Sistemas Híbridos de Recomendación de Recetas Saludables y Personalizadas

Francisco Arenas * 1 Sofía Rebolledo 1 Alvaro Romero * 1

Abstract

En este estudio evaluamos distintos enfoques para un sistema híbrido de recomendación de recetas que integre tanto interacciones de usuarios y recetas, características semánticas de las recetas sobre ingredientes y etiquetas, y metadatos nutricionales. Se compararon métodos sobre dos grandes colecciones de datos; Food.com y MealRecPlus. Comparamos métodos clásicos como KNN, Most Popular y LightFM, y una arquitectura avanzada como DeepFM. Además, proponemos un post-procesado por etiquetas dietéticas que permite al usuario personalizar sus recomendaciones. Los experimentos muestran que DeepFM alcanza los mejores resultados, con una Precision@10 de 0.9430 y una Recall@10 de 0.5705 en Food.com y altos niveles de novedad (15.16) y diversidad (1.00), mientras que los enfoques clásicos se quedan muy por debajo. Finalmente, discutimos limitaciones como la falta de métricas dietéticas y validación con usuarios reales, y proponemos futuras líneas de trabajo.

1. Introducción

En la última década, el crecimiento de plataformas digitales de recetas ha dado lugar a bibliotecas masivas con millones de usuarios, recetas y reseñas disponibles. Paralelamente, los sistemas de recomendación han evolucionado desde métodos puros de filtrado colaborativo, que explotan patrones de comportamiento de usuarios similares, hasta enfoques híbridos que incorporan contenido semántico y metadatos nutricionales. Rostami et al. (2023) demostraron que los modelos híbridos que combinan señales de usuario, atributos de ítem y características de contenido logran mejoras significativas en precisión y diversidad en comparación con técnicas convencionales.

Sin embargo, en el dominio de recetas persisten tres retos principales:

Proyecto de curso de IIC3633 Sistemas Recomendadores 2025-1.

- Relevancia versus salud: equilibrar la probabilidad de interacción con la calidad nutricional de la recomendación.
- 2. **Novedad y diversidad:** exponer al usuario a recetas variadas y menos populares sin sacrificar relevancia.
- Personalización dietética: ofrecer recomendaciones filtradas por etiquetas nutricionales ("alto en proteína", "bajo en carbohidratos", "vegano", etc.) de forma intuitiva y eficaz.

En este trabajo abordamos estos retos mediante:

- Implementación y comparación de modelos de filtrado colaborativo (LightFM) y de factorización profunda (DeepFM) en los conjuntos de datos de Food.com y MealRecPlus.
- Enriquecimiento de LightFM con contenido (TF-IDF + SVD de ingredientes) y metadatos de salud (FSA, WHO), mejorando Precision@10 de 0.09 a 0.145 en MealRecPlus.
- Desarrollo de una estrategia híbrida por etiquetas que permite a los usuarios guiar la recomendación según sus preferencias dietéticas, combinando similitud de etiquetas y embeddings del modelo.
- Evaluación exhaustiva con métricas de relevancia (Precision@K, Recall@K), novedad (self-information) y diversidad intra-lista.

2. Datasets

Para abordar el objetivo de esta investigación, se utilizaron 2 datasets que contienen información relevante para los sistemas recomendadores a evaluar.

2.1. Food.com - Recipes and Reviews

El dataset de Food.com consta de dos tablas principales:

Recipes: 522517 recetas clasificadas en 312 categorías. Cada receta contiene nombre, categoría y otros metadatos como ingredientes, palabras clave y valores nutricionales.

^{*}Contribución equivalente ¹Departamento de Ciencia de la Computación, Pontificia Universidad Católica de Chile, Santiago, Chile.

Reviews: 1 401 982 reseñas escritas por 271 907 usuarios distintos. Cada reseña contiene autor, rating y metadatos adicionales.

2.2. MealRecPlus

El dataset de MealRecPlus consta de dos tablas principales:

- Recipes: 11 253 recetas. Cada receta contiene nombre nombre, rating, categoria, ingredientes y tags
- **Reviews**: 151 148 reseñas escritas por 1 576 usuarios distintos. cada reseña contiene autor, *rating* y critica

3. Metodología

En esta sección describimos el diseño experimental, los modelos evaluados y las métricas utilizadas.

