



Pontificia Universidad Católica de Chile
Departamento de Ciencias de la Computación
Sistemas Recomendadores

Transparent and Scrutable Recommendations Using Natural Language User Profiles

Pablo Leiva
Manuel Muñoz
Pascual Sepúlveda



Definición y Motivación del Estudio

El rol de los sistemas de recomendación

- Personalizan contenido en plataformas como Netflix y Amazon.
- Basados en representaciones latentes de interacciones previas.
- Tienen como objetivo sugerir contenido relevante.



Definición y Motivación del Estudio

Limitaciones de los sistemas actuales

- Representaciones opacas (embeddings latentes).
- Falta de transparencia en las recomendaciones.
- Difícil reflejar cambios dinámicos en las preferencias.



Definición y Motivación del Estudio

Propuesta del paper

- Uso de Modelos de Lenguaje de Gran Escala (LLMs).
- Creación de perfiles de usuario en lenguaje natural.
- Beneficios:
 - Perfiles legibles y comprensibles.
 - Ajustes directos por parte de los usuarios.



Definición y Motivación del Estudio

Conceptos clave: Transparencia y Escrutinio

- **Transparencia:** Representación fiel de los mecanismos de recomendación.
- **Escrutinio:** Permite inspeccionar y modificar preferencias.



Trabajos Relacionados

Enfoques previos en sistemas de recomendación

- 1 Plantillas predefinidas
- 2 Preferencias basadas en conjuntos
- 3 Recomendaciones basadas en conjuntos



Trabajos Relacionados

Uso de plantillas predefinidas

Generate Neural Template Explanations

Uso de redes neuronales para aprender plantillas.

Personalized Transformer for Explainable Recommendation

Mejoras mediante transformadores.

Personalized Prompt Learning for Explainable Recommendation

Introducción de prompt learning en modelos preentrenados.

Limitaciones:

- Solo explican recomendaciones a nivel de ítem.
- No permiten entender o modificar el perfil completo del usuario.



Trabajos Relacionados

Preferencias basadas en conjuntos

Learning from Sets of Items in Recommender Systems

Mejoran recomendaciones a nivel de ítems.

Transparent, Scrutable and Explainable User Models for Personalized Recommendation
Utilizan etiquetas para inferir preferencias.

Using Groups of Items for Preference Elicitation in Recommender Systems

Clustering para agrupar ítems automáticamente.

Limitaciones:

- Dependencia de etiquetas predefinidas.
- No reflejan preferencias más complejas o matices únicos.



Trabajos Relacionados

Preferencias basadas en contexto

**Large Language Model
Augmented Narrative Driven
Recommendations**
Convierte reseñas en narrativas personalizadas.

Limitaciones:

- Narrativas no editables, limitan el control del usuario.
- Dependencia de un historial significativo de reseñas.



Método Propuesto

User Profile Recommendation (UPR)

- Generación de perfiles personalizados
 - Comprensibles
 - Personalizables
- Uso de lenguaje natural
- Mejora en transparencia y ajustes



Método Propuesto

Procedimiento: Cálculo de puntaje promedio

- Análisis de elementos con los que el usuario interactuó
- Normalización en intervalo $[-1, 1]$
- Identificación de palabras clave en reseñas
- Cálculo de puntajes promedio.



Método Propuesto

Procedimiento: Cálculo de puntaje promedio

$$\bar{r}_f = \sum_{i \in \mathcal{I}_f} r_i$$

Donde:

r : recomendación

f : palabra característica



Método Propuesto

Procedimiento: Función de utilidad

$$U(f) = |\bar{r}(f)| \times cov(f) \times sig(f)$$

$r(f)$: Preferencia del usuario calculada anteriormente por el elemento f .

$conv(f)$: Cuánto aparece en las reseñas el elemento f vs el total de reseñas.

$Sig(f)$: Peso que se le da a las características que más menciona el usuario.



