



User Diversity: Una Métrica para Diversidad en Recomendaciones

Fernanda Pérez Hargreaves, Mathias Madsen Sánchez, Felipe Olivares Labarca, Beatriz Errázuriz Camus



Motivación

- Los sistemas de recomendación se enfocan en el comportamiento pasado del usuario.
- Priorizan categorías ya exploradas, dejando de lado áreas relevantes poco desarrolladas.
- User Diversity** fomenta la exploración de nuevas áreas, equilibrando **Personalización** y **Diversidad**
- Ofrece recomendaciones variadas que enriquecen la experiencia del usuario y fomentan el descubrimiento de nuevos intereses.

Objetivo

Objetivo Principal: Desarrollar una métrica que evalúe la diversidad en las recomendaciones, promoviendo la exploración de áreas menos exploradas por el usuario.

Objetivos Específicos:

- Definir formalmente la métrica **User Diversity**.
- Implementarla en datasets como **MovieLens**, **Last.fm** y **Yelp** (data de restaurantes).
- Compararla con métricas existentes para identificar **ventajas y limitaciones**.

Dataset

- MovieLens:** Calificaciones de usuarios sobre películas.
- Last.fm y Music-dataset-1950-to-2019:** Información sobre preferencias musicales de usuarios y metadatos de canciones (géneros y características auditivas).
- Yelp (restaurantes):** Datos de negocios y reseñas, enfocados en categorías gastronómicas.

Conjunto	Nº Usuarios	Nº Items	Géneros
Música	11	3037	7
Películas	138,493	22,854	20
Restaurantes	112,330	49,362	12

Modelos

- Baselines:** **Random** (recomendaciones aleatorias) y **Most Popular** (ítems más populares).
- Modelos avanzados:** **iKNN** (similitud entre ítems), **FastFM** (factorización de máquinas), **DeepFM** (redes profundas y factorización), y **Hybrid** (precisión y diversidad).

Métricas

- De rendimiento:** MAP, NDCG@10, Precision@10 y Recall@10
- De diversidad:** Entropía de Shannon, Long Tail, Intra-List Diversity (ILD), Diversity Coverage, Inverse Propensity Score (IPS)

Fórmula de User Diversity

$$UD = 1 - \frac{\left| \sum_{j=1}^k \left(\frac{R_j}{R} \right) \log \left(\frac{R_j}{R} \right) \right|}{\log(k)}$$

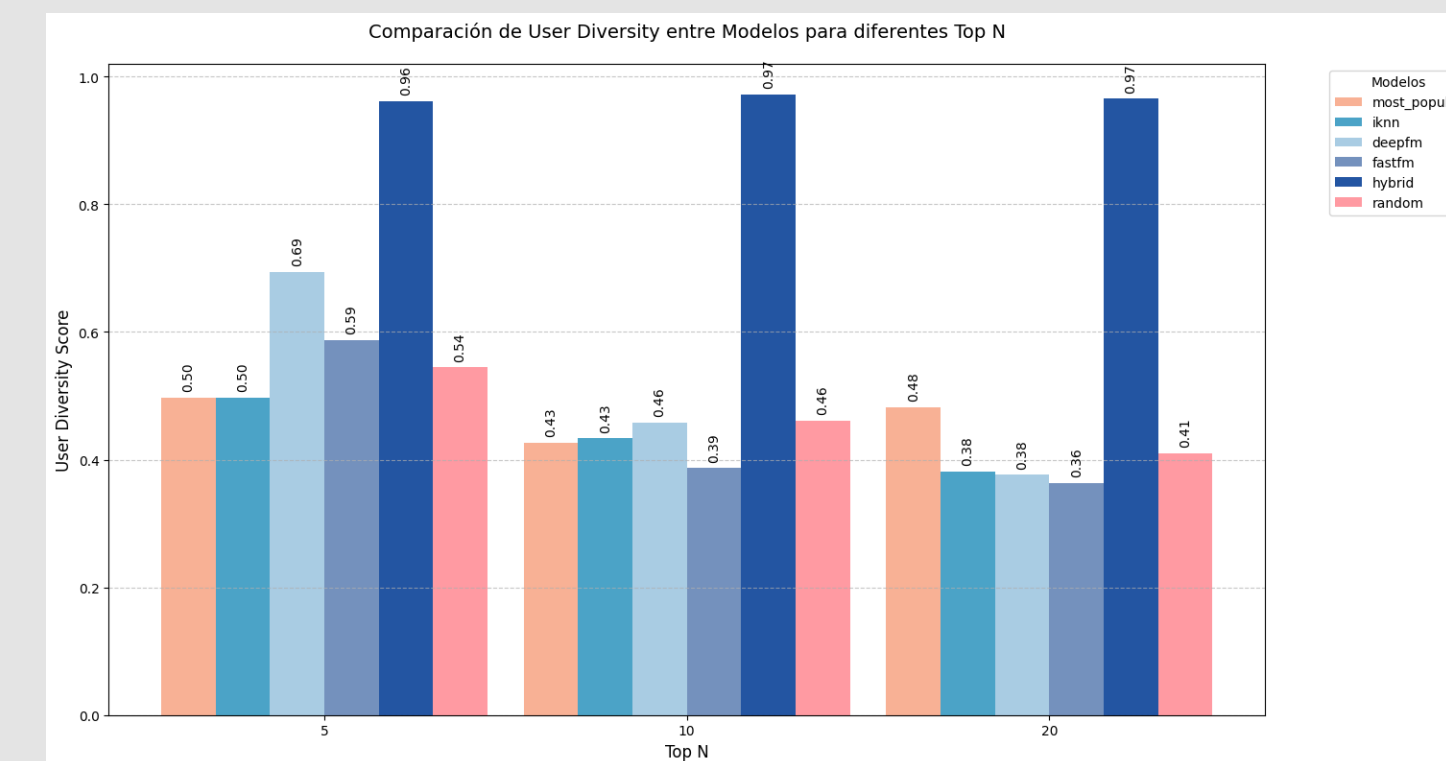
Donde:

- R_j : Número de recomendaciones en la categoría j .
- R : Total de recomendaciones relevantes para el usuario.
- k : Total de categorías consideradas.
- $\log(k)$: Factor de normalización para garantizar que $UD \in [0, 1]$

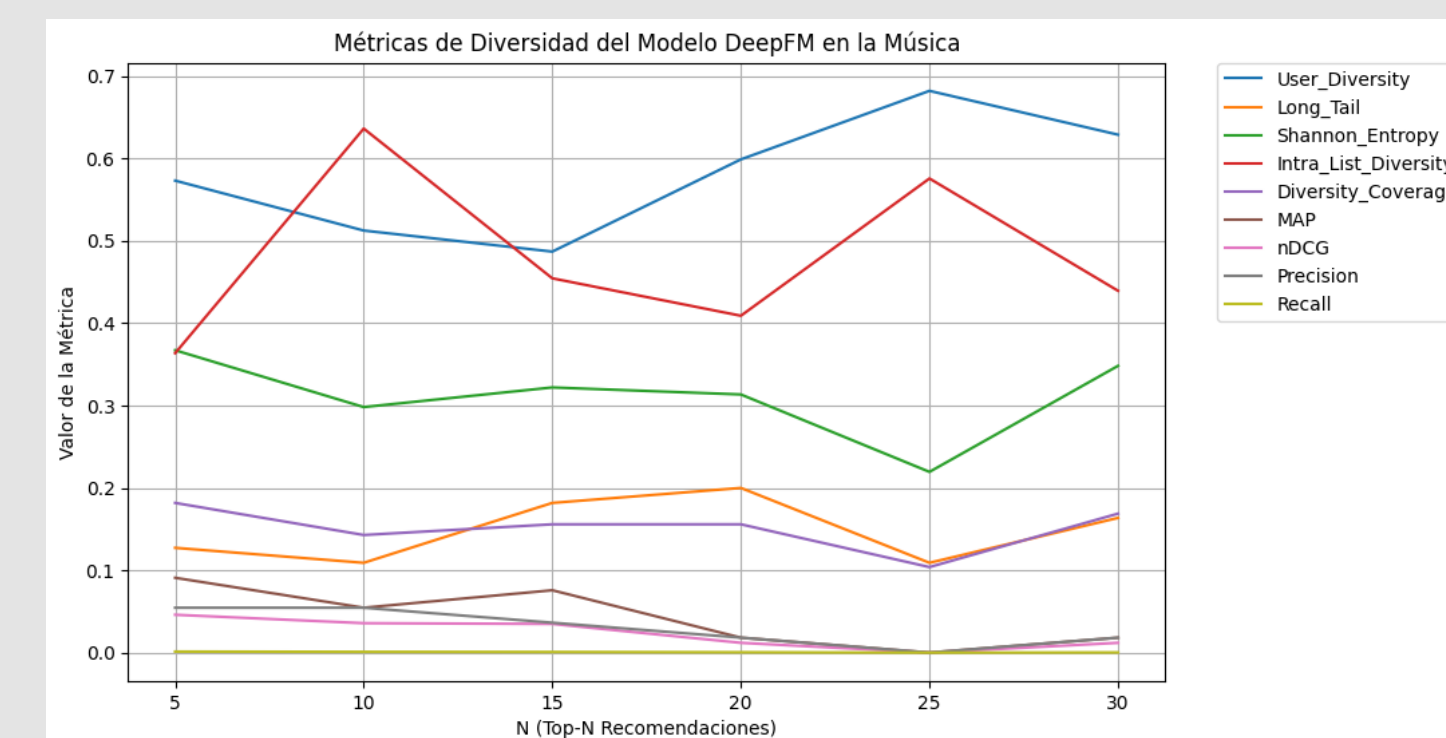
Interpretación:

- $UD = 1$: Máxima diversidad. Las recomendaciones están equilibradas entre las categorías consideradas.
- $UD = 0$: Sin diversidad. Las recomendaciones están concentradas en una única categoría.

Resultados



El modelo **Hybrid** alcanzó los valores más altos de **User Diversity**, pero sacrificó el rendimiento en otras métricas como precisión y recall, lo que lo hace menos práctico para aplicaciones generales. Por otro lado, **DeepFM** mostró un equilibrio entre diversidad y métricas de rendimiento, siendo el modelo más consistente.



- User Diversity** se comportó de manera más robusta que otras métricas como **Long Tail** e **Intra-List Diversity (ILD)**, al combinar personalización y variedad en las recomendaciones.
- Shannon Entropy** y **Diversity Coverage** también reflejan tendencias similares, pero no ajustan tan bien la diversidad personalizada como lo hace User Diversity.
- User Diversity frente a métricas tradicionales:** **Long Tail** fomenta la inclusión de ítems menos populares, pero no considera relevancia personalizada. **Shannon Entropy** mide dispersión general, pero no distingue entre categorías relevantes y no relevantes. **Intra-List Diversity (ILD)** analiza la variedad dentro de las listas, pero carece de enfoque en preferencias específicas del usuario. **Diversity Coverage** evalúa amplitud de categorías, pero puede generar listas desbalanceadas.

Conclusiones

- Fortalezas de User Diversity:**
 - Combina diversidad y relevancia, adaptándose a las categorías principales del usuario.
 - Normaliza resultados para diferentes valores de k , ofreciendo un análisis robusto.
 - Supera métricas tradicionales al integrar personalización explícita con diversidad.
- Limitaciones de la métrica:**
 - Depende de datos detallados del usuario, lo que dificulta su uso en escenarios con poca información.
 - Requiere mayor poder computacional que las métricas tradicionales.
 - Enfocada en categorías conocidas, puede limitar la exploración de nuevas áreas
- Conclusión general:** **User Diversity** es una métrica innovadora que equilibra diversidad y personalización, destacándose como una herramienta poderosa para enriquecer la experiencia del usuario en sistemas de recomendación.

Dificultades

- Datos de Last.fm sin rankings, lo que requirió estrategias personalizadas.
- Incompatibilidad con pyRecLab, obligando a desarrollar herramientas desde cero.
- Limitaciones de memoria en Google Colab para procesar grandes datasets.
- Simplificación de procedimientos y ajustes en el plan original.

Trabajo Futuro

- Estudiar el impacto de diferentes valores de k .
- Analizar cómo top_n afecta el balance entre precisión y diversidad.
- Incorporar pesos personalizados según las preferencias del usuario.

Referencias

- Kunaver, M., & Požrl, T. (2017). *Diversity in recommender systems – A survey*. Knowledge-Based Systems, 123, 154-162. <https://doi.org/10.1016/j.knosys.2017.02.009>
- Adomavicius, G., & Kwon, Y. (2012). Improving aggregate recommendation diversity using ranking-based techniques. *IEEE Transactions on Knowledge and Data Engineering*.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*.