



Zero-Shot Next-Item Recommendation using Large Pretrained Language Models

IIC3633 - Sistemas Recomendadores

Jairo Navarro - Begoña Pendas - Fabián Riveros

Noviembre 2023

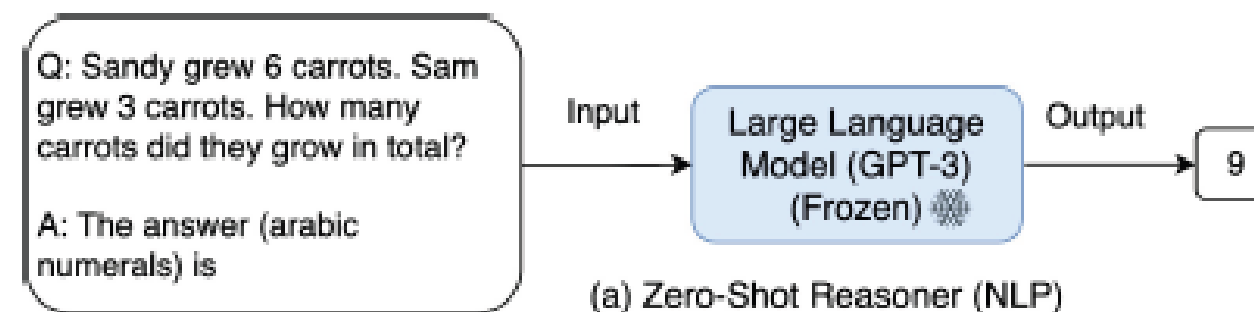
Outline

- **Contexto, Problema y Contribución**
- Marco teórico y Estado del Arte
- Solución: Zero-Shot Next Item Recommendation
- Evaluación, Conclusiones y Críticas

Contexto

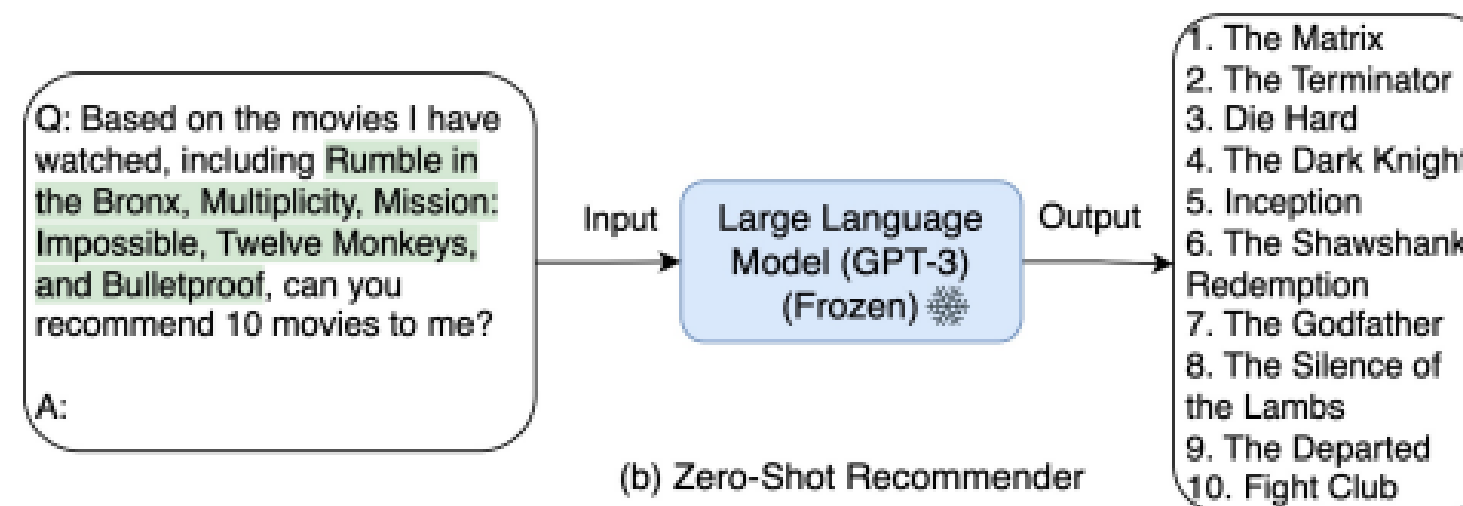
Los **Large Language Models (LLM)** han sido capaces de desempeñarse bien en tareas relacionadas a **Natural Language Processing (NLP)** por medio de *prompts* ajustado en *zero-shot*.

El uso de LLM como recomendador de *next item recommendation* ajustado en *zero-shot* está en desarrollo



Problema

El problema del NLP es que requieren de entrenamiento con datos de interacción de usuario con ítems, lo cual precisa conocer el universo de ítems y el método de recomendación. Esto limita su capacidad para realizar recomendaciones de Zero-Shot, lo cual se ve reflejado en el **bajo desempeño esperado** de estas.



Contribución

Se propone un acercamiento al problema de next item recommendation con el uso de LLM: **Zero-Shot Next Item Recommendation (NIR) prompting**, la cual incluye una estrategia de prompting de 3 pasos.

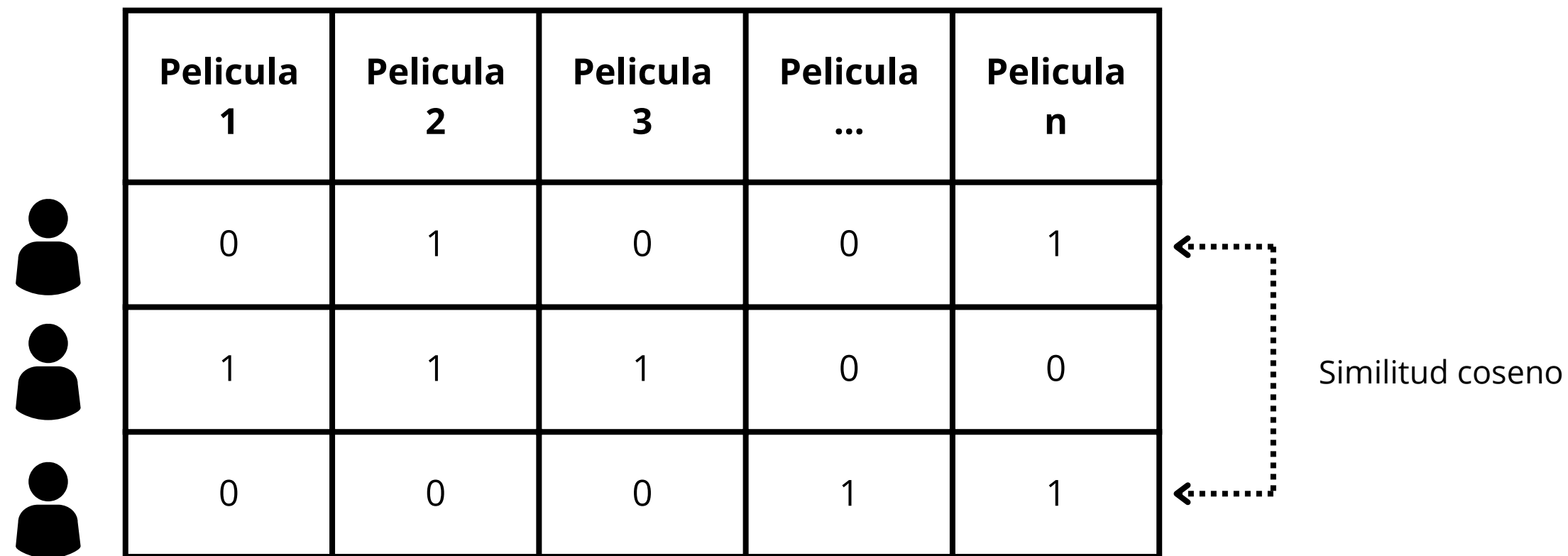
Además, se muestran resultados empíricos por medio de la evaluación usando MovieLens 100K y GPT text-davinci-003




Outline

- Contexto, Problema y Contribución
- **Marco teórico y Estado del Arte**
- Solución: Zero-Shot Next Item Recommendation
- Evaluación, Conclusiones y Críticas

User Filtering

Para el User Filtering, se utiliza la similitud coseno entre el vector de películas vistas por el usuario objetivo y los vectores de otros usuarios para seleccionar los usuarios mas similares.




	Pelicula 1	Pelicula 2	Pelicula 3	Pelicula ...	Pelicula n
	0	1	0	0	1
	1	1	1	0	0
	0	0	0	1	1

Similitud coseno


Item Filtering

Similar a el User Filtering, se utiliza la similitud coseno entre dos películas, se seleccionan una cantidad de películas mas similares para cada película que fue vista por el usuario objetivo.



Pelicula 1	Pelicula 2	Pelicula 3	Pelicula ...	Pelicula n
0	1	0	0	1
1	1	1	0	0
0	0	0	1	1

Similitud coseno



Estado del Arte

Next Item Recommendation

Cadenas de
Markov

Modelos
Neuronales

Data
Augmentation
& Constrative
Learning

Todos estos modelos requieren ser entrenados por las interacciones usuario-ítems. En otras palabras, **no son capaces de realizar recomendaciones en zero-shot**

Estado del Arte

LLM Based Recommendation



The diagram consists of a central orange rectangle labeled 'LLM Based Recommendation'. Below this rectangle are two orange circles. The left circle is labeled 'GPT-2' and the right circle is labeled 'BERT'.

GPT-2

BERT

Se propuso su uso como recomendador principal para next-item recommendation, pero debido a la cantidad de ítems a recomendar y el modelamiento inadecuado de las preferencias del usuario, **no obtuvieron buenos desempeños**

Outline

- Contexto, Problema y Contribución
- Marco teórico y Estado del Arte
- **Solución: Zero-Shot Next Item Recommendation**
- Evaluación, Conclusiones y Críticas

Zero Shot Next Item Recommendation

Objetivo: Recomendar ítems (películas) al usuario por medio de LLM en zero-shot.

Problema encontrado: Los LLM han alcanzado un buen desempeño en tareas del Procesamiento del Lenguaje Natural (NLP) en zero-shot, pero no se ha investigado su uso en recomendaciones en zero-shot.

Solución propuesta: Se propone una estrategia de prompting, Zero Shot Next Item Recommendation (NIR), que guíe al LLM (GPT-3) para realizar recomendaciones de ítems.

Zero Shot Next Item Recommendation

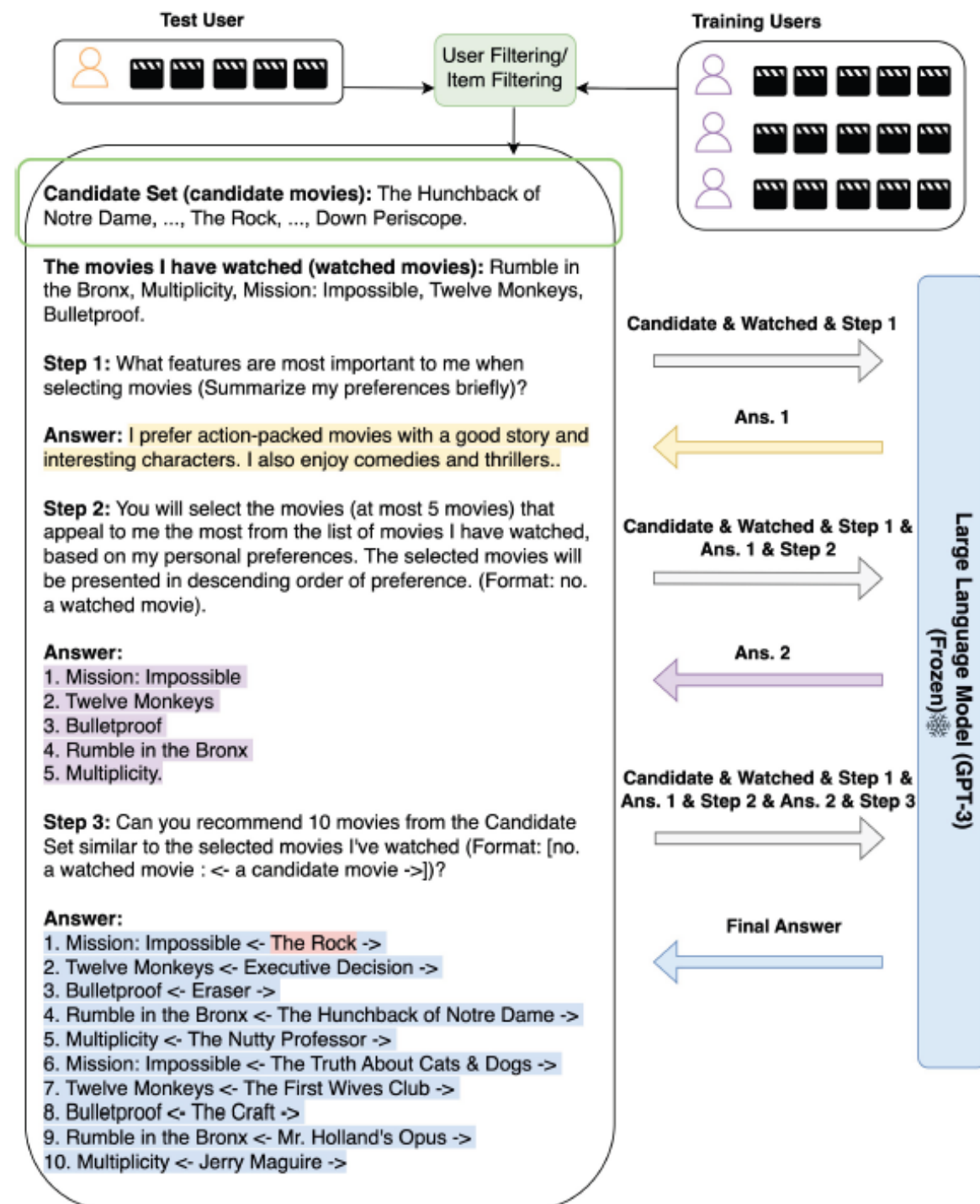
Zero Shot NIR Prompting Strategy:

Estrategia que sigue múltiples pasos que permiten que GPT-3 actúe como recomendador en un ambiente zero-shot.

3 componentes: construcción del set de candidatos, 3 pasos de GPT-prompting y la extracción de respuesta.

Construcción del set de candidatos: se utiliza user-filtering y/o item-filtering para crear un set de candidatos para cada usuario. Este set se utiliza en el siguiente componente.

Zero Shot Next Item Recommendation



3 pasos de GPT prompting

1. Preferencias del usuario: prompt que recibe el set de candidatos y las películas que ha visto el usuario y espera que GPT-3 resuma las preferencias del usuario en base a eso.
2. Películas representativas: prompt combina las preferencias del usuario y solicita que GPT-3 elija películas representativas ordenadas de forma descendente.
3. Recomendación: se integra lo anterior a prompt que pide recomendación de 10 películas del set de candidatos.

Zero Shot Next Item Recommendation

Extracción de respuesta

Se extraen los ítems recomendados de los resultados textuales de los 3 pasos de GPT-3 prompting usando el método simple de extracción basado en reglas.

Outline

- Contexto, Problema y Contribución
- Marco teórico y Estado del Arte
- Solución: Zero-Shot Next Item Recommendation
- **Evaluación, Conclusiones y Críticas**

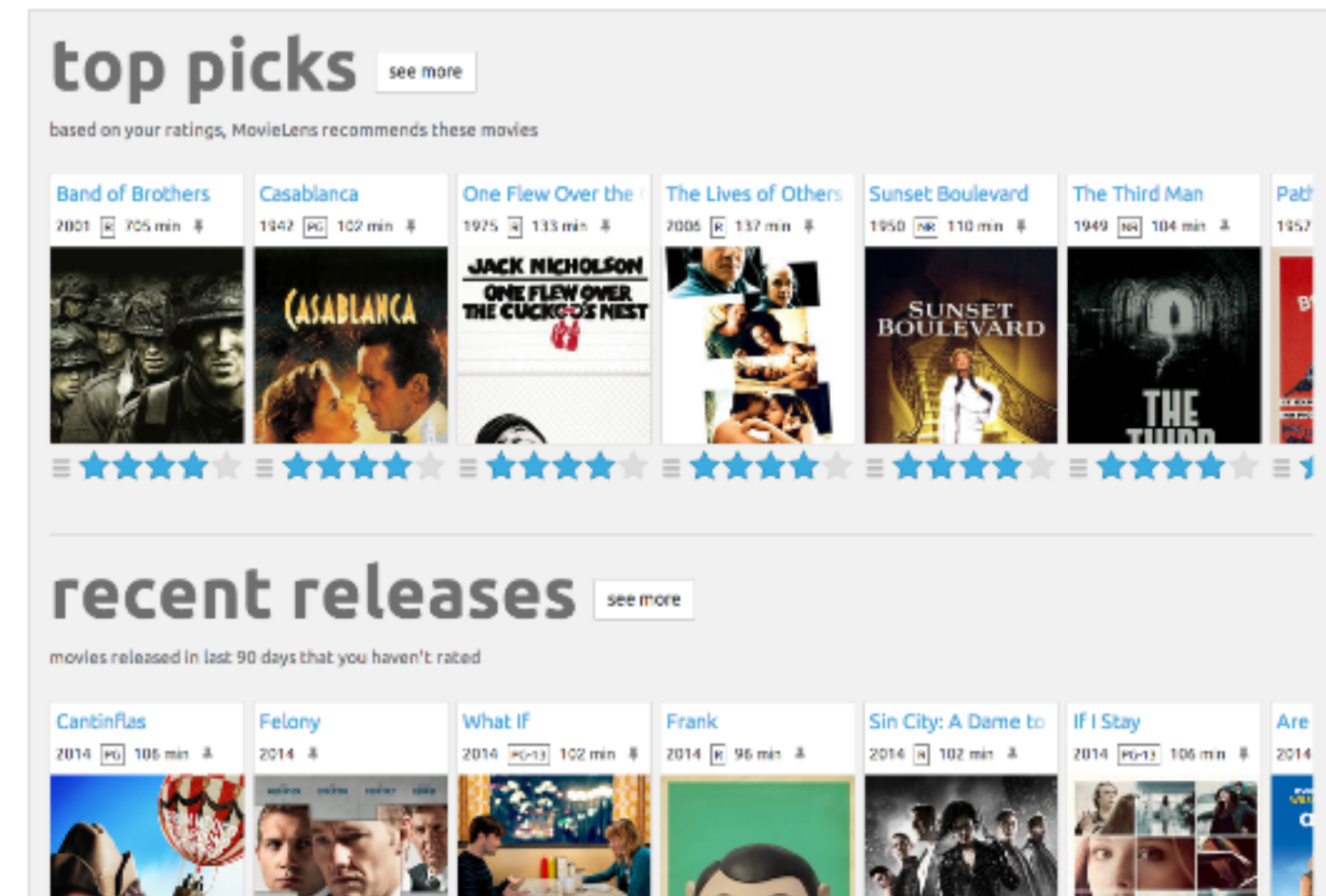
Zero Shot Next Item Recommendation

Se utilizó el dataset **MovieLens 100K**, el cual contiene:

- 944 usuarios
- 1.683 películas

Para la implementación, se usó:

- GPT-3 text-davinci-003 (175B)



Fuente: movielens.org

Zero Shot Next Item Recommendation

Se propone comparar el desempeño en 2 baselines: strong next-item recommendation y zero-shot baselines.

Full Training

- **POP**: Popularity based model
- **FPMC**: Combinación entre factorización matricial y cadenas de Markov
