www.elementos.buap.mx

Sistemas de recomendación: el motor invisible de la era digital

Víctor Giovanni **Morales Murillo**
Karina **Cancino Villatoro**

Los sistemas de recomendación son una de las aplicaciones más importantes de la inteligencia artificial (IA), el análisis de datos y la ciencia de datos, debido a que se han integrado en nuestra vida diaria a través de múltiples servicios en línea, incluyendo la mayoría de los principales sitios de comercio electrónico (e-commerce), plataformas multimedia (servicios de streaming) y redes sociales. De esta manera, compañías de tecnología como Google, Amazon, Spotify, Facebook, Netflix, Uber, Airbnb, LinkedIn y muchas más, han utilizado los motores de recomendación como uno de sus principales servicios para mejorar la experiencia de los usuarios cuando navegan en sus sitios web, con la finalidad de incrementar los indicadores clave de rendimiento (KPIs), como el número de ventas y la retención de clientes (Castells y Jannach, 2023).

Esta tecnología es una herramienta potente que facilita la toma de decisiones informadas, efectivas y eficientes dentro de una amplia variedad de artículos y servicios conocidos como ítems, los cuales pueden ser productos comerciales, listas de reproducción, series o películas, amigos en redes sociales, vacantes o aplicantes de empleo, puntos de interés (atractivos turísticos, hoteles, restaurantes, aparcamientos, etc.), publicaciones (libros, revistas, artículos, etc.), profesiones o carreras profesionales, cursos académicos y muchos más ítems (Morales-Murillo *et al*., 2024).

De esta manera, las recomendaciones personalizadas y automatizadas son omnipresentes en la era digital, en cualquier parte del ciberespacio nos encontramos con sugerencias sobre artículos que podemos comprar, aplicaciones novedosas para descargar y nuevos artistas o música por descubrir. Estas sugerencias personalizadas son generadas por los sistemas de recomendación, que son componentes de software que identifican los ítems más adecuados para un determinado usuario, usando estadísticas y modelos de aprendizaje automático, llamados en inglés *machine learning* (Castells y Jannach, 2023).

El campo de investigación sobre los sistemas de recomendación es muy amplio y continúa creciendo rápidamente, debido a los casos de éxito que han beneficiado a múltiples negocios y consumidores. Asimismo, el continuo desarrollo del aprendizaje automático y de la era digital es un fuerte impulso para esta área de investigación. Actualmente, la población mundial es de aproximadamente 8 mil millones de personas, de las cuales, el 66.2 % son usuarios de internet y el 62.3 % son usuarios de redes sociales (Kemp, 2024). Mientras que el tiempo promedio que un ser humano pasa en internet es de 6 horas y 40 minutos, en el uso de redes sociales es de 2 horas y 23 minutos, y en servicios de *streaming* es de 3 horas y 6 minutos. Es difícil calcular la enorme cantidad de información que se produce en la era digital, lo que ha intensificado exponencialmente el problema de la sobrecarga de información.

En respuesta, surgieron los sistemas de recomendación con una tendencia en investigaciones y recursos en idioma inglés. Sin embargo, el desarrollo de nuevos y diferentes recursos para el área de sistemas de recomendación es fundamental para generar nuevas oportunidades en diversos dominios de aplicación y en diferentes regiones del mundo, especialmente en los países que están en vías de desarrollo y que requieren del alcance de este tipo de tecnología, debido a la globalización.

La ética en inteligencia artificial es un campo multidisciplinario que tiene el objetivo de regular el desarrollo y uso de la IA; no obstante, las investigaciones en esta área son limitadas. Por lo tanto, es un área de oportunidad para atender los desafíos éticos de los sistemas de recomendación, como son las recomendaciones inapropiadas, la recopilación y almacenamiento de datos no autorizados, las fugas de datos, las inferencias no autorizadas, las trampas de comportamiento del usuario, los algoritmos de caja negra, las explicaciones de recomendaciones poco informativas y el sesgo de datos.

