frequency score:

$$\sum_{i \in I} u_i \cdot ing\_frequency_i \geq \tau$$

E. 10: Minimum cumulative popularity score across selected ingredients

where τ is a tunable threshold (e.g. 50), chosen to guarantee a minimum cumulative popularity score across all selected ingredients.

SOLUTION POST-PROCESSING

Once the optimization model has been solved, the output generated specifies the daily grams of each ingredient that should be consumed. Although this information meets the defined nutritional targets, it is not directly interpretable by an end user, as it is not structured in the form of dishes or meals. For this reason, a post-processing phase was designed with the purpose of converting the lists of ingredients into specific culinary proposals, compatible with actual eating habits.

To organize the daily ingredients into distinct meals, the KMeansConstrained algorithm, an extension of the classic KMeans algorithm, was used. This variant allows restrictions to be imposed on the minimum size of the clusters, which is particularly useful in this context, as the aim is to obtain exactly three groups of ingredients per day (corresponding to the three main meals of the day), each with at least four ingredients. This minimum constraint prevents the creation of empty or insignificant clusters, while allowing flexibility in the maximum number of ingredients per meal.

The algorithm is applied to a vector representation of the ingredients obtained from a co-occurrence matrix. This symmetric matrix collects the frequency with which each pair of ingredients appears together in the recipes in the dataset, thus capturing real culinary patterns. Before clustering, the matrix is normalized using standard scaling to prevent ingredients with higher absolute frequencies from dominating the grouping. The result is three daily clusters; each interpreted as a meal.

Mathematically, KMeansConstrained solves the same intra-cluster variance minimization problem as traditional KMeans, but with an additional constraint on the minimum cluster size:

$$\min_{C_1,\ldots,C_k} \sum_{j=1}^{k} \sum_{x \in C_j} \left\|x - \mu_j\right\|^2 \ \textit{subject to}$$

$$\left|C_j\right| \geq s_{min} \quad \forall j \in \{1, \ldots, \mathrm{k}\}$$

E. 11: K-means objective and minimum size constraint

where μ_j is the centroid of the cluster C_j, and $s_{min} = 3$ represents the minimum number of elements allowed per group.

Once the ingredients have been grouped into three meals, actual recipes that resemble each group as closely as possible are identified. To do this, Jaccard similarity, a standard metric for comparing sets, is used. Given a set of grouped ingredients C, and a recipe R, both considered subsets of the universe of ingredients I, similarity is defined as:

$$J(C, R) = \frac{|C \cap R|}{|C \cup R|}$$

E. 12: Jaccard similarity between clustered and reference recipe ingredients

This metric, which takes values between 0 and 1, indicates the degree of coincidence between the two sets. For each meal of the day, the three recipes with the highest Jaccard similarity are selected as proposals for the user.

This system allows the outputs of the mathematical model to be translated into realistic culinary proposals, substantially improving the interpretability of the system and its applicability in everyday life.

4. EXPERIMENTS

PURPOSE OF THE EXPERIMENT

The objective of this experiment is to evaluate whether the proposed system can generate weekly nutritional plans that comply with official dietary recommendations using only age and gender as input parameters. A baseline profile-a 21-year-old woman, not pregnant or lactating-was selected to validate the feasibility and effectiveness of the model under standard conditions.

The experiment assesses whether the optimization problem is solvable in practice and whether the resulting plans meet nutritional and food group constraints. It also allows preliminary evaluation of the meal distribution and recipe suggestions to ensure the results are nutritionally valid and gastronomically coherent.

EXPERIMENTAL SETUP

COMPUTATIONAL ENVIRONMENT

The experiments were conducted on a personal laptop with an Intel® Core™ i7-1280P processor, 14 physical cores, 20 logical processors, and Windows 11 Pro as the operating system. The entire project was developed in Python 3.11.0, using libraries such as Pyomo, Gurobi, pandas, scikit-learn, and k-means-constrained.

The optimization model is solved in 10 seconds. Daily clustering using KMeansConstrained and recipe selection via Jaccard similarity also showed very low execution times (1.18 s and 2 s respectively), confirming the feasibility of the system without the need for specialized hardware.

SOLVER CONFIGURATION

The model was solved using Gurobi 10.0.3 through Pyomo, with an optimality tolerance of 0.1% (mipgap = 0.001). Symbolic model export (.lp) and live solver output (tee=True) were enabled. No additional parameters were set, as the problem was solved efficiently without further tuning.

DATASET AND MODEL SIZE

The model was executed for a 21-year-old woman, not pregnant and not lactating, generating a full weekly plan (7 days). The instance includes 223 ingredients and active constraints on 33 nutrients, 8 food groups, and 33 subgroups. In total, it contains 3,946 variables (1783 binary) and 4,276 constraints, according to the Gurobi log.

SYSTEM EFFICIENCY OVERVIEW

The system showed excellent computational performance throughout the entire pipeline. The optimization model was solved in 10 seconds, while the clustering and recipe matching stages executed in approximately 1.18 and 2 seconds, respectively. These results confirm that the proposed method is computationally lightweight and suitable for execution on standard personal hardware without requiring specialized resources. This efficiency is largely due to the compact structure of the model and the careful formulation of its constraints, which make it not only robust but also highly scalable to different user profiles and scenarios.

SYSTEM FLOW OVERVIEW

The complete process follows a modular and sequential structure, combining data preprocessing, optimization, and postprocessing. The system is designed as a pipeline with clearly defined inputs and outputs, ensuring both transparency and adaptability.

Figure 5 shows a schematic view of the system's internal flow, from user input to the final output, including the intermediate stages where ingredient data is filtered, the optimization model is solved, and postprocessing steps such as clustering and recipe matching are applied.

Inputs and Outputs:

- Input: Age and gender of the user.
- Internal Data: Nutrient intake recommendations (AESAN/EFSA), nutritional composition of ingredients, food group and frequency constraints, and a cleaned recipe database.
- Output: A weekly ingredient plan, grouped into three daily clusters, with the top matching recipes based on Jaccard similarity.

This structure ensures that the system remains efficient, modular, and adaptable to a wide range of user profiles and dietary scenarios.

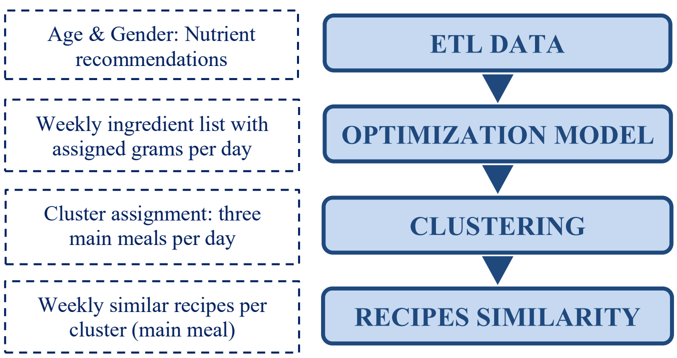

Figure 5: System pipeline

5. RESULTS

To evaluate the system, we applied the model to a user profile corresponding to a 21-year-old woman, not pregnant and not lactating. The nutritional recommendations were taken from the AESAN and EFSA guidelines and incorporated into the optimization model as lower and upper bounds on daily nutrient intake. The results presented in this section correspond to a one-week plan generated for this user and are analyzed from a nutritional, computational, and culinary perspective.

ANALYSIS OF THE GENERATED SOLUTION

The output (see Table A1 from Appendix A) includes, for each day, a set of ingredients along with the daily amount (in grams) to be consumed of each, with the objective of satisfying both nutritional recommendations and serving constraints within food groups.

In total, the model selected 66 unique ingredients, distributing them across the seven days of the week. This number reflects a balance between dietary variety and compliance with the nutritional and structural restrictions imposed by the model.

According to the optimization problem's constraints, each ingredient may be selected on a maximum of two different days throughout the week. As a result of this limitation, 53 ingredients are used on two days, while 13 ingredients appear on only one. This structure ensures a minimum level of diversity in the weekly selection, encouraging a varied combination of foods while respecting nutritional boundaries.