- **GRU4Rec**: GRU based sequential recommendation model
- **SASRec**: Sequential recommendation model with self-attention
- **CL4SRec**: Contrastive learning based sequential recommendation model

Zero-shot

- **Simple Prompting**: Se realiza un prompt
- **CS-Combine-IF**: Selecciona 10 películas random del set item filtering-based
- **CS-Combine-UF**: Selecciona 10 películas random del set user filtering-based
- **NIR-Single-IF**: Combina los 3 pasos en un solo prompt
- **NIR-Single-UF**: Combina los 3 pasos en un solo prompt
- **NIR-Multi-IF**: Usa 3 prompt por separado para guiar al LLM
- **NIR-Multi-UF**: Usa 3 prompt por separado para guiar al LLM

Resultados

Para la evaluación se utilizaron 2 métricas: **HR@10** (Hit Rate at 10) y **ndcg@10** (Normalized Discounted Cumulative Gain at 10).

De la tabla se puede apreciar:

- NIR-based methods superan el desempeño de POP
- Comparado con modelos de recomendación secuencial, NIR-based methods son competitivos
- CS-Random-UF/IF se desempeña mejor que Simple Prompting
- NIR-based methods superan el desempeño de CS-Random-UF/IF y Simple Prompting
- Multi-IF/UF superan el desempeño de Combine-IF/UF
- UF-based NIR-based prompting tienen mejor desempeño que IF-based prompting

Table 1: Main result comparison on MovieLen 100K.

Setting	Method	MovieLens 100K	
		HR@10	NDCG@10
Full Training	POP	0.0519	0.0216
	FPMC	0.1018	0.0463
	GRU4Rec	0.1230	0.0559
	SASRec	0.1241	0.0573
	CL4SRec	0.1273	0.0617
Zero-Shot	Simple Prompting	0.0297	0.0097
	CS-Combine-IF	0.0805	0.0352
	CS-Combine-UF	0.0954	0.0457
	NIR-Single-IF	0.0975	0.0501
	NIR-Single-UF	0.1135	0.0529
	NIR-Multi-IF	0.1028	0.0505
	NIR-Multi-UF	0.1187	0.0546

Conclusiones

Análisis detallado

1. Efectos de los componentes de prompting

La siguiente tabla muestra que todos los componentes del prompting contribuyen al desempeño de las recomendaciones.