3.1. Diseño Experimental

Para cada conjunto de datos (Food.com y MealRecPlus) seguimos el siguiente protocolo:

1. Preprocesamiento de datos:

- Eliminación de usuarios e ítems con menos de 5 interacciones.
- Limpieza y tokenización de los campos de texto (*ingredientes*, *tags*).
- Conversión de metadatos nutricionales (FSA, WHO) a *buckets* categóricos.

2. División Train/Test:

- Partición aleatoria de interacciones: 75% para entrenamiento, 25% para test.
- Uso de random_train_test_split de LightFM, asegurando que todos los usuarios de test aparezcan en train.

3. Búsqueda de Hiperparámetros (solo LightFM y DeepFM):

- Grid search sobre no_components, learning_rate, item_alpha, user_alpha y loss (warp, bpr).
- Criterio de selección: maximizar Precision@10
 en un conjunto de validación interno.

3.2. Modelos Evaluados

3.2.1. SELECCIÓN DE MODELOS

La elección de los modelos empleados en este trabajo responde a la necesidad de abordar distintas características observadas en los datos, así como a la intención de comparar enfoques clásicos, híbridos y de aprendizaje profundo dentro del contexto de sistemas de recomendación de recetas.

LightFM fue seleccionado como modelo base debido a su capacidad de combinar el enfoque de filtrado colaborativo con características adicionales tanto de usuarios como de ítems. En particular, su arquitectura basada en *factorization machines* permite incorporar información semántica (como embeddings de ingredientes o etiquetas nutricionales) junto con la matriz de interacciones usuario-receta. Esta flexibilidad resulta ideal para escenarios como *MealRec-Plus*, donde existe abundante metainformación sobre salud y composición de los alimentos. Además, LightFM permite explorar múltiples funciones de pérdida como WARP y BPR, adecuadas para tareas de ranking implícito, donde no se cuenta con ratings explícitos sino solo con señales de consumo.

Por otro lado, se empleó **DeepFM** exclusivamente sobre el dataset *Food.com*, el cual posee una mayor densidad de interacciones y una estructura más adecuada para modelos complejos. DeepFM extiende el paradigma de *factorization machines* al combinarlo con redes neuronales profundas, lo que permite capturar no solo interacciones lineales sino también patrones de mayor orden entre múltiples características heterogéneas. Esta arquitectura híbrida es particularmente ventajosa cuando se dispone de gran variedad de atributos categóricos (como categorías de receta, ingredientes, puntuaciones nutricionales, etc.), y se busca modelar tanto dependencias explícitas como representaciones latentes más abstractas.

La combinación de estos dos enfoques nos permite explorar el espectro entre modelos ligeros con capacidad explicativa (LightFM) y arquitecturas más expresivas orientadas a rendimiento predictivo (DeepFM), con el objetivo de evaluar no solo la precisión de las recomendaciones, sino también aspectos clave como diversidad, novedad y capacidad de personalización.

3.2.2. BASELINES

- Random Predictor: asigna predicciones aleatorias dentro del rango de ratings.
- Most Popular: recomienda ítems ordenados por número de interacciones.
- KNN: filtrado colaborativo basado en vecinos más cercanos.

3.2.3. LIGHTFM

- Factorization Machine con pérdida WARP/BPR.
- Versión pura: usa solo la matriz de interacciones.
- Versión enriquecida: añade como features:

- Embeddings de contenido (TF–IDF + SVD de ingredientes y tags).
- Buckets nutricionales (FSA, WHO) para usuarios e ítems.

3.2.4. DEEPFM (SOLO FOOD.COM)

- Arquitectura híbrida que combina un componente lineal y una red neuronal profunda.
- Input: vectores one-hot de usuario e ítem + embeddings de contenido y nutrición.
- Entrenamiento con optimizador Adam y pérdida logloss.

3.2.5. ESTRATEGIA HÍBRIDA POR TAGS

• Post-procesado que reordena el top-N de LightFM según

$$score_{hybrid}(i) = \alpha sim_{tag}(i) + (1 - \alpha) sim_{model}(i).$$

- sim_{tag}: similitud de Jaccard entre las etiquetas de la receta y el conjunto de tags preferidos.
- sim_{model}: producto escalar entre el embedding del ítem y el embedding promedio de usuarios.