Figure 6 displays the number of distinct ingredients used each day, which ranges approximately from 14 to 23. This information highlights that the daily planning is not homogeneous, as a direct consequence of the frequency constraints and the various combinations required to meet daily nutritional goals. This variability adds value from both a nutritional and practical perspective, allowing for different daily structures without compromising overall consistency.

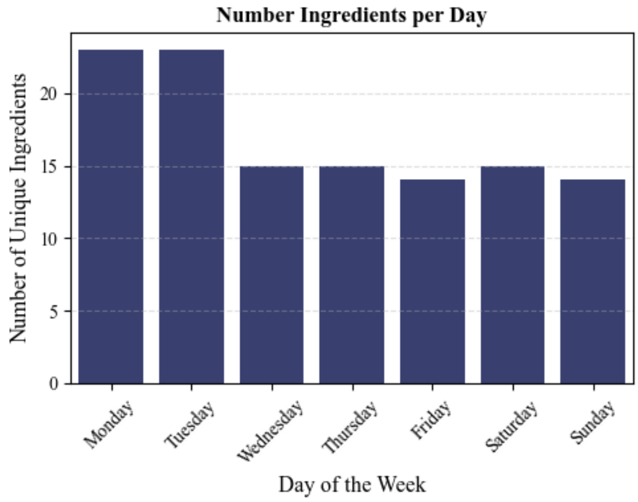

Figure 6: Total food weight per day

Figure 7 shows the total weight of food planned per day. While some day-to-day variability is observed, the values remain within a stable and reasonable range of approximately 1050 to 1500 grams. This suggests that the model has found balanced solutions in terms of total food quantity, consistently distributed across the week.

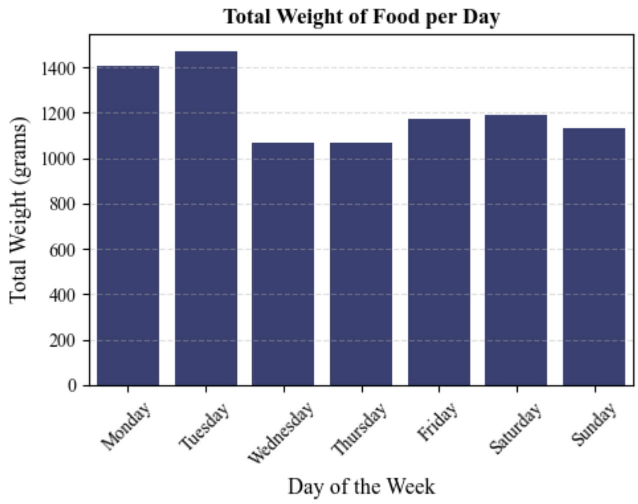

Figure 7: Number of distinct ingredients per day

While many ingredients show up on two days—including olive oil, potato starch, eggplant, bulgur, spelt flour, low-fat yogurt, and garlic, ingredients such as salt and soy sauce appear only once. This may be due to their high salt content and high culinary frequency in recipes. Therefore, including them more than once could lead to exceeding the highly restrictive sodium limits, whereas including a single portion helps increase the cumulative frequency required by E. 10.

EVALUATION OF NUTRITIONAL COMPLIANCE

A fundamental part of the analysis is to assess the extent to which the generated solution meets the established nutritional targets. The model is designed to minimize both positive and negative deviations from recommended intake values, weighting each by a penalty assigned to each nutrient based on its clinical importance.

Figure 8 shows the daily proportion of nutrients that fall within the recommended range (OK), above the maximum allowed, or below the minimum required. A relatively consistent distribution can be observed across the week, although with some variation between days. On each day, approximately 30–40% of nutrients fall below the minimum, while around 20–35% meet the target exactly. This pattern suggests that, while the model has succeeded in producing a globally consistent and balanced solution, strictly meeting minimum requirements remains the main challenge, possibly due to the concentration of certain nutrients in specific ingredients or the difficulty of adjusting multiple targets simultaneously in a multi-nutrient context.

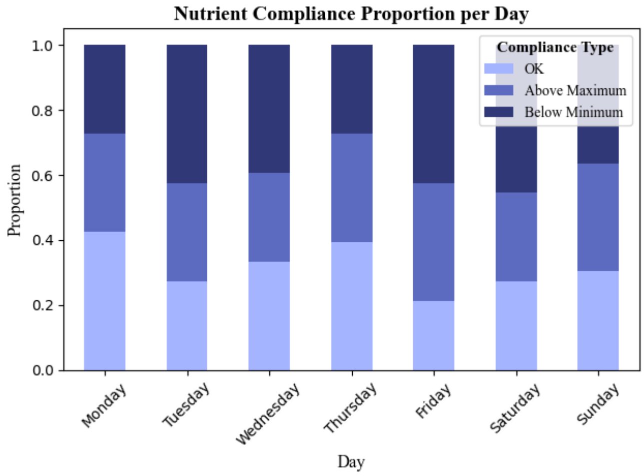

Figure 8: Daily nutrient compliance by deviation type

To further explore the individual behavior of each nutrient across the week, Figure 9 classifies nutrients into two categories: those with consistent compliance, meaning they show the same deviation type (below, above, or within range) throughout all seven days, and those with mixed compliance, where the deviation type changes over time. The pie chart shows that 15 nutrients with consistent patterns, 0% were always below the recommended minimum, 40% always exceeded the maximum, and only 20% remained within the recommended range. The bar chart focuses on nutrients with mixed behavior, such as fiber, magnesium, polyunsaturated fatty acids, salt, or zinc, which vary considerably from day to day. This variability may stem from the high sensitivity of these nutrients to small dietary changes or from their concentration in ingredients that are not used frequently, making it difficult to meet their targets consistently without affecting others.

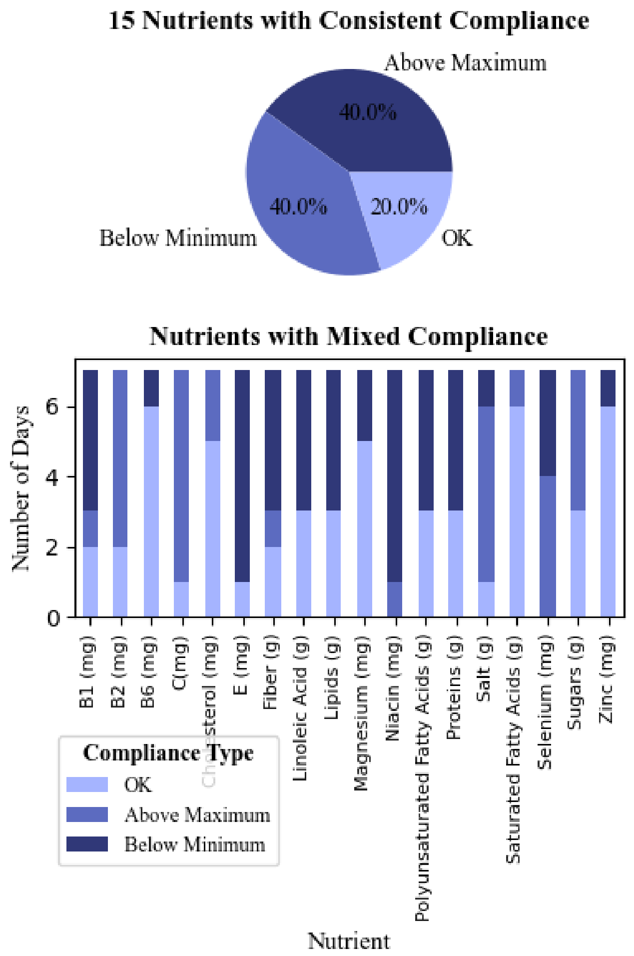

Figure 9: Nutrients with consistent vs. mixed compliance

Figure 10 presents the 15 nutrients with the highest accumulated absolute deviation over the week. The most notable deviations are observed in folates, vitamin A, and vitamin B12, followed by iodine and vitamin D. These deviations indicate that, despite the optimizer's efforts, some nutrients remain difficult to adjust accurately. This may be due to tight tolerance thresholds, limited availability in the ingredient set, or high concentration in very few foods. However, the presence of a nutrient in this ranking does not necessarily imply a major impact on the overall solution quality.

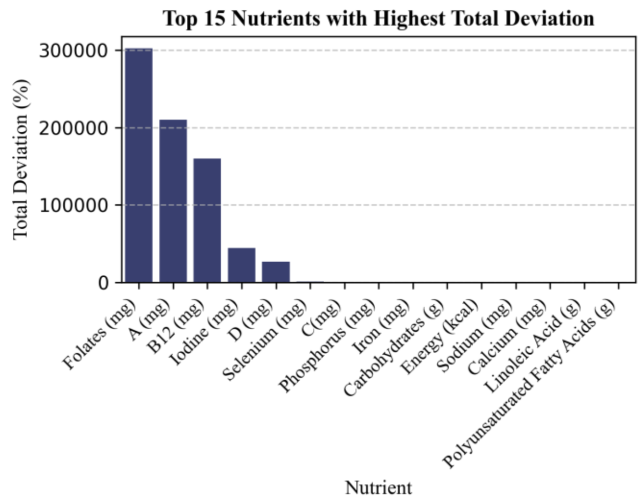

Figure 10: Top 15 nutrients with highest total weekly deviation

Figure 11 complements this analysis by showing the 15 nutrients that contribute the most to the objective function, i.e., those whose deviation, weighted by penalty, has had the greatest impact on the model. While some nutrients such as folates, vitamin A, and vitamin B12 appear again (also present in Figure 10) their relative order and prominence shift when taking into account how much each deviation affects the optimization. This demonstrates that the model does not simply react to the magnitude of the deviation, but rather to the relative importance of each nutrient. Therefore, small deviations in highly penalized nutrients may influence optimization more than large deviations in less critical ones.

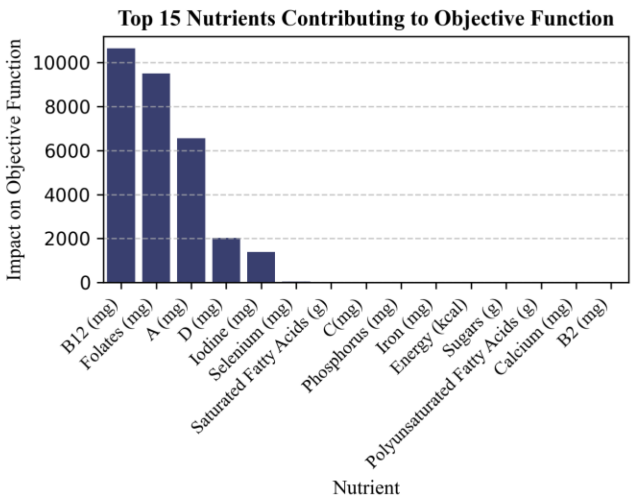

Figure 11: Top 15 nutrients with highest impact on the objective function (see E. 3)

Finally, to better understand the origin of the observed deviations and the actual influence of each ingredient on the most problematic nutrients, Figure 12 presents a heatmap showing the relative contribution of each ingredient to the nutrients most penalized by the model. This visualization helps identify which foods have the greatest nutritional impact on the most sensitive targets.

For example, horseradish is practically the only relevant source of selenium, and together with lemon and lemon juice, it also dominates the supply of vitamin C, while low-fat yogurt stands out as the main provider of iodine. In the case of vitamin B12, the contribution comes almost exclusively from crayfish, with a minor contribution from egg. Vitamin A is highly concentrated in various types of cheese, which are also the primary contributors of saturated fats.

These findings confirm that certain nutrients depend heavily on very specific ingredients, which limits the model's flexibility. Additionally, some ingredients do not stand out for any one nutrient in particular but contribute small amounts to many of them. When selected, these ingredients can cause a generalized increase in the total nutrient content. This multi-nutrient contribution can complicate the optimizer's task, since adjusting a single nutrient may unintentionally disturb other targets.

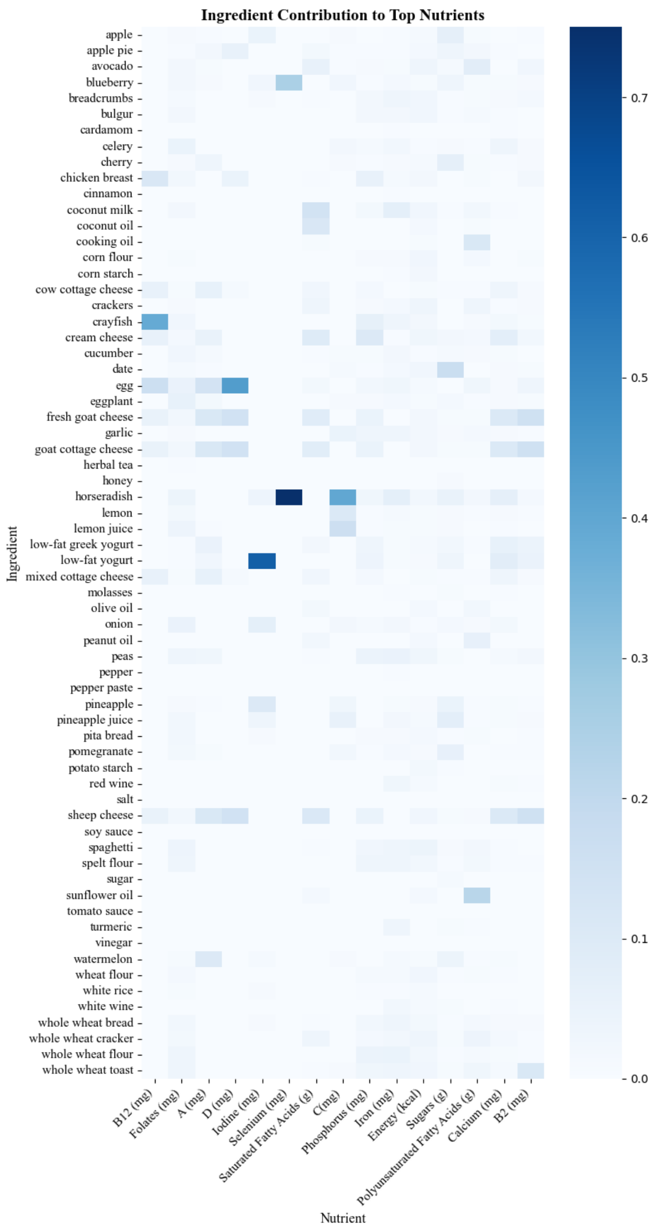

Figure 12: Relative contribution of ingredients to key nutrients

This analysis not only helps explain why certain nutrients show high deviations or strong objective impact but also reveals critical dependencies that can be addressed in future versions of the model. For instance, limiting the dominance of specific nutrients by a single ingredient or encouraging diversification through additional constraints may lead to improved balance.

EVALUATION OF MINIMUM FREQUENCY COMPLIANCE

To assess the impact of the ingredient popularity constraint (Equation E. 10) on the model's behavior, five versions of the problem were executed by varying the parameter τ between 30% and 70%. For each case, a set of key indicators was analyzed, including the objective function value, the total weekly food weight, the diversity of ingredients used, the number of ingredients used only once, as well as nutritional compliance metrics. Quantitative results are summarized in Table A2 from Appendix A meanwhile the main trends are illustrated in the following figures.

As shown in Figure 13, as the value of τ increases, the objective function value also increases. This is consistent with the fact that additional constraints are being introduced, which limit the model's flexibility and force it to select more frequent ingredients, even if they are less efficient nutritionally. However, this increase is not linear: between τ = 30% and τ = 50%, the objective value remains nearly constant, but from 60% onwards, it rises significantly.

On the other hand, the total weekly food intake also reflects this shift, though in a more irregular way. The model maintains a relatively stable intake between τ = 30% and τ = 60%, with a slight dip at τ = 50%. However, at τ = 70%, the total weight increases sharply. This may be related to the inclusion of more voluminous ingredients or the need to compensate for nutritional deviations with larger amounts.

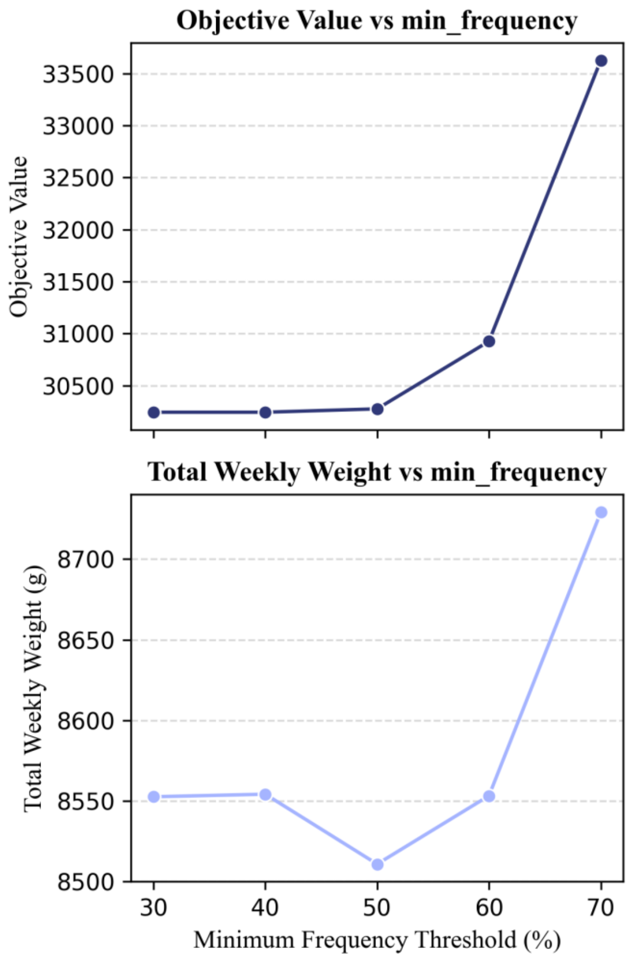

Figure 13: Evolution of objective value and weekly food weight by frequency threshold

Figure 14 shows how increasing the value of τ leads to a clear increase in the total number of unique ingredients used in the solution. This result was expected, as increasing the minimum culinary frequency threshold forces the model to diversify its selection to include ingredients present in a higher percentage of real recipes.

The bottom part of the figure shows that from τ = 50% onward, there is a significant increase in the number of ingredients that appear only once during the week, while the number of ingredients used twice slightly decrease. This phenomenon reveals that, since the frequency constraint is computed globally, the model can fulfill it by including an ingredient just once during the week without using it on more days. In this way, very frequent ingredients are added in minimal amounts solely to increase the covered frequency percentage, without significantly affecting the objective value.

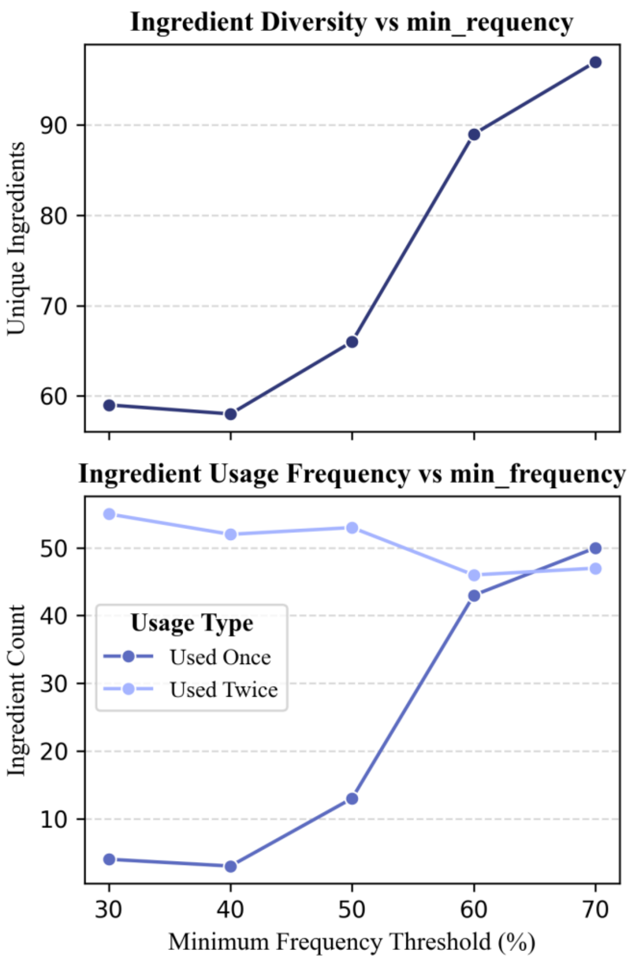

Figure 14: Weekly ingredient diversity and usage frequency by threshold

From a nutritional perspective, Figure 15 shows the average percentage of nutrients that fall below the minimum, within the acceptable range, or above the maximum threshold, for each minimum frequency value (τ). Contrary to expectations, the percentage of nutrients "in range" does not decline as restrictions increase. In fact, it remains relatively stable and even shows a slight upward trend, reaching around 36% at τ = 70%. Meanwhile, the proportion of nutrients "below minimum" decreases moderately from around 39% to 29%, suggesting that requiring the inclusion of more common ingredients helps the model avoid severe nutritional deficits

However, this improvement comes with a trade-off: the proportion of nutrients "above maximum" increases, especially from τ = 50% onward. This indicates a growing tendency to exceed upper nutritional limits as the model loses flexibility.

Although the total percentage of deviated nutrients does not rise significantly, it is possible that individual deviations - both above and below - are larger, helping explain the increase observed in Figure 13. Since the objective function penalizes not only the number of deviations but also their magnitude (weighted by the importance of each nutrient), a small group of nutrients with large deviations can greatly impact the overall cost.

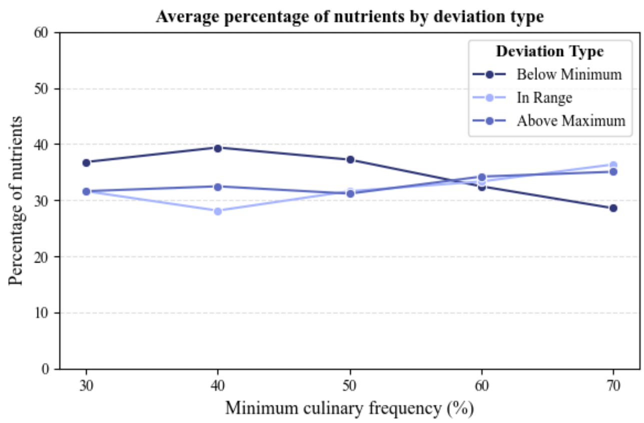

Figure 15: Average nutrient compliance by deviation type and frequency

Figure 16 shows the three nutrients that contribute the most to the objective function at each value of τ. Across all frequency thresholds, vitamin B12, folates, and vitamin A consistently dominate the objective impact. This reflects two key factors: these nutrients have relatively high penalty weights in the model due to their clinical importance, and they are also among the most difficult to meet within the available ingredient space.

Their influence remains stable across all frequency levels, which suggests that increasing culinary plausibility does not reduce their structural difficulty. In fact, small deviations in these nutrients can significantly increase the total cost of the solution.

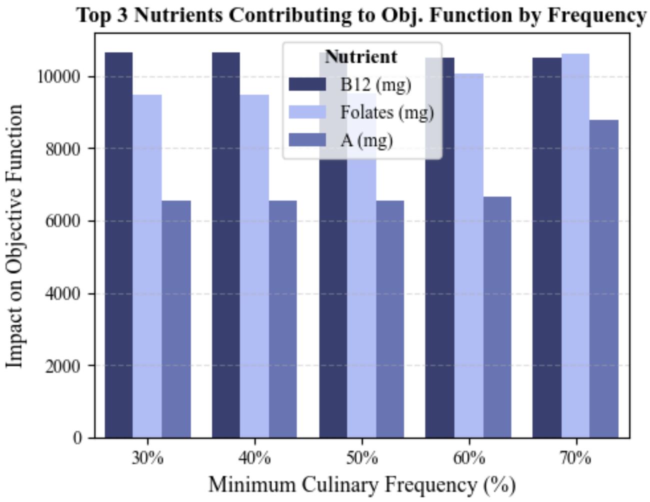

Figure 16: Top 3 nutrients contributing to the objective function by frequency

In parallel, Figure 17 presents the nutrients with the highest cumulative deviations, regardless of their penalty weight in the model. Once again, folates, vitamin A, and vitamin B12 show the highest deviations across all values of τ, with minimal variation as the frequency threshold increases. This suggests that these nutrients are systematically difficult to adjust, possibly due to their high concentration in a limited set of ingredients or because their requirements are harder to meet through the available food options.

Although Figure 15 showed that the overall percentage of nutrients within the acceptable range remains stable or improves slightly, the model continues to struggle with these specific nutrients. Their consistently high deviation values help explain the increase in the objective function observed in Figure 13, as large deviations—even if limited to a few nutrients—generate a significant penalty under the current formulation. This highlights a structural limitation: meeting the minimum frequency constraint τ pushes the model to use very common ingredients, which may be inadequate to correct certain nutrient imbalances, especially those that depend on rare or difficult-to-substitute foods.

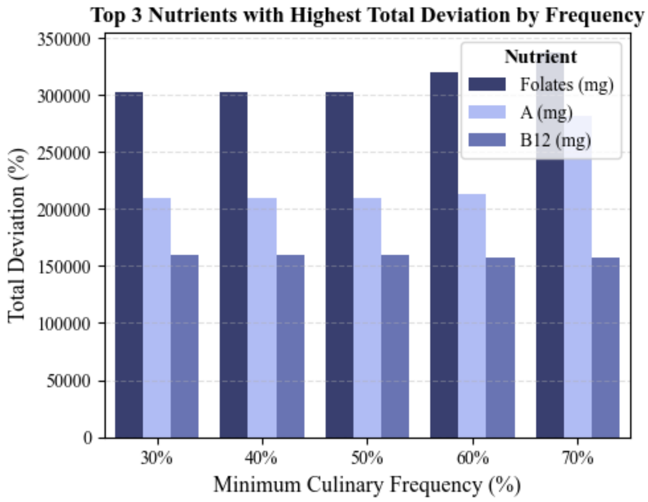

Figure 17: Top 3 nutrients with highest cumulative deviation by frequency

This analysis shows that imposing a minimum culinary frequency threshold τ has a clear impact on the model's optimal solution. While it helps improve the culinary plausibility of generated diets, it also introduces new nutritional challenges. From τ = 60% onwards, these tensions intensify, significantly affecting both the model's cost and its nutritional precision. As a result, this parameter has been set to an intermediate value (50%) to ensure a good balance between culinary realism and nutritional quality.

CLUSTERING AND RECIPE SUGGESTIONS

After solving the optimization problem, it is essential to assess whether the ingredient combinations generated by the model can be transformed into coherent culinary proposals. This section aims to evaluate whether the ingredients selected for each day can be grouped in a way that makes sense as possible meals, and to what extent these groups resemble real recipes. To support this analysis, the effect of the minimum allowed cluster size (*min_size*) is also studied, both in terms of internal structure and similarity to actual recipes, in order to find a balance between nutritional organization and culinary feasibility.

Figure 18 shows how the distribution of ingredients across the three daily clusters varies depending on the *min_size* parameter. In the upper section of the figure, corresponding to *min_size = 3*, the distribution is uneven: on several days, one cluster contains most of the ingredients while the other two barely reach the minimum threshold. This imbalance can make it difficult for each cluster to represent a complete and coherent meal. In contrast, the lower section of the figure, associated with *min_size = 4*, presents a more balanced distribution of ingredients. While some variation between days still exists, setting a higher minimum size clearly helps achieve a more even and structured segmentation, making the interpretation of the clusters as actual meals more feasible from both a nutritional and culinary standpoint.

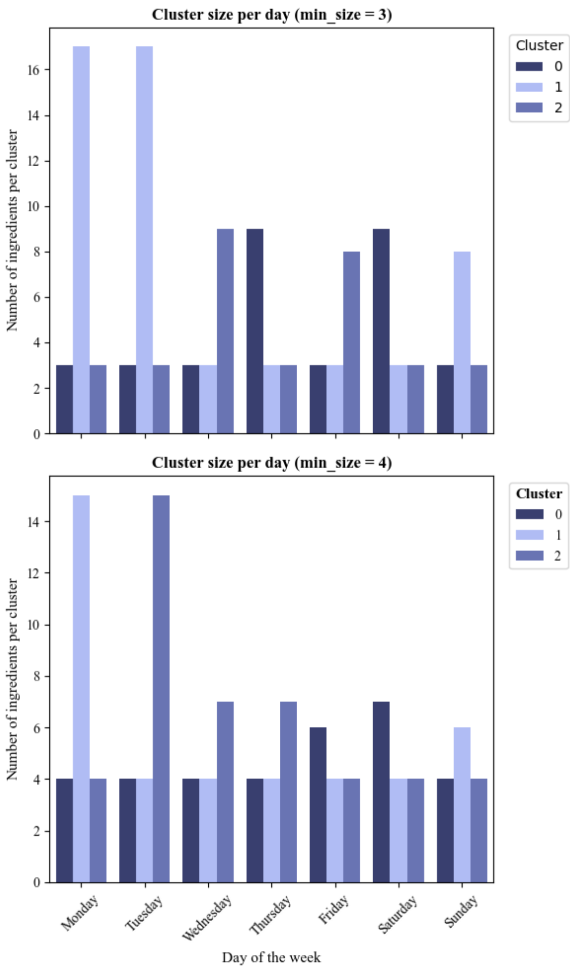

Figure 18: Distribution of ingredient count per cluster and day

Figure 19 displays the Jaccard similarity values between the clusters and the closest matching recipes from the database. In the *min_size = 3* configuration, similarity scores are significantly higher, in some cases reaching up to 70%, and generally remaining above 30% throughout the week. This is explained by the fact that smaller clusters contain simpler ingredient combinations, which increases the likelihood of matching with real recipes. However, this setting also shows greater variability: some clusters are highly similar to known recipes, while others are not, reflecting heterogeneity in group composition. On the other hand, with *min_size = 4*, similarity scores are lower overall and less dispersed. It is rare to see values above 25%, and the maximum scores are noticeably lower. This suggests that more complex and structured clusters tend to deviate more uniformly from the actual recipes found in the dataset.

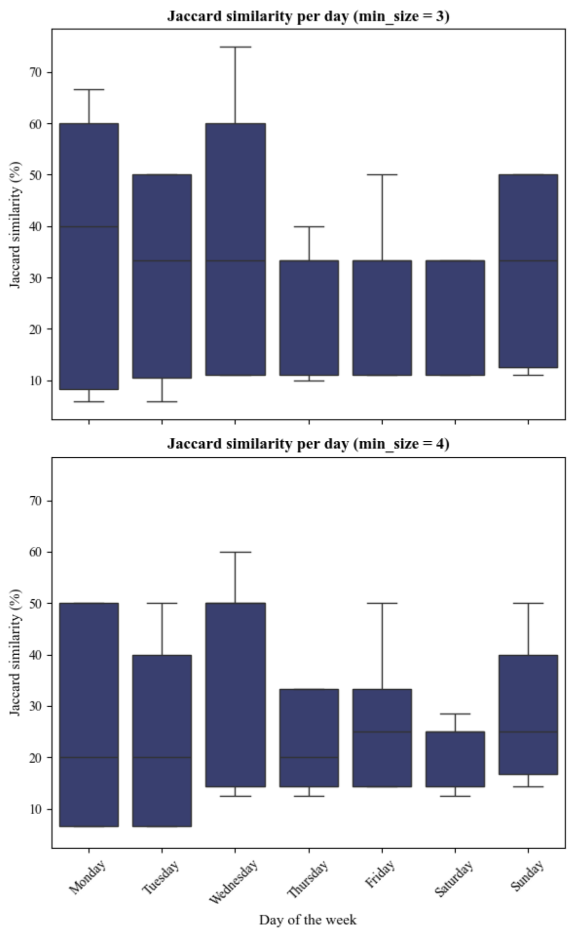

Figure 19: Jaccard similarity comparison per day

Nevertheless, these results should be interpreted carefully. First, it is important to note that the recipe ingredients were simplified during the preprocessing stage, with specific terms grouped into more generic versions to facilitate comparison. While this strategy enables compatibility between data from different sources, it also leads to a considerable loss of information. The simplification process omits relevant components or merges multiple items under a broad category. An illustrative case is presented in Table 1, where a recipe originally containing up to 17 detailed ingredients - such as “canned salmon,” “chili powder,” “Greek yogurt,” “kalamata olives,” and “panko” - is ultimately reduced to only five simplified ingredients: “olive oil,” “Greek yogurt,” “egg,” “lemon,” and “pepper”. This transformation significantly alters the culinary and nutritional interpretation of the dish.

ORIGINAL ING	***SIMPLIFIED ING***
6 teaspoons brie	Olive oil
1 can canned salmon	Greek yogurt
1 teaspoon chili powder	Egg
2 cups cornmeal	Lemon
1 egg	Pepper
2 tbs flat leaf parsley	
1 tbsp chopped garlic	
1/2 cup Greek yogurt	
3 tbsp green onion	
5 kalamata olives	
1 lemon	
2 tbsp mayonnaise	
Olive Oil	
1 cup panko	
1 teaspoon of paprika	
Pepper	
1 1/2 teaspoons salt	

Table 1: Example of ingredient loss due to simplification

In addition, the similarity metric used—the Jaccard index—considers only the presence or absence of ingredients, ignoring their quantity. This means that an ingredient used in a very small amount (e.g., 5 grams) contributes equally to the similarity score as one used in large amounts (e.g., 200 grams). As a result, recipes and clusters that merely share minor components (such as spices or condiments) may appear highly similar, even though they do not truly overlap in their key ingredients.

Figure 20 allows for a more direct examination of the relationship between cluster size and recipe similarity. A clear inverse trend emerges in both configurations: smaller clusters (containing 3 to 5 ingredients) tend to achieve similarity scores over 60%, while clusters with more than 10 ingredients rarely exceed 15%. This pattern confirms that as the complexity and size of the ingredient group increases, the likelihood of finding a significantly matching recipe decreases, especially when simplification reduces the ingredient list in recipes, as shown in Table 1.

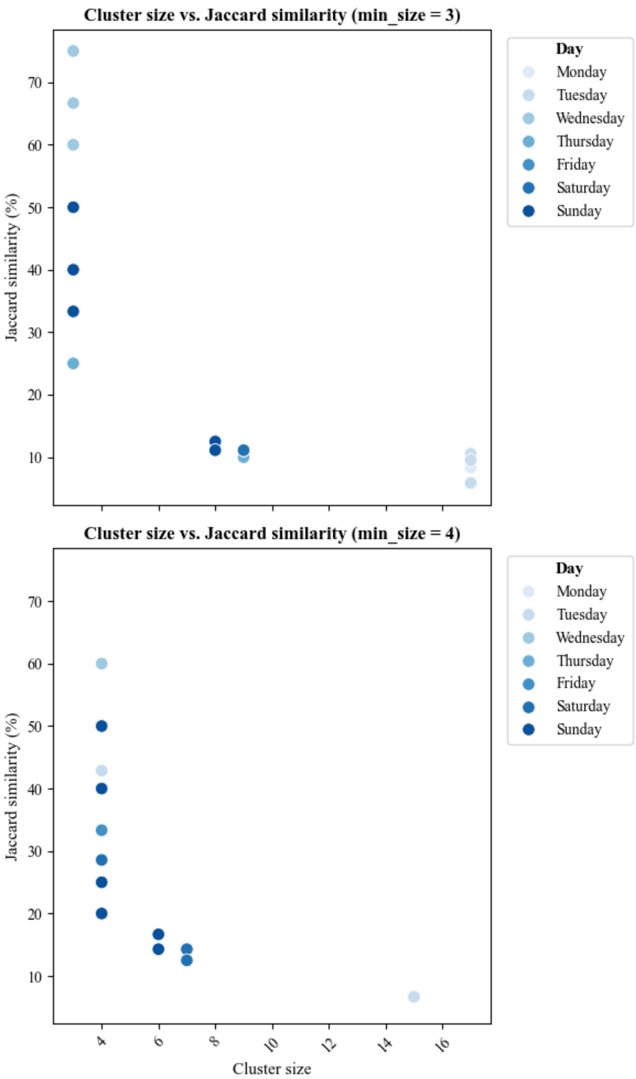

Figure 20: Relationship between cluster size and Jaccard similarity

Overall, the results reveal a necessary trade-off between similarity and structural balance. A minimum size of three achieves higher alignment with existing recipes but often results in unbalanced clusters where one group dominates the daily ingredient distribution. In contrast, a minimum size of four ensures more even and coherent grouping, though at the expense of lower similarity scores. This difference is further amplified by the simplification of ingredients and the limitations of the Jaccard metric, which does not account for nutritional weight or significance. Therefore, while three may yield more favorable numerical results, it does not always lead to better culinary or nutritional outcomes.

For this reason, the final configuration chosen for the clustering process uses four. The resulting daily ingredient clusters and their most similar recipe matches are included in Table A3 from Appendix A.

EVALUATION OF MODEL FEASIBILITY

Once the base case was thoroughly evaluated, an additional sensitivity analysis was conducted to assess the robustness and generality of the model across different population profiles. To this end, the problem was solved for all age and gender groups included in the reference database, using the same set of ingredients, frequency restrictions ($\tau = 50\%$), and model parameters in all cases.

The results, presented in Table A4 form Appendix A, show that the model is feasible for all analyzed groups, ranging from toddlers (<0.5–3 years) to older adults (70+ years), including both males and females. This confirms that the system is sufficiently flexible to adapt to profiles with very diverse nutritional needs, without requiring additional manual adjustments.

From a computational perspective, solution times remain within a reasonable range (between 5 and 10 seconds), with slight variations related to the differences in nutrient requirements and the complexity of matching them with the available ingredients.

Additionally, the value of the objective function varies significantly across profiles, which is to be expected: groups with stricter requirements - such as children - tend to produce solutions with higher penalties. This does not indicate a worse solution but rather reflects the greater nutritional complexity of certain groups. In contrast, profiles with broader or more flexible requirements, such as young adults without specific conditions, tend to yield lower objective values.

In summary, the results confirm that the model is feasible, scalable, and adaptable to different nutritional profiles, which is essential for its practical application in real-world contexts.

6. CONCLUSIONS AND FUTURE DEVELOPMENT

This project has achieved its main objective of designing a system capable of generating weekly combinations of ingredients that, at least from a nutritional standpoint, align with official nutritional recommendations based on the user's age and gender profile. Although full weekly menus with concrete recipes have not been generated, a solid foundation has been established that opens the path toward that goal.

In addition, an analysis of the culinary viability of the generated combinations has been conducted, grouping the daily ingredients into potential meals and evaluating their similarity to real recipes.

Possible future developments include:

- Using more complete and standardized data sources for both ingredients and recipes, which would improve preprocessing and reduce the loss of relevant information during ingredient simplification.
- Incorporating an intelligent system that, by learning from traditional recipes, can approximate an optimal solution for a complete weekly menu.
- Developing a visual application that makes the system accessible to any user, allowing them to input their data and receive personalized suggestions in an intuitive way, with the potential to include allergies or individual preferences.

7. BIBLIOGRAPHY

[1] American Journal of Clinical Nutrition. *Dietary optimization models*. *Am. J. Clin. Nutr.*, vol. 112, no. 5, 2022. [Online]. Available: https://ajcn.nutrition.org/article/S00029165(22)10525-3/fulltext.

[2] E. Bar, "Gut Israeli research: How microbiome-based personalized diets can improve health," NoCamels, Jun. 2021. [Online]. Available: https://nocamels.com/2021/06/gut israeli-research-microbiome-personalized-diets/.