Table 2: Ablation study of the impact of different components in the proposed prompting on MovieLens 100K.

Candidate Set	User Preference	Representative Movies	HR@10
-	-	-	0.0297
✓	-	-	0.1071
✓	✓	-	0.1136
✓	-	✓	0.1082
✓	✓	✓	0.1187

Conclusiones

Análisis detallado

2. Impacto del tamaño del conjunto de candidatos.

Los resultados de la siguiente figura indican que el rendimiento de las recomendaciones es sensible al tamaño del conjunto de candidatos.

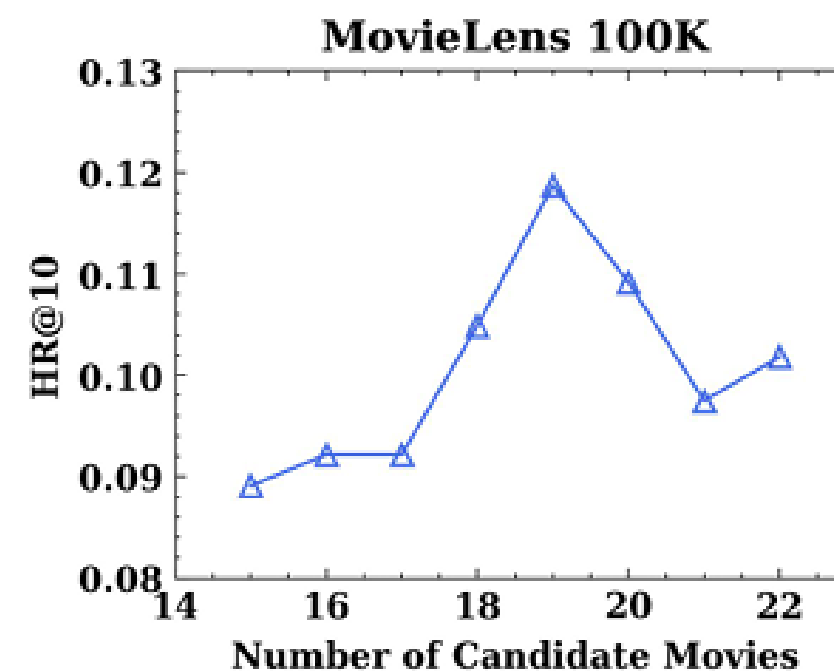


Figure 3: Results for different candidate set sizes.

Crítica

- Los autores no entregan un fundamento matemático que valide lo propuesto.
- Se presenta un enfoque prometededor para recomendaciones de películas, sin embargo sería interesante analizar sus resultados en otros contextos de recomendación y evaluar si es posible generalizar y extender este enfoque.
- Sería interesante estudiar el impacto del sesgo de generar la recomendación en base a un set de candidatos definido utilizando las técnicas mencionadas previamente.



Zero-Shot Next-Item Recommendation using Large Pretrained Language Models

IIC3633 - Sistemas Recomendadores

Jairo Navarro - Begoña Pendas - Fabián Riveros

Noviembre 2023

Referencias

- [1] Jianxin Chang, Chen Gao, Yu Zheng, Yiqun Hui, Yanan Niu, Yang Song, Depeng Jin, and Yong Li. 2021. Sequential recommendation with graph neural networks. In Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 378–387.
- [2] Ruining He and Julian McAuley. 2016. Fusing similarity models with markov chains for sparse sequential recommendation. In 2016 IEEE 16th International Conference on Data Mining (ICDM). IEEE, 191–200.
- [3] Zhiwei Liu, Yongjun Chen, Jia Li, Man Luo, S Yu Philip, and Caiming Xiong. 2021. Self-supervised Learning for Sequential Recommendation with Model Augmentation. (2021).
- [4] Yuhui Zhang, Hao Ding, Zeren Shui, Yifei Ma, James Zou, Anoop Deoras, and Hao Wang. 2021. Language Models as Recommender Systems: Evaluations and Limitations. In I (Still) Can't Believe It's Not Better! NeurIPS 2021 Workshop.