ALGORITMOS DE RECOMENDACIÓN

Las técnicas clásicas y más populares de recomendación son el filtrado basado en contenido, el filtrado colaborativo y los enfoques híbridos de recomendación (Morales *et al*., 2022; Castells y Jannach, 2023). Mientras que las últimas tendencias son los sistemas de recomendación secuenciales/basados en sesiones (Moreira *et al*., 2021), las recomendaciones explicables (Zhang y Chen, 2020) y los sistemas de recomendación conversacionales (CRS) (Jannach *et al*., 2020), que han sido impulsados por el rápido desarrollo del campo del procesamiento de lenguaje natural (*NLP: Natural Language Processing*), el aprendizaje profundo (*Deep Learning*), el aprendizaje por refuerzo (*Reinforcement Learning*) y los grandes modelos de lenguaje (*LLMs: Large Language Models*).

FILTRADO BASADO EN CONTENIDO

Esta técnica de recomendación utiliza como principal fuente de información las descripciones textuales de los ítems que representan sus características; así, las preferencias pasadas de un usuario pueden hacer match con esas características. Este algoritmo se basa en el campo de recuperación de información (IR: Information Retrieval), en donde el objetivo es identificar y ponderar un conjunto de documentos relevantes dada una consulta, mientras que en los motores de recomendación los documentos representan a los ítems y se ponderan con base en un perfil del usuario. En la Figura 1 se presenta un ejemplo de filtrado basado en contenido, en el que

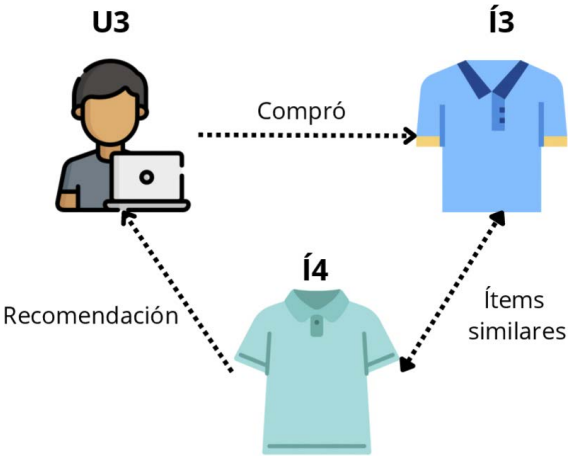

Figura 1. Ejemplo de filtrado basado en contenido. Imagen elaborada con recursos gráficos de flaticon.com.

	Í1	Í2	Í3	Í4	Í5
U1	3	4	3	?	?
U2		4	3		5
U3		1	3	1	
U4	4		3	2	3
U5	3		3	2	

Tabla 1. Ejemplo de una matriz de calificaciones con 5 usuarios y 5 ítems (*User-Item Rating Matrix*) para el filtrado colaborativo.

al usuario U3 se le recomienda una camisa similar (ítem Í4) a la que ha comprado en el pasado (ítem Í3). Una alternativa es usar un modelo espacio vectorial de IR para representar los artículos como vectores de documentos que contienen las descripciones de los ítems y emplear la similitud del coseno para estimar los productos más similares al perfil del usuario.

FILTRADO COLABORATIVO

Este enfoque de recomendación se ha utilizado ampliamente en el estado del arte con la idea de aprovechar las opiniones o comentarios de otros usuarios. La principal fuente de información es un conjunto de calificaciones de los usuarios hacia los ítems que han comprado o consumido en el pasado. Estas calificaciones se almacenan en una matriz de calificaciones, conocida en inglés como *User-Item Rating Matrix*, la cual generalmente usa una escala de Likert de 1 a 5 en las calificaciones para representar el grado de interés que tiene un usuario hacia un ítem. De esta manera, las calificaciones representan los siguientes niveles de interés que puede tener un usuario hacia un ítem.

1. Muy negativo
2. Negativo
3. Neutro
4. Positivo
5. Muy positivo

En la Tabla 1 se muestra un ejemplo de una matriz de calificaciones que incluye 5 usuarios y 5 ítems. En este caso, el usuario 1 ha otorgado una calificación de 3 al ítem I1, una de 4 al ítem I2 y otra de 3 al ítem I3. Sin embargo, las calificaciones correspondientes a los ítems I4 e I5 aún no están disponibles y se busca predecirlas usando las calificaciones de los usuarios más similares al usuario U1. Supongamos que los usuarios U2 y U4 son los más similares al usuario U1; entonces, la predicción para el ítem I5 del usuario U1 sería el promedio de la calificación del ítem I5 del usuario U2 y U4, que corresponde a una calificación de 4.