Método Propuesto

Procedimiento: Convertir features con mayor utilidad a un perfil

- Uso de modelos preentrenados
 - Llama y Mistral, ajustados para seguir instrucciones.
- Generación del perfil
 - Resumen de reseñas por características.
- Límite en prompts
 - Máximo 5 reseñas por característica.
 - Perfil limitado a 200 tokens.



Método Propuesto

Entrenamiento: Datasets

- Amazon Movies and TV (Amazon-TV)
- TripAdvisor

	Amazon-MT	TripAdvisor
#users	4,811	9,765
#items	5,459	6,280
#records	288,693	320,023
#features	5,114	5,069
#records per user	60.01	32.77
#records per item	52.88	50.96
#words per explanation	14.14	13.01



Método Propuesto

Entrenamiento: Datasets

- Amazon Movies and TV (Amazon-TV)
- TripAdvisor

Separación		
Entrenamiento	Validación	Pruebas
80%	10%	10%



Método Propuesto

Experimentación: Entrenamiento y modelos utilizados

- Se utilizó el modelo GPT-2
- Uso del perfil en lenguaje natural y los ítems de la base de datos.
- Ajuste de predicciones de calificaciones basándose en preferencias.
- Normalización de calificaciones:
 - Escala de entrenamiento: [0, 1].
 - Escala de salida: [1, 5].



Método Propuesto

Experimentación: Modelos utilizados para comparación

- Most Popular
 - Recomienda los ítems con más calificaciones.
 - No considera preferencias individuales.
- UserKNN-BM25
 - Filtrado colaborativo.
 - Vecinos más cercanos (KNN) entre usuarios.
 - Ponderado con BM25.



Método Propuesto

Experimentación: Modelos utilizados para comparación

- Item-KNN-cosine
 - Filtrado colaborativo basado en ítems.
 - Similaridad coseno entre ítems.
- Matrix Factorization (MF)
 - Aprende incrustaciones latentes de usuarios e ítems.
- NeuMF
 - Modelo híbrido: factorizar matrices y perceptrones multicapa.



Método Propuesto

Experimentación: Modelos explicables

- PETER+
 - Genera explicaciones junto con las predicciones.
- PEPLER-MLP
 - Combina predicción de calificaciones con generación de explicaciones a nivel de ítem.