[3] M. Kanjilal, U. Kumar, G. K. Gupta, D. Agrawal, R. K. Arya, and J. Batra, "Dietary habits and their impact on the physical status of school-going adolescents in Delhi: A cross-sectional study," Journal of Clinical and Diagnostic Research, vol. 15, no. 7, pp. OC43–OC47, 2021. [Online]. Available: JCDR - Body mass index, Food preferences, Junk food, Undernutrition.

[4] MDPI, "Polyphenol-rich foods for cardiovascular health," Nutrients, vol. 14, no. 3, pp. 545, 2023. [Online]. Available: https://www.mdpi.com/2072-6643/14/3/545.

[5] Instituto Nacional de Saúde Doutor Ricardo Jorge, I.P., *Tabela da Composição de Alimentos – v6.0*, 2023. [Online]. Available: https://portfir-insa.min-saude.pt/

[6] Comité Científico AESAN, *Informe del Comité Científico de la Agencia Española de Seguridad Alimentaria y Nutrición (AESAN) de revisión y actualización de las Recomendaciones Dietéticas para la población española*, Revista del Comité Científico nº 32, 2020, pp. 11–58. [Online]. Available: https://www.aesan.gob.es/

[7] Comité Científico AESAN, *Informe del Comité Científico de la Agencia Española de Seguridad Alimentaria y Nutrición (AESAN) sobre Ingestas Nutricionales de Referencia para la población española*, Revista del Comité Científico nº 29, 2019. [Online]. Available: https://www.aesan.gob.es/

[8] R. M. Ortega, A. M. López-Sobaler, A. Aparicio, E. Rodríguez-Rodríguez, L. G. González-Rodríguez, J. M. Perea, and B. Navia, *Objetivos nutricionales para la población española*, Departamento de Nutrición, Facultad de Farmacia, Universidad Complutense de Madrid, 2021. [Online]. Available: https://www.ucm.es/idinutricion/file/objetivos-nutricionales-2021-1?ver

[9] EFSA Panel on Dietetic Products, Nutrition, and Allergies (NDA), *Scientific Opinion on Dietary Reference Values for fats, including saturated fatty acids, polyunsaturated fatty acids, monounsaturated fatty acids, trans fatty acids, and cholesterol*, EFSA Journal, vol. 8, no. 3, 2010, Art. no. 1461. doi: 10.2903/j.efsa.2010.1461. [Online]. Available: https://efsa.onlinelibrary.wiley.com/doi/10.2903/j.efsa.2010.1461

[10] EFSA Panel on Nutrition, Novel Foods and Food Allergens (NDA), *Scientific Opinion on the dietary reference values for sodium*, EFSA Journal, vol. 17, no. 9, 2019, Art. no. 5778. doi: 10.2903/j.efsa.2019.5778. [Online]. Available: https://efsa.onlinelibrary.wiley.com/doi/10.2903/j.efsa.2019.5778

[11] Spoonacular API, *Spoonacular – The Food API*, [Online]. Available: https://spoonacular.com/food-api

APPENDIX A

LUNES	*MARTES*	*MIÉRCOLES*	*JUEVES*	*VIERNES*	*SÁBADO*	*DOMINGO*
Herbal tea (200)	Herbal tea (200)	Eggplant (150)	Cucumber (150)	Cucumber (150)	Eggplant (150)	Coconut milk (150)
Celery (150)	Celery (150)	Lemon juice (150)	Garlic (150)	Horseradish (150)	Lemon juice (150)	Garlic (150)
Watermelon (150)	Watermelon (150)	Onion (150)	Red wine (109)	Pineapple juice (150)	Onion (150)	Horseradish (150)
Low-fat yogurt (120)	Low-fat yogurt (120)	Crayfish (125)	Cherry (100)	Low-fat Greek yogurt (120)	Crayfish (125)	Pineapple juice (150)
Apple (100)	White wine (104.167)	Avocado (100)	Date (100)	Red wine (109)	White wine (104)	Low-fat Greek yogurt (125)
Blueberry (100)	Apple (100)	Date (100)	Pineapple (100)	Avocado (100)	Blueberry (100)	Pineapple (100)
Chicken breast (100)	Chicken breast (100)	Goat cottage cheese (60)	Cream cheese (60)	Cherry (100)	Pomegranate (100)	Fresh goat cheese (60)
Lemon (100)	Lemon (100)	Mixed cottage cheese (60)	Fresh goat cheese (60)	Cream cheese (60)	Cow cottage cheese (60)	Egg (53)
Sheep cheese (60)	Pomegranate (100)	Corn flour (40)	Spaghetti (60)	Spaghetti (60)	Goat cottage cheese (60)	Breadcrumbs (40)
White rice (60)	Sheep cheese (60)	Whole wheat bread (40)	Peas (50)	Egg (53)	White rice (60)	Crackers (40)
Peas (50)	Apple pie (40)	Whole wheat cracker (40)	Breadcrumbs (40)	Crackers (40)	Whole wheat bread (40)	Pita bread (40)
Apple pie (40)	Bulgur (40)	Whole wheat flour (40)	Wheat flour (40)	Pita bread (40)	Whole wheat cracker (40)	Wheat flour (40)
Bulgur (40)	Corn flour (40)	Honey (5)	Whole wheat toast (40)	Whole wheat toast (40)	Whole wheat flour (40)	Potato starch (30)
Spelt flour (40)	Spelt flour (40)	Sugar (5)	Sunflower oil (10)	Pepper paste (1)	Peanut oil (10)	Sunflower oil (10)
Corn starch (30)	Corn starch (30)	Tomato sauce (1)	Vinegar (1)		Turmeric (3)	
Potato Starch (30)	Coconut oil (10)					
Coconut oil (10)	Cooking oil (10)					
Peanut oil (10)	Olive oil (10)					
Olive oil (10)	Molasses (5)					
Cardamom (3)	Vinegar (5)					
Cinnamon (1)	Cardamom (3)					
Pepper (1)	Turmeric (3)					
Soy sauce (1)	Salt (1)					

Table A1: Problem solution - weekly list of ingredients and grams

FREQUENCY (%)	*OBJ. VALUE*	*TOTAL GRAMS*	*NUM. ING*	*USED ONCE*	*USED TWICE*
30%	30243.7	8552.52	59	4	55
40%	30243.5	8554.10	58	3	52
50%	30275.8	8510.72	66	13	53
60%	30925.6	8553.12	89	43	46
70%	33626.2	8729.11	97	50	47

FREQUENCY (%)	**NUTR. BELOW MIN (%)**	**NUTR. IN RANGE (%)**	**NUTR. ABOVE MAX (%)**
30%	36.8 %	31.6 %	31.6 %
40%	39.4 %	28.1 %	32.5 %
50%	37.2 %	31.6 %	31.2 %
60%	32.5 %	33.3 %	34.2 %
70%	28.6 %	36.4 %	35 %

Table A2: Quantitative analysis of the impact of the minimum ingredient frequency threshold

DAY	*CLUSTER*	*CLUSTER ING*	SIMILARITY	RECIPE	SIMPLIFIED *ING*
Monday	0	Celery, Soy sauce, Cardamom, Peas	20	Ginger Garlic Chili Salmon	Soy sauce, Vinegar
			20	Kimchi/Kimchee / (Korean Fermented Spicy Cabbage)	Soy sauce, Crackers
			20	Ground Pork Ramen	Soy sauce, Garlic
	1	Blueberry, Apple, Bulgur, Potato starch, Apple pie, Corn starch, Low-fat yogurt, Watermelon, Spelt flour, Coconut oil, Herbal tea, Cinnamon, Sheep cheese, White rice, Peanut oil	6.67	Vegetable Dip	Low-fat yogurt
			6,67	Yogurt Parfait	Low-fat yogurt
			6,67	Chicken Enchilada Casserole	Sheep cheese
	2	Lemon, Chicken breast, Olive oil, Pepper	50	Quinoa, Tomato, Green Onion Side Salad	Lemon, Olive oil
			50	Watermelon, Feta and Mint Salad	Olive oil, Cucumber, Salt, Lemon, Pepper
			50	Fried Salmon Cakes	Olive oil, Greek yogurt, Egg, Lemon, Pepper
Tuesday	0	Celery, Lemon, Corn flour, Turmeric	25	Spice-Rubbed Lemon Barbecue Salmon	Lemon
			20	Doughnuts	Corn flour, Honey