3.3. Métricas de Evaluación

- Precision@K y Recall@K: proporción de ítems relevantes en el top-K.
- **Novedad:** Para cada ítem *i* se define su novedad como:

$$nov(i) = -\log_2(pop(i)),$$

donde pop(i) representa la proporción de usuarios en el sistema que han interactuado con el ítem i. Cuanto menor es la popularidad, mayor es la novedad. Para un usuario u con su lista de recomendación $R_u = [i_1, i_2, \ldots, i_K]$, la novedad promedio es:

$$\operatorname{novelty}(R_u) = \frac{1}{K} \sum_{k=1}^{K} \operatorname{nov}(i_k),$$

donde K es la cantidad de ítems recomendados y i_k es el k-ésimo ítem en la lista recomendada. Finalmente, la novedad global se obtiene promediando sobre todos los usuarios.

• Diversidad intra-lista: Se calcula como:

$$D(R_u) = \frac{2}{K(K-1)} \sum_{1 \le a < b \le K} (1 - \operatorname{sim}_{\cos}(i_a, i_b)),$$

donde:

- $R_u = [i_1, i_2, \dots, i_K]$ es la lista de recomendaciones para el usuario u,
- K es la longitud de dicha lista,
- $sim_{cos}(i_a, i_b)$ es la similitud coseno entre los vectores de representación (embeddings) de los ítems i_a e i_b .

Esta métrica promedia la disimilitud (1 - similitud) de todos los pares únicos en la lista. Un valor alto indica mayor diversidad (menor similitud promedio entre ítems recomendados).

4. Análisis

En esta sección examinamos detalladamente el rendimiento de los modelos y el comportamiento de los hiperparámetros clave.

4.1. Food.com

Modelo	Prec@10	Rec@10	Nov@10	Div@10
KNN	0.0024	0.0010	_	_
LightFM	0.0102	0.0221	_	_
LightFM				
(+ nutrición)	0.0102	0.0221	1.3431	0.0414
DeepFM	0.9430	0.5705	15.1629	1.0037

Table 1. Comparativa de métricas en Food.com

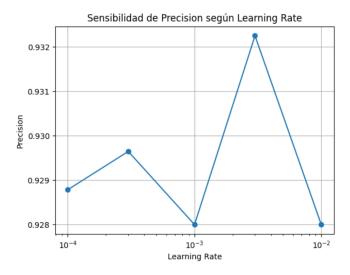
- KNN vs LightFM: KNN ofrece precisión casi nula; LightFM duplica el desempeño, pero sigue siendo bajo.
- Efecto de nutrición: añadir buckets FSA/WHO no modifica la precisión, aunque sí aumenta novedad y diversidad.
- **DeepFM:** es robusto frente a variaciones de hiperparámetros y logra un salto considerable en *Prec@10* y *Rec@10*.

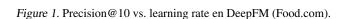
Sensibilidad de DeepFM A continuación, la variación de *Precision@10* según el *learning rate* en DeepFM:

4.2. MealRecPlus

- Impacto de salud: incorporar FSA/WHO eleva Prec@10 de 0.0954 a 0.1372.
- Contenido semántico: añadir TF–IDF+SVD aumenta Prec@10 a 0.1453, pero la diversidad empeora.

Sensibilidad de LightFM Variación de *Precision@10* según el número de componentes latentes y función de pérdida:





Configuración	Prec@10	Rec@10	Nov@10	Div@10	
LightFM	0.0954	0.0746	8.791	0.0	
(puro)					
LightFM	0.1372	0.0749	8.7002	0.0205	
(FSA/WHO)					
LightFM					
(TF-IDF	0.1453	0.0745	8.0028	0.0020	
+SVD + FSA/W	VHO)				

Table 2. Comparativa de métricas en MealRecPlus

5. Resultados

Para ilustrar la diferencia cualitativa entre modelos, la Tabla 3 presenta ejemplos reales de las 5 recetas más recomendadas por los sistemas DeepFM y LightFM (con y sin tags dietéticos). Se observa que DeepFM recomienda ítems diversos y ricos en categorías, mientras que LightFM sin tags tiende a priorizar platos populares y postres. La versión híbrida de LightFM permite al usuario guiar el filtrado temático —por ejemplo, eligiendo recetas etiquetadas como vegetarianas— lo cual se refleja en la coherencia de los resultados obtenidos.

Ahora para evaluar la eficiencia y capacidad de personalización del enfoque por tags, presentamos a continuación el desglose de las Top-10 recomendaciones de LightFM al aplicar distintos filtros nutricionales Tabla 4.

Los resultados confirman que:

 Relevancia vs saludabilidad: DeepFM en Food.com domina en precisión y novedad, mientras que LightFM enriquecido en MealRecPlus equilibra mejor salud y

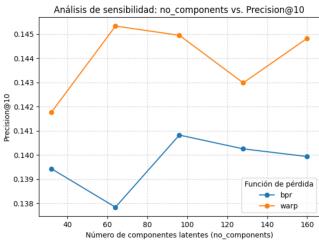


Figure 2. Precision@10 vs. no_components en LightFM (Meal-RecPlus), comparando warp y bpr.

relevancia.

- Novedad y diversidad: DeepFM mantiene altos valores de Nov@10 y Div@10; LightFM requiere diversificación adicional.
- Personalización dietética: la estrategia híbrida por tags muestra potencial, pendiente de métricas específicas.

6. Conclusiones

En este trabajo se compararon distintos enfoques de sistemas recomendadores de recetas sobre dos grandes colecciones de datos (dataset de Food.com y MealRecPlus), evaluando métodos clásicos de filtrado colaborativo como LightFM y modelos basados en aprendizaje profundo como DeepFM, utilizando una estrategia híbrida que combina tags, keywords, categorías y otros metadatos de salud para lograr crear un sistema recomendador balanceado. En base a nuestros resultados, podemos concluir:

- Desempeño global: DeepFM alcanza una Precision@10 de 0.9430 y una Recall@10 de 0.5705 en Food.com, superando con creces a los baselines y a las versiones de LightFM, lo que evidencia la capacidad de los modelos de *deep learning* para capturar interacciones complejas entre usuarios, ítems y metadatos, que los modelos clásicos como LightFM no captan muy bien.
- Saludabilidad y relevancia: En MealRecPlus, enriquecer LightFM con buckets nutricionales (FSA/WHO) eleva la *Precision@10* de 0.0954 a 0.1372, y al incorporar características de contenido

Modelo	Nombre de la Receta	Categoría / Tags
	Granny's Greek Fried Potatoes	Potato
DeepFM	Seasoned Goldfish Crackers	Toddler Friendly
	Fried Yellow Squash	Vegetable
	Polenta Bacon Goat Cheese Sauce	Sauces
	Banana Bread w/Chocolate Chips	Quick Breads
	Peanut Butter Cup Cookies	Dessert
LightFM (sin tags)	Chicken Pot Pie IX	North American
	Chewy Choc. Oatmeal Cookies	Dessert
	World's Best Lasagna	Pasta
	Chicken Enchiladas II	Mexican
	Spinach Artichoke Dip II Banana Crumb Muffins	vegetarian, dip vegetarian, muffins
LightFM (vegetarian)	Old Fashioned Pancakes	vegetarian, breakfas
	Lentil Soup	vegetarian, healthy
	Black Bean Veggie Burgers	vegetarian, burgers

Table 3. Top-5 recetas recomendadas para un usuario según el modelo utilizado.

(TF-IDF + SVD) alcanza 0.1453, demostrando que la integración de señales de salud y semántica de ingredientes mejora la utilidad de las recomendaciones.

- Novedad y diversidad: DeepFM mantiene altos valores de Nov@10 (15.16) y Div@10 (1.00) en Food.com, mientras que LightFM, pese a beneficiarse de la información nutricional, muestra aún baja diversidad intra-lista, indicando la necesidad de mecanismos adicionales de diversificación para modelos clásicos, por lo que reduce su utilidad en la práctica frente a modelos de aprendizaje profundo.
- Personalización dietética: La estrategia híbrida por tags permite guiar el ranking según preferencias dietéticas (por ejemplo "vegetariano"), logrando recomendaciones coherentes con las etiquetas seleccionadas.

No obstante, este trabajo presenta ciertas limitaciones:

- El análisis cuantitativo no incluyó pruebas de significancia estadística que confirmen la robustez de las diferencias observadas.
- La evaluación se basó únicamente en métricas offline y desactualizadas. Además, faltarían estudios con usuarios reales para validar la satisfacción y aceptación práctica.
- Los modelos más complejos (DeepFM) requieren mayor tiempo de entrenamiento y recursos computacionales, lo cual podría limitar su adopción en la

Tag aplicado	Nombre de la Receta	Categoría / Tags
gluten-free	Flourless Peanut Butter Cookies Guacamole Spinach Caprese Salad Lemon Garlic Tilapia White Chili I Mandarin Almond Salad Roasted Pumpkin Seeds Turkey Sloppy Joes 3-Ingredient PB Cookies Fresh Pear Pie	snacks, dietary, gluten-free dip, north-american, gluten-free salads, dietary, gluten-free seafood, main-dish, gluten-free chili, poultry, gluten-free salad, gluten-free, easy snacks, gluten-free main-dish, gluten-free dessert, easy, gluten-free dessert, fruit, gluten-free
high-protein	Sticky Spicy Chicken Tuna Salad Brown Sugar Pork Chops Salisbury Steak Famous Pork Chops Pasta Chicken Salad Unbelievable Chicken Rotisserie-Style Chicken Key West Chicken Baked Pork Chops I	poultry, main-dish, high-protein seafood, lunch, high-protein pork, easy, high-protein beef, dinner, high-protein pork, 5-ingredients, high-protein salad, chicken, high-protein oven, poultry, high-protein main-dish, high-protein poultry, grilled, high-protein pork, baked, high-protein
healthy	Potato Salad Pizza Dough I Cranberry Sauce Honey Chicken Kabobs Easy Pumpkin Muffins Chicken and Dumplings Honey Wheat Bread Cranberry Spinach Salad Southwestern Egg Rolls Lemon Blueberry Bread	salad, low-protein, healthy low-fat, healthy, easy condiments, low-calorie, healthy poultry, grill, healthy vegetarian, dessert, healthy main-dish, healthy bread, low-cholesterol, healthy salad, healthy, low-fat healthy, main-dish, baked bread, dessert, healthy
low-fat	Pizza Dough I Cranberry Sauce Spinach Cranberry Salad Corn Black Bean Salad II Slow Cooker Pulled Pork Chicken Tortilla Soup Orange Cream Fruit Salad Salsa Chicken Buttery Soft Pretzels Lentil Soup	main-dish, beginner-cook, low-fat condiment, low-fat, low-sodium salad, low-fat, healthy salad, low-fat, easy crockpot, low-fat, simple soup, low-fat, slow-cooker fruit, salad, low-fat poultry, low-fat, quick bread, snack, low-fat soup, low-fat, vegetarian

Table 4. Top-10 recetas recomendadas por LightFM al aplicar diferentes filtros de tags nutricionales.

práctica o al menos requeriría de optimizaciones adicionales, como transformaciones de datos o selección previa a su entrenamiento.

• Falta evaluar más modelos avanzados como (DeepFM) para identificar mejores alternativas (quizás más rápidas de entrenar).

Como propuestas de trabajo futuro, se podría:

- Incorporar nuevas métricas para medir mejor la diversidad de las recetas (y cobertura de las menos populares) e identificar con una escala razonable si las recetas efectivamente pertenecen a las preferencias dietéticas del usuario (esto de la mano con datasets especializados o curados para esto).
- Desarrollar un protocolo de evaluación con usuarios finales que permita medir efectivamente si las recomendaciones son útiles y balanceadas para los usuarios.

- Explorar distintas técnicas más óptimas para encontrar los mejores hiperparámetros del modelo en función de cada perfil dietético.
- Extender las estrategias híbridas para mejorar el ajuste a las preferencias nutricionales de los usuarios.

En conjunto, nuestros resultados muestran que la combinación de señales colaborativas, semánticas y nutricionales, junto con mecanismos de personalización explícita, permite identificar mejoras para construir sistemas recomendadores de recetas tanto precisos como alineados con objetivos de salud y hábitos de los usuarios.

Referencias

Food, of the United Nations, A. O., for Agricultural Development, I. F., Fund, U. N. C., Programme, W. F., and Organization, W. H. The state of food security and nutrition in the world 2019: Safeguarding against economic slowdowns and downturns. Technical report, Food and Agriculture Organization of the United Nations, Rome, 2019. URL https://openknowledge.fao.org/items/f73edb75-4017-497b-8ca3-e66ea2109560. CC BY-NC-SA 3.0 IGO.

James, S. L., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., others, and Abdelalim, A. Global, regional, and national incidence, prevalence, and years lived with disability for 354 diseases and injuries for 195 countries and territories, 1990–2017: A systematic analysis for the global burden of disease study 2017. *The Lancet*, 392:1789–1858, 2018. URL https://www.healthdata.org/sites/default/files/files/policy_report/2019/GBD_2017_Booklet.pdf.

Li, D., Chen, C., Gong, Z., Lu, T., Chu, S. M., and Gu, N. Collaborative filtering with noisy ratings. In *Proceedings of the 2019 SIAM International Conference on Data Mining*, pp. 747–755. SIAM, 2019.

Rostami, M., Farrahi, V., Ahmadian, S., Jalali, S. M. J., and Oussalah, M. A novel healthy and time-aware food recommender system using attributed community detection. *Expert Systems with Applications*, 221:119719, 2023. doi: 10.1016/j.eswa.2023.119719. URL https://doi.org/10.1016/j.eswa.2023.119719.

World Health Organization. Global status report on noncommunicable diseases 2014. Technical Report WHO/NMH/NVI/15.1, World Health Organization, October 2014. URL https://www.who.int/publications/i/item/9789241564854.

A. Repositorio

El código para replicar los resultados de este trabajo se encuentra en este repositorio de GitHub

B. Tablas anexadas

Modelo	ID Receta	Nombre	Categoría / Tags	Rating	Score
DeepFM	428343	Granny's Greek Fried Potatoes	Potato	5.0	1.0000
	148446	Seasoned Goldfish Crackers	Toddler Friendly	5.0	1.0000
	317234	Fried Yellow Squash	Vegetable	5.0	1.0000
	524958	Polenta Crusted Bacon Goat Cheese Sauce	Sauces	5.0	1.0000
	19522	Banana Bread W/Chocolate Chips	Quick Breads	5.0	1.0000
	302801	Shishlik (Israeli Skewers)	Veal	5.0	0.9999
	285831	Whole Wheat Dinner Rolls	Yeast Breads	5.0	0.9999
	494438	Athenian Ranch Pork Empanadas #RSC	Weeknight	5.0	0.9999
	173712	Olive Garden Alfredo Sauce	Sauces	5.0	0.9999
	194158	Mama's Fry Bakes	Yeast Breads	5.0	0.9999
	9471	Peanut Butter Cup Cookies	Dessert	_	_
	17652	Banana Crumb Muffins	Breads, Vegetarian	_	-
	26317	Chicken Pot Pie IX	Savory Pies	-	-
	25037	Big Chewy Chocolate Chip Cookie	Dessert	-	-
LightFM (sin tags)	6865	To Die For Blueberry Muffins	Breads	_	-
	24445	Chocolate Oatmeal Cookies	Dessert	_	-
	16066	Slow Cooker Pot Roast	Beef	_	-
	23600	World's Best Lasagna	Pasta	-	-
	20144	Banana Banana Bread	Breads	_	-
	8694	Chicken Enchiladas II	Mexican	-	-
LightFM + Tag 'vegetarian'	17652	Banana Crumb Muffins	vegetarian, muffins	5.0	0.9352
	14231	Guacamole	vegetarian, dip	-	0.9819
	26819	Spinach Artichoke Dip II	vegetarian, dip	_	0.6714
	21014	Old Fashioned Pancakes	vegetarian, breakfast	_	0.8155
	20446	Cilantro Cream Sauce	vegetarian, sauce	-	0.6853
	13045	Broccoli Cheese Soup	vegetarian, soup	_	0.6623
	7958	Pumpkin Choc Chip Muffins	vegetarian, muffins	-	0.7586
	13978	Lentil Soup	vegetarian, healthy	-	0.6800
	85452	Black Bean Veggie Burgers	vegetarian, burgers	-	0.5757
	14144	Bean Quesadillas	vegetarian, mexican	-	0.3138

Table 5. Recomendaciones detalladas por modelo para un usuario seleccionado. Se incluyen categorías/tags y scores de relevancia.