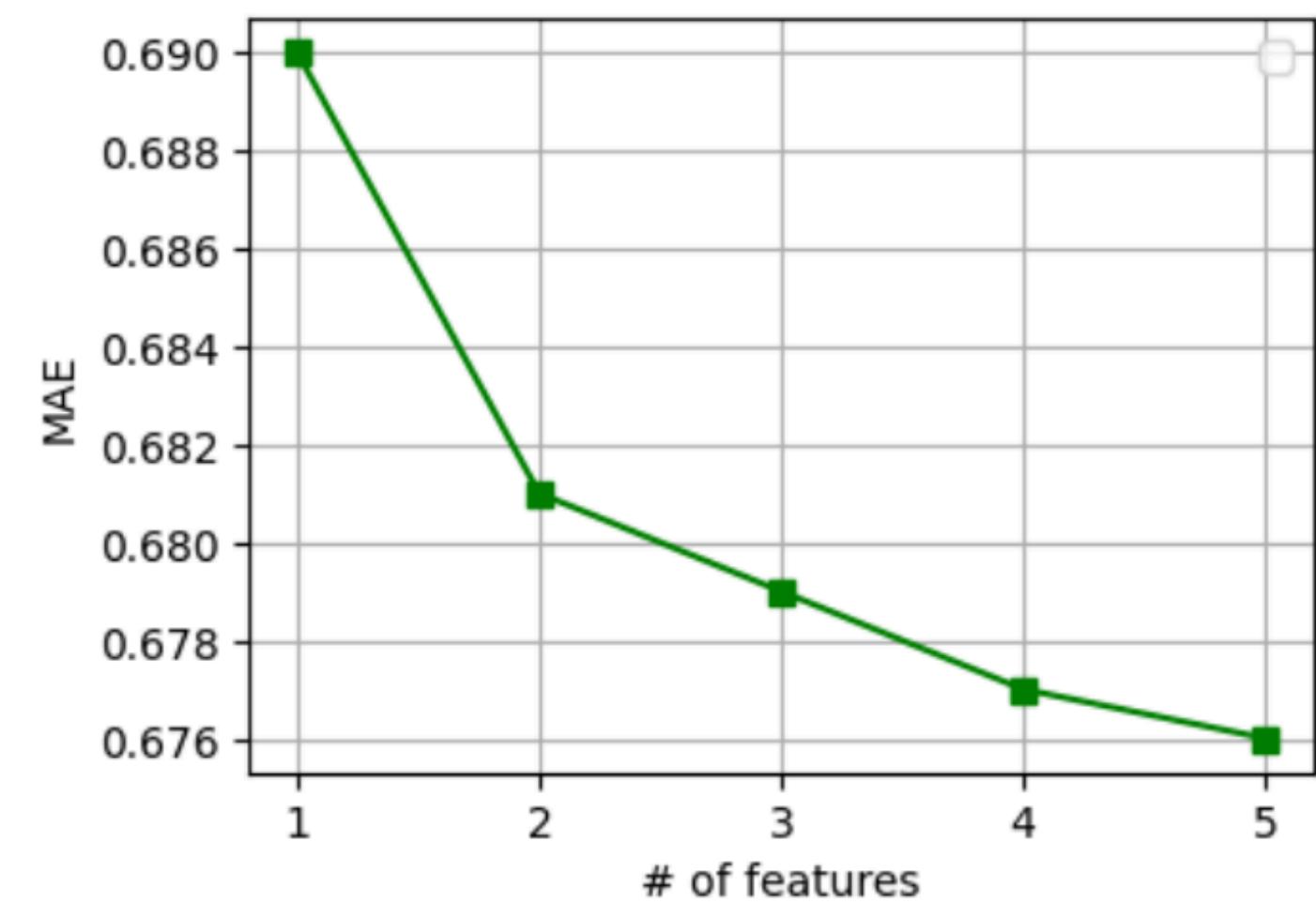
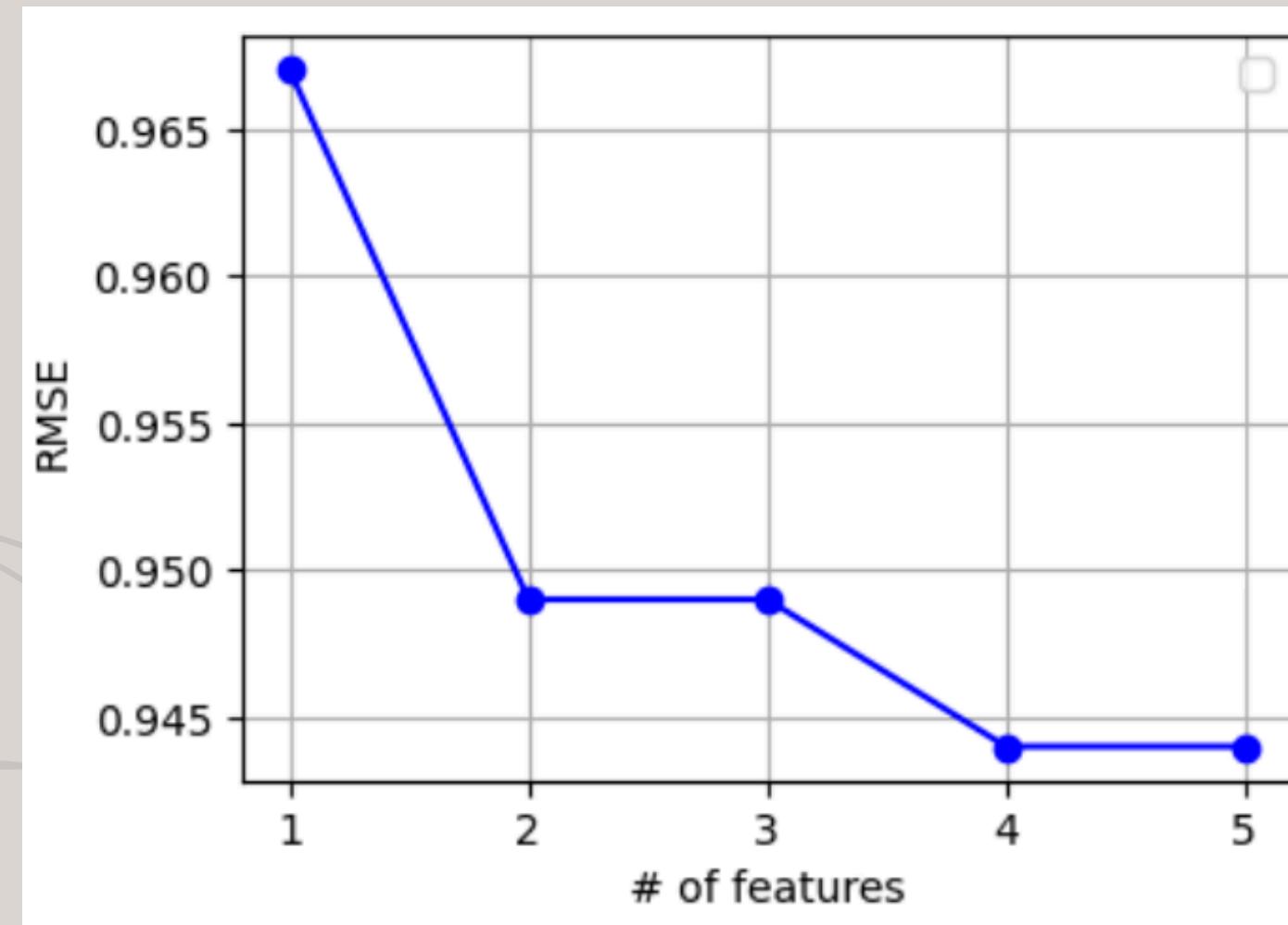
Resultados

RMSE, MAE, nDCG@10, MAP

Model	Amazon-MT				TripAdvisor			
	RMSE	MAE	nDCG@10	MAP	RMSE	MAE	nDCG@10	MAP
MostPop	1.505	0.962	0.896	0.815	1.277	0.902	0.932	0.857
UserKNN	0.960	0.707	0.935	0.864	0.836	0.634	0.952	0.883
Item-KNN	1.045	0.790	0.897	0.823	0.890	0.683	0.953	0.852
MF	0.925	0.686	0.941	0.870	0.786	0.599	0.960	0.891
NeuMF	0.943	0.694	0.936	0.866	0.819	0.634	0.955	0.886
PETER+	0.924	0.685	0.941	0.870	0.803	0.621	0.958	0.889
PEPLER-MLP	0.925	0.672	0.941	0.869	0.793	0.606	0.959	0.889
UPR (Llama2)	0.944	0.678	0.938	0.866	0.804	0.616	0.955	0.885
UPR (Mistral)	0.941	0.679	0.940	0.870	0.804	0.610	0.952	0.887

Resultados

Estudio de ablación





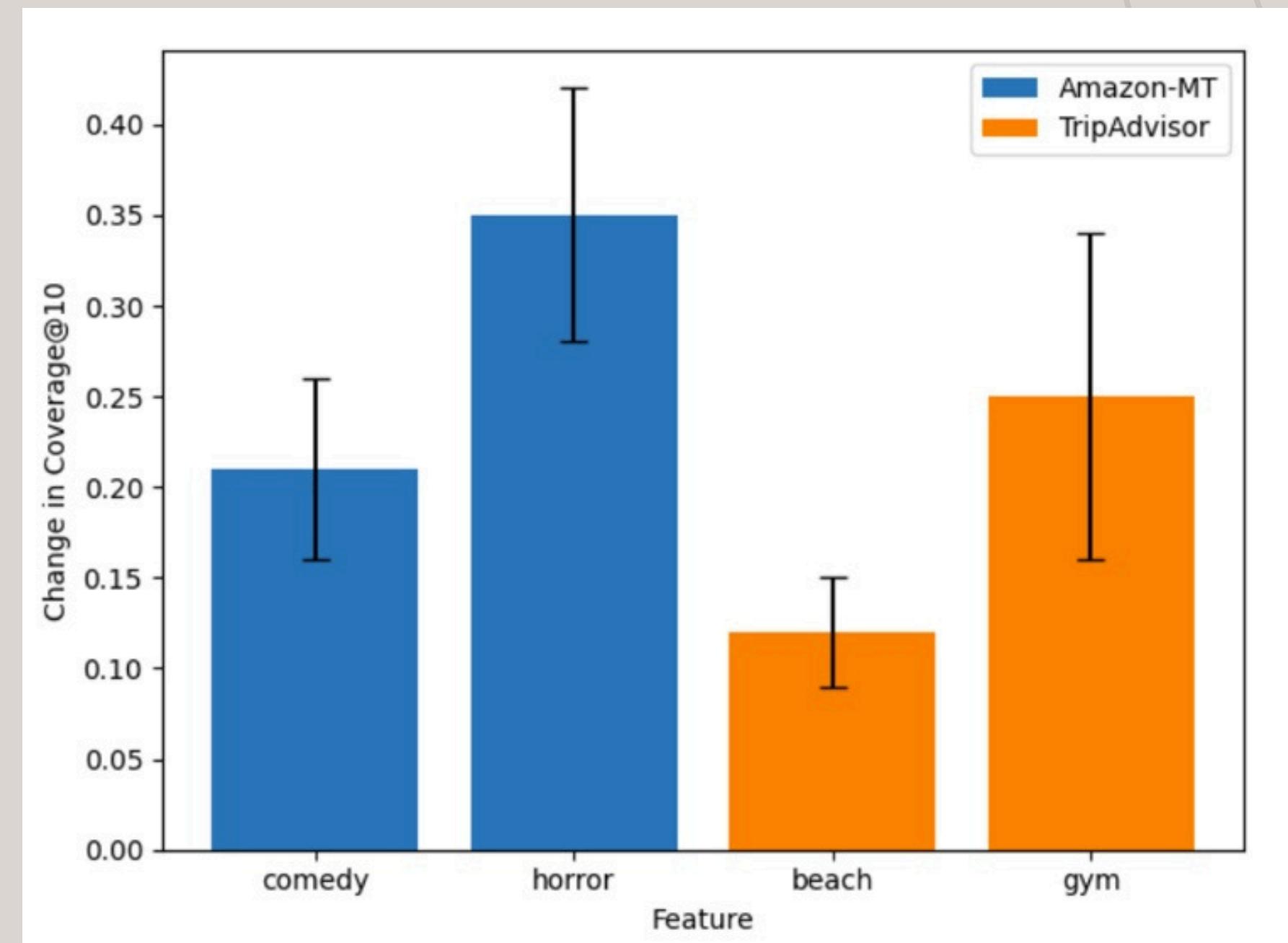
Resultados

Estudio cualitativo de perfiles

Metric	Amazon-MT	TripAdvisor
Fluency	95%	92%
Informativeness	86%	82%
Conciseness	75%	72%
Relevance	90%	87%

Resultados

Capacidad de ajustar preferencias en tiempo real





Conclusiones y Limitaciones

- Transparente y escrutable
- Métricas competitivas
- Recomendaciones actualizadas al instante
- Depende de la calidad de las reseñas
- Predominancia de reseñas positivas
- Significativamente más lento



Trabajos a futuro

- Reducir posibles sesgos en las recomendaciones
- Equilibrio óptimo entre longitud de perfil y carga cognitiva



Pontificia Universidad Católica de Chile
Departamento de Ciencias de la Computación
Sistemas Recomendadores

Transparent and Scrutable Recommendations Using Natural Language User Profiles

Pablo Leiva
Manuel Muñoz
Pascual Sepúlveda