

# Recomendación de Videojuegos : Análisis comparativo de modelos

René Saavedra, Geraldine Coli Acevedo y Carlos Olguín

Pontificia Universidad Católica de Chile

## Abstract

En este artículo estudiamos la generación de recomendaciones de videojuegos basadas principalmente en el uso del contexto usuario-ítem. Para esto, se plantea el uso de diversos modelos que aprovechan información subyacente de ítems y usuarios asociados (como pueden ser géneros de ítems, preferencias del usuario y demás). Para ello, se estudia un dataset Game Recommendations on Steam, el cual contiene diversa información que servirá para entrenar los diversos modelos. El objetivo de esto es analizar el rendimiento de estos modelos en busca de una calidad y diversidad en las recomendaciones. Los detalles en torno a los modelos y la metodología asociada se encuentran en su respectiva sección.

## Introducción

En los sistemas de recomendación, se utilizan diversas técnicas computacionales y matemáticas para aprovechar los datos disponibles y generar recomendaciones variadas y de calidad. Una estrategia clave es incorporar contexto (del usuario o del ítem), que permite mejorar los modelos al extraer información subyacente. Sin embargo, esto requiere transformar los datos a un formato adecuado y resolver problemas como el cold-start. Para abordar el cold-start, Rahman et al. (2022) proponen usar contexto del usuario (datos demográficos, situación de recomendación, etc.) junto con historiales previos de interacciones (mínimo cierta cantidad de ratings). Además, para extraer información semántica de atributos categóricos, se aplican técnicas como:

- Codificación binaria (one-hot) de categorías en dimensiones separadas.
- Métodos basados en semántica (TF-IDF, Cosine Similarity), como sugiere Yuen et al. (2023).

Para esto se proponen modelos de este tipo tal que:

- CNN-Graph Based: Combina redes neuronales convolucionales (CNN) con estructura de grafos para modelar relaciones complejas entre usuarios e ítems, capturando tanto características y patrones detrás de los datos.
- LightFM: Modelo híbrido que integra factorización matricial con embeddings de usuarios/ítems, permitiendo incluir features contextuales para mejorar recomendaciones en escenarios sparse.

Esta variedad permitirá identificar qué combinación de técnicas y modelos ofrece mejores resultados para distintos tipos de datos y requisitos del sistema de recomendación.

## Dataset

Ya mencionado, el dataset ocupado finalmente fue el de Game Recommendations on Steam, el cual contiene información de usuarios y videojuegos en conjunto con su metadata, es decir, información adicional que sirve para hacer cross y extraer contexto.

En específico, el dataset involucra diversos subdatasets que servirán para realizar cruces de información en busca de diversa metadata asociada para usuario e ítem, a continuación se presentan los distintos conjuntos de datos a emplear:

games (50.872×13)		users (14.306.064×3)	
app_id: int64	title: object	user_id: int64	products: int64
date_release: object	win: bool	reviews: int64	
mac: bool	linux: bool		
rating: object	positive_ratio: int64		
user_reviews: int64	price_final: float64		
price_original: float64	discount: float64		
steam_deck: bool			

Table: Tablas de juegos y usuarios

recommendations (41.154.794×8)		games_metadata (50.872×3)	
app_id: int64	helpful: int64	app_id: int64	description: object
funny: int64	date: object	tags: object	
is_recommended: bool	hours: float64		
user_id: int64	review_id: int64		

Table: Tablas de recomendaciones y metadatos

Dado el tamaño del dataset combinado (usuarios, recomendaciones, ítems y metadata de videojuegos), se opta por tomar una muestra aleatoria para realizar testeos adecuados. Además, para abordar mejor el problema de cold-start, se fija  $MIN\_REVIEWS\_USER = 150$  y  $MIN\_REVIEWS\_ITEM = 200$ .

## Metricas

Para evaluar los distintos modelos, se proponen diversas métricas que miden la calidad y variedad de las recomendaciones. De esta forma, se define:

- NDCG (Normalized DCG):**

$$NDCG@K = \frac{DCG@K}{IDCG@K}$$

con IDCG siendo el DCG ideal (ordenamiento óptimo).

- MAP (Mean Average Precision):**

$$MAP = \frac{1}{|U|} \sum_{u=1}^{|U|} \frac{1}{|T_u|} \sum_{k=1}^K P_u@k \cdot rel_u(k)$$

- Diversidad:**

$$\text{Diversidad} = 1 - \frac{1}{|U| \cdot K \cdot (K-1)} \sum_{u \in U} \sum_{i,j \in R_u, i \neq j} \text{sim}(i, j)$$

donde  $\text{sim}(i, j)$  es la similitud entre ítems  $i$  y  $j$ .

- Novedad:**

$$\text{Novedad} = \frac{1}{|U| \cdot K} \sum_{u \in U} \sum_{i \in R_u} \text{pop}(i)$$

con  $\text{pop}(i) = -\log_2\left(\frac{\#\text{usuarios que interactuaron con } i}{\#\text{usuarios}}\right)$

## Metodología

Se compararon cuatro métodos de recomendación sobre particiones del 10% al 100% del dataset:

- LightFM**, modelo de factorización de matrices híbrido que combina señales colaborativas y de contenido, permitiendo incorporar atributos de usuarios e ítems para mejorar la precisión.
- LightGCN**, red de convolución en grafos simplificada que difunde representaciones latentes de usuarios e ítems a través de varias capas, capturando relaciones de vecindad en la estructura de interacción sin el uso de funciones de activación innecesarias.
- Baselines**, “Most Popular” recomienda los ítems con mayor número de interacciones globales y “Random” elige ítems al azar como referencia de rendimiento mínimo.

## Resultados

Evaluamos el modelo **LightGCN** con múltiples combinaciones de *dimensión*, *número de capas* y *learning rate*, manteniendo constante el número de **lotes** en 50, ya que no presentó variabilidad significativa.

Modelo	NDCG	MAP	Diversidad	Novedad
D256_L5_LR0.01	<b>0.7802</b>	<b>0.7098</b>	0.9520	10.7191
D64_L2_LR0.001	0.2988	0.2502	0.9073	<b>11.2608</b>
D128_L3_LR0.005	0.6007	0.5235	0.9410	10.9912
D256_L2_LR0.005	0.6640	0.5961	<b>0.9645</b>	10.8838
D128_L5_LR0.001	0.5563	0.4762	0.9420	11.2620

Table: Comparación de métricas entre cinco configuraciones del modelo LightGCN.

Por otro lado, el modelo **LightFM** no logró entregar resultados competitivos, a pesar de las distintas combinaciones de parámetros y modificaciones probadas.

Modelo	NDCG	MAP	Diversidad	Novedad
Popular	0.051	0.034	0.894	<b>11.872</b>
Random	0.005	0.004	<b>0.994</b>	10.720
LightGCN (D256_L5_LR0.01)	<b>0.7802</b>	<b>0.7098</b>	0.9520	10.7191
LightFM	0.0055	0.0027	0.0493	1.0493

Table: Comparación de métricas entre modelos recomendadores.

## Análisis

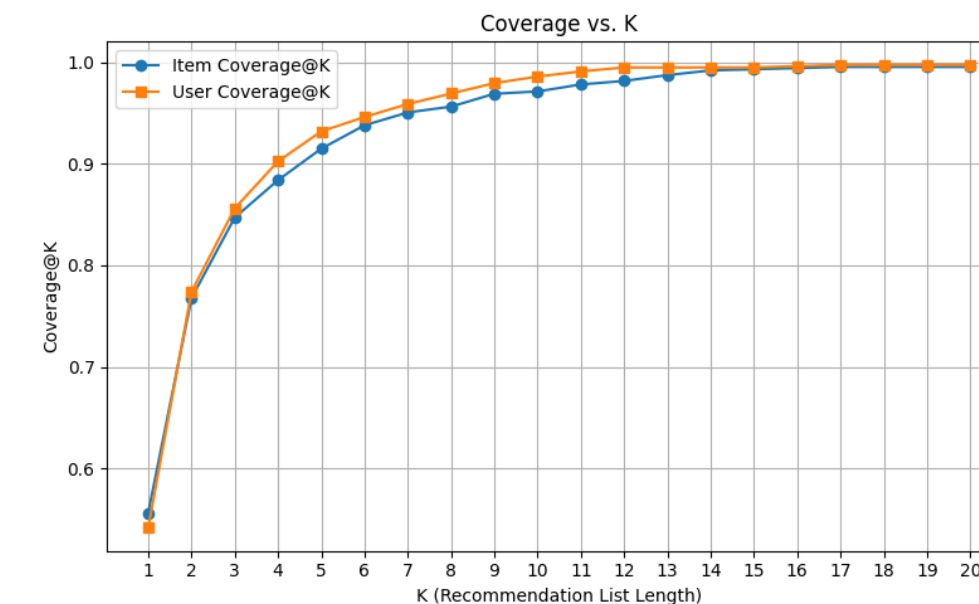


Figure: Item & User Coverage vs. K

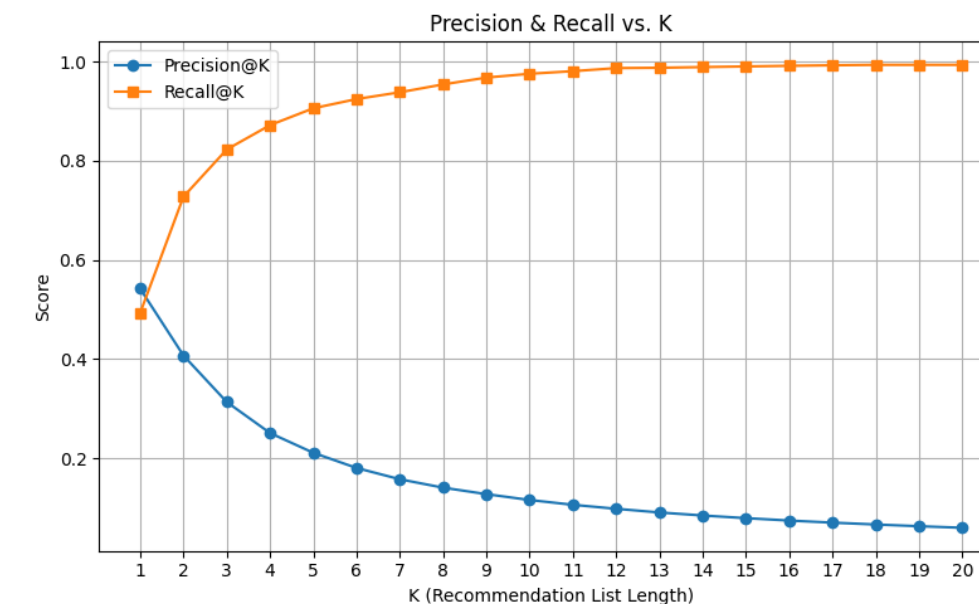


Figure: Precision@K & Recall@K vs. K

Utilizando el modelo con mejores resultados de LightGCN, se evaluó cómo varían la cobertura, así como la precisión y el recall, a medida que aumenta el tamaño de la lista de recomendaciones ( $K$ ).

## Conclusiones

- LightFM muestra un desempeño insuficiente, probablemente por la alta sparsidad y complejidad del dominio.
- La arquitectura de LightGCN de convolución en grafos captura mejor las relaciones colaborativas, logrando los mejores resultados, superando a una recomendación popular.
- El cold-start es un gran obstáculo dentro de la recomendación. Usuarios o ítems nuevos requieren señales adicionales (metadatos, enfoque híbrido) para superar el *popular baseline*.

## Referencias

Guy Shani and Asela Gunawardana. *Evaluating Recommendation Systems*. En: Ricci, F., Rokach, L., Shapira, B., Kantor, P. (eds) *Recommender Systems Handbook*. Springer, Boston, MA, 2011, pp. 257–297. DOI: [https://doi.org/10.1007/978-0-387-85820-3\\_8](https://doi.org/10.1007/978-0-387-85820-3_8)

Md Rahman, Wei Liu, and Hao Peng. *A Survey of Cold-Start Problems in Recommender Systems: Challenges, Strategies, and Trends*. Expert Systems with Applications, vol.198, 2022, 116834 DOI: <https://doi.org/10.1016/j.eswa.2022.116834>

Xiang Wang, Xiangnan He, Yixin Cao, Meng Liu, and Tat-Seng Chua. *LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation*. Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '20), 2020, pp. 639–648. DOI: <https://doi.org/10.1145/3397271.3401063>

Yehuda Koren, Robert Bell, and Chris Volinsky. *Matrix Factorization Techniques for Recommender Systems*. IEEE Computer, vol.42, no.8, 2009, pp.30–37. DOI: <https://doi.org/10.1109/MC.2009.263>