			20	Quinoa, Tomato, Green Onion Side Salad	Lemon, Olive oil
	1	Chicken breast, Olive oil, White wine, Salt	50	Classic French Mussels	Olive oil, White wine
			42,86	Chicken, Red Pepper, and White Bean Chili	Chicken breast, Olive oil, Chicken broth cube, Bell pepper, Salt, Garlic
			40	Herb chicken with sweet potato mash and sautéed broccoli	Chicken breast, Olive oil, Sweet potato
	2	Apple, Bulgur, Cardamom, Low-fat yogurt, Corn starch, Cooking oil, Apple pie, Watermelon, Pomegranate, Spelt flour, Coconut oil, Herbal tea, Molasses, Vinegar, Sheep	6,67	Vegetable Dip	Low-fat yogurt
			6,67	Fall Farro Salad with Pomegranate, Walnut & Truffles	Pomegranate
			6,67	Yogurt Parfait	Low-fat yogurt
Wednesday	0	Corn flour, Whole wheat flour, Honey, Sugar	50	Doughnuts	Corn flour, Honey
			50	Cherry-Berry Pie With Agave Nectar All-Butter Crust	Corn flour, Whole wheat flour
			33,33	Strawberry Shortcake Pancakes	Baking powder, Greek yogurt, Whole wheat flour, Sugar
	1	Tomato sauce, Avocado, Lemon juice, Onion	60	Vegan Taco bowls with	Tomato, Onion, Avocado, Lemon juice

				Cilantro Lime Cauliflower Rice	
			60	Pulled Pork Sandwich with Mango BBQ sauce	Tomato sauce, Onion, Lemon juice, Coriander seeds
			50	(Houston's) Spinach Dip	Chicken broth cube, Unsalted butter, Lemon juice, Tomato sauce, Onion
	2	Crayfish, Whole wheat bread, Eggplant, Whole wheat cracker, Mixed cottage cheese, Goat cottage cheese, Date	14,29	Coconut and Whole Wheat Chicken Tenders	Whole wheat bread
			14,29	Fettuccine With Smashed Peas	Mixed cottage cheese
			12,5	How Sweet It Is Sweet Potato Lasagne	Eggplant, Sweet potato
Thursday	0	Cream cheese, Red wine, Vinegar, Peas	25	Tomato, Cucumber & Onion Salad with Feta Cheese: Real Convenience Food	Cream cheese
			20	Spinach Salad with Strawberry Vinaigrette	Bell pepper, Cream cheese
			20	Ginger Garlic Chili Salmon	Soy sauce, Vinegar

	1	Cucumber, Sunflower oil, Garlic, Breadcrumbs	33.33	Sweet Pepper and Heirloom Tomato Gazpacho	Cucumber, Olive oil, Garlic, Parsley
			33.33	Bobby Flay's Meatball & Sauce	Tomato sauce, Olive oil, Garlic, Breadcrumbs
			33.33	Italian Beef Braciole	Red wine, Olive oil, Garlic, Breadcrumbs
	2	Spaghetti, Wheat flour, Pineapple, Whole wheat toast, Cherry, Fresh goat cheese, Date	14.29	Almond Crusted Salmon Fillets with Roasted Broccolini	Wheat flour
			14.29	TROPICAL BANANA GREEN SMOOTHIE	Pineapple
			12.5	Mediterranean Spinach Artichoke Dip	Fresh goat cheese, Low-fat Greek yogurt
	0	Spaghetti, Cherry, Horseradish, Red wine, Crackers, Pineapple juice	14.29	Kimchi/Kimchee/Gimchi (Korean Fermented Spicy Cabbage)	Soy sauce, Crackers
Friday			14.29	Easy Beef Bourguignon	Beef broth cube, Red wine
			14.29	Braised Lamb Chops	Pepper, Pineapple juice
	1		50	Chopped Beet Salad	Cream cheese, Pepper paste

	2	Cream cheese, Cucumber, Low-fat Greek yogurt, Pepper paste	40	Spring Salad with Radishes and Beets	Spinach, Cream cheese, Pepper paste
			33.33	Layered Chicken Salad With Couscous	Chicken breast, Cucumber, Sweet potato, Pepper paste
		Pita bread, Egg, Avocado, Whole wheat toast	25	Caramel Peanut Fudge Cake	Egg
			25	Classic Eggs Benedict	Egg
			20	Ginger Lentils With Carrots and Fresh Herbs	Avocado, Lentils
Saturday	0	Crayfish, Whole wheat flour, Whole wheat cracker, Pomegranate, Goat cottage cheese, White rice, Peanut oil	14.29	Fall Farro Salad with Pomegranate, Walnut & Truffles	Pomegranate
			14.29	Chimichurri Skirt Steak with Grilled Asparagus	Whole wheat flour
			12.5	Sweet-N-Smoky Salmon With Ginger Mahogany Rice	Chicken broth cube, Peanut oil
	1	Blueberry, White wine, Turmeric, Onion	28.57	Greek-Style Baked Fish: Fresh, Simple, and Delicious	Bell pepper, White wine, Oregano, Tomato, Onion

			25	Asparagus and Pea Soup: Real Convenience Food	Onion
			25	Heirloom Tomato Basil and Olive Oil Wine Sauce over Pasta	White wine
	2	Eggplant, Whole wheat bread, Cow cottage cheese, Lemon juice	25	Chimichurri Skirt Steak with Grilled Asparagus	Lemon juice
			25	Roasted Alaskan Salmon with Cherries	Lemon juice
			25	Aloo Baingan	Lemon juice
Sunday	0	Crackers, Horseradish, Pineapple juice, Pineapple	40	Powerhouse Almond Matcha Superfood Smoothie	Crackers, Pineapple, Coriander seeds
			25	TROPICAL BANANA GREEN SMOOTHIE	Pineapple
			20	Kimchi/Kimchee/Gimchi (Korean Fermented Spicy Cabbage)	Soy sauce, Crackers

	1	Wheat flour, Potato starch, Fresh goat cheese, Sunflower oil, Pita bread, Coconut milk	16.67	Rice and Peas with Coconut Curry Mackerel	Coconut milk
			16.67	Almond Crusted Salmon Fillets with Roasted Broccolini	Wheat flour
			14.29	Lamb Coconut Tomato Curry Sauce	Vegetable cream, Coconut milk
	2	Egg, Low-fat Greek yogurt, Garlic, Breadcrumbs	50	Breaded Shrimp and Spicy Mayo Appetizer	Egg, Breadcrumbs
			40	Corned Beef Cakes	Chicken broth cube, Egg, Breadcrumbs
			40	Bread Omlette	Whole milk, Egg, Breadcrumbs

Table A3:Daily ingredients clusters and top-3 matched recipes with simplified ingredients

GROUP	*AGE*	*OBJ.VALUE*	*EXCUTION TIME*	*NUM SOLUTIONS*
	< 0.5	2707557.1	7,8 s	2
	0.5 - 1	2632800.4	5,95 s	4
Children	2 - 3	66835	5,88 s	7
	4 - 5	58657.9	7,06 s	7
	6 - 10	43605.3	6,59 s	5
	10 - 13	30548.2	7,2 s	8
	14 - 19	26034.3	6,85 s	8
	20 - 39	24833.1	7,32 s	9
Males	40 - 49	24830.9	6,81 s	4
	50 - 59	24828	6,45 s	5
	60 - 69	24651.2	8,04 s	5
	≥70	24503	7,08 s	5
	10 - 13	30597.1	9,09 s	9
	14 - 19	30511.7	9,79 s	10
	20 - 39	30275.8	8,14 s	7
Women	40 - 49	25675.5	8,4 s	5
	50 - 59	25678.6	7,88 s	7
	60 - 69	25496.1	8,56 s	9
	≥70	25359.3	7,92 s	4
Pregnancy (2º half)	Any	26864.8	7,52 s	6
Pregnancy	Any	19235.9	8,04 s	5

Table A4: Madel performance across age and gender groups