FILTRADO HÍBRIDO

El filtrado híbrido consiste en combinar dos o más técnicas de recomendación para complementar los algoritmos y resolver sus limitaciones al implementarse de manera individual.

El filtrado basado en contenido presenta un problema de novedad, en donde las sugerencias de los motores de recomendación se basan exclusivamente en el perfil del usuario; es decir, si al usuario le gustan las películas de terror, entonces se le recomendarán películas del mismo género sin tomar en cuenta otros géneros. Mientras que en el filtrado colaborativo se presenta el problema de arranque en frío, en donde generar predicciones es complejo

Figura 2. Ejemplo de información secuencial para el usuario U1 y ejemplo de información de sesiones para el usuario U2. Imagen elaborada con recursos gráficos de flaticon.com y Amazon.com.

cuando los usuarios tienen un bajo número de ítems calificados o no han calificado ningún ítem.

RECOMENDACIONES SECUENCIALES Y BASADAS EN SESIONES

Recientemente, los sistemas de recomendación secuenciales y basados en sesiones (Wang *et al*., 2022) han surgido como un nuevo paradigma que captura las preferencias dinámicas y a corto plazo del usuario para generar recomendaciones más precisas y oportunas. Estos algoritmos se basan en las interacciones del usuario hacia los ítems a lo largo del tiempo.

Una interacción está constituida por un usuario, un ítem y una acción, por ejemplo, un clic, una vista o una compra. Una secuencia o sesión es un conjunto de interacciones con propiedades específicas, como la longitud (número de interacciones) de la secuencia o sesión, el orden interno, el tipo de acción (vista, compra o clic), la información disponible del usuario y la estructura de los datos.

A diferencia de los sistemas de recomendación secuenciales, que solo tienen una secuencia de interacciones por usuario, en las recomendaciones basadas en sesiones los usuarios pueden tener más de una sesión, las cuales están separadas por intervalos de tiempo.

En la Figura 2 se presenta un ejemplo de información secuencial para el usuario U1, que ha leído una saga de libros de *Los Juegos del Hambre* y la información de sesiones del usuario U2, que ha comprado o visto diferentes ítems en tres sesiones distintas separadas por intervalos de tiempo; es decir, la sesión 1 y 2 pueden tener una diferencia de 2 o 3 semanas, y la sesión 2 y 3 pueden tener una diferencia de tiempo de N semanas.

La idea de estos algoritmos es usar un conjunto de datos sobre interacciones de los usuarios hacia los ítems para predecir las siguientes acciones de los usuarios como compras, vistas o clics.

Además, estas técnicas son una tendencia en los sistemas de recomendación debido a que los últimos avances en *NLP*, Deep Learning y *LLMs* se están empleando para generar nuevos modelos de recomendación (Vats *et al*., 2024).

Explicación basada en el usuario	Explicación basada en el ítem
El libro *Los Detectives Salvajes* de Roberto Bolaño fue recomendado a ti porque tus usuarios similares Pedro y Pablo han comprado ese mismo ítem antes.	El libro *Los Detectives Salvajes* de Roberto Bolaño fue recomendado a ti porque tú leíste un libro del mismo autor antes.

Tabla 2. Ejemplos de explicación de recomendaciones basadas en el usuario y basadas en el ítem.

RECOMENDACIÓN EXPLICABLE

Otra tendencia de investigación en los sistemas de recomendación es generar nuevos modelos de explicación (Zhang y Chen, 2020), también conocidos como modelos interpretables o transparentes, los cuales describen por qué específicos ítems fueron recomendados a ciertos usuarios.

Esto es para mejorar la transparencia, persuasión, eficacia, confiabilidad y satisfacción del usuario en los sistemas de recomendación. Asimismo, garantiza una mejor depuración de errores para el algoritmo de recomendación.

Un ejemplo de una explicación de recomendación tradicional se muestra en la Tabla 2, en donde en la parte izquierda se presenta una explicación basada en el usuario y en la parte derecha se observa una explