



Evaluación de modelos de recomendación de recetas en base a filtrado colaborativo, metadatos y métricas de salud



Grupo 7 - Alvaro Romero, Sofía Rebolledo y Francisco Arenas Pontificia Universidad Católica de Chile

Motivación

Los **sistemas de recomendación de recetas** deberían considerar más factores que solo la popularidad a la hora de recomendar platos a sus usuarios. En particular, deberían aportar al bienestar individual, promoviendo hábitos alimentarios saludables y adaptándose a necesidades personales como alergias, dietas y creencias. Nuestra propuesta busca mejorar la experiencia del usuario mediante un enfoque híbrido.

Problema

Los sistemas recomendadores de recetas buscan, por lo general, sugerir preparaciones según las preferencias de sus usuarios. Sin embargo, las soluciones clásicas basadas en filtrado colaborativo enfrentan ciertas limitaciones clave:

- Falta de criterios nutricionales y de salud**
Las recomendaciones priorizan la popularidad de las recetas por sobre el valor nutricional de estas, lo que lleva a sugerencias poco saludables para los usuarios.
 - Preferencias estáticas**
Los sistemas suponen que los gustos de los usuarios son estáticos en el tiempo, sin adaptarse a los cambios que podrían darse por objetivos físicos, etapas de vida o nuevos hábitos de estos.
 - Desatención de restricciones personales**
Estas soluciones ignoran alergias, dietas especiales (por ejemplo, sin gluten, vegana, etc.) y preferencias éticas o culturales, afectando así a la relevancia y seguridad misma de estas recomendaciones (Rostami et al., 2023).
- Además, la sobrecarga de opciones en plataformas con miles de recetas fuerza al usuario a filtrar manualmente, perdiendo tiempo y recibiendo recomendaciones poco útiles (Li et al., 2019).

Metodología

Datos y metadatos utilizados

Se trabajó con dos *datasets*:

- **Food.com – Recipes and Reviews:**

Se extrajeron *keywords* relacionadas con ingredientes y restricciones alimentarias (ej. “vegano”, “sin gluten”, “halal”), junto a metadatos nutricionales como calorías, grasas, proteínas, colesterol, etc... También, se incorporó en algunos experimentos algunos metadatos adicionales proporcionados por el *dataset*, como las categorías.

- **MealRecPlus:**

Se utilizaron scores nutricionales de salud (FSA y OMS), transformados en *buckets* que indican cuando un alimento es saludable o no, en distintas escalas. También, se consideraron *tags* generales de la receta, para comparar distintos enfoques de modelado.

Modelos y entrenamiento

Se entrenaron dos tipos de modelos:

- **KNN:**

Aplicado al *dataset* de *Food.com* usando *scikit-learn*, con representación basada en las características diseñadas anteriormente.

- **LightFM:**

Modelo híbrido colaborativo que incluye metadatos. Se entrenó sobre ambos *datasets* y se exploró automáticamente el espacio de hiperparámetros para optimizar su rendimiento.

Los modelos fueron evaluados usando Precision@10 y Recall@10, comparando distintas combinaciones de *features* e hiperparámetros.

Resultados

Dataset food.com

Método	Precision@10	Recall@10
KNN*	0.0024	0.0010
LIGHTFM*	0.0103	0.0219
HAFR	0.0629	0.0612
CFRR	0.0624	0.0608
FGCN	0.0631	0.0615
HTFRS	0.0654	0.0637

Dataset MealRecPlus

Método	Precision@10	Recall@10
LIGHTFM*	0.1403	0.0760

Discusión

- Comparación con el estado del arte: Nuestro LightFM con metadatos FSA/OMS logra $Precision@10 = 0.1403$ en *MealRecPlus*, más del doble que el HTFRS (0.0654) de Zhang et al. (2023), demostrando el claro beneficio de usar información oficial de salud.
- Impacto de los metadatos nutricionales: Al incorporar scores de la FSA y la OMS, el modelo alcanza $Precision@10 = 0.14$, frente a 0.0024 del KNN y 0.0103 del LightFM sin metadata, lo que confirma que dichos atributos mejoran significativamente la relevancia de las recomendaciones.
- Limitaciones del KNN y del LightFM “puro”: Ambos métodos en Food.com registraron rendimientos muy bajos debido al ruido en ingredientes y etiquetas inconsistentes, así como a la falta de criterios nutricionales en su entrenamiento.

Conclusión

El enfoque híbrido desarrollado cumple con éxito los tres objetivos iniciales:

- Personalización: al combinar filtrado colaborativo y basado en etiquetas, las recomendaciones se ajustan estrechamente a las preferencias de cada usuario.
- Calidad nutricional: la integración de metadatos de la FSA y la OMS asegura que todas las recetas sugeridas satisfacen criterios de salud validados.
- Relevancia y precisión: con una $Precision@10$ de 0.14, nuestro modelo LightFM enriquecido supera ampliamente a los métodos clásicos, priorizando opciones más adecuadas.
- En conjunto, este sistema ofrece sugerencias de recetas más personalizadas, saludables y pertinentes que las alternativas tradicionales.

Referencias

- [1] irkaal. (s.f.). Food.com – Recipes and reviews [Conjunto de datos]. Kaggle. Recuperado el 2 de julio de 2025, de <https://www.kaggle.com/datasets/irkaal/foodcom-recipes-and-reviews/data>
- [2] Li, D., Chen, C., Gong, Z., Lu, T., Chu, S. M., & Gu, N. (2019). Collaborative filtering with noisy ratings. En *Proceedings of the 2019 SIAM International Conference on Data Mining* (pp. 747–755). SIAM. <https://doi.org/10.1137/1.9781611975673.84>
- [3] Li, M., Li, L., Tao, X., & Huang, J. X. (2024). MealRec+: A meal recommendation dataset with meal-course affiliation for personalization and healthiness. En *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24)* (pp. 564–574). ACM. <https://doi.org/10.1145/3626772.3657857>
- [4] Rostami, M., Farrahi, V., Ahmadian, S., Jalali, S. M. J., & Oussalah, M. (2023). A novel healthy and time-aware food recommender system using attributed community detection. *Expert Systems with Applications*, 221*, 119719. <https://doi.org/10.1016/j.eswa.2023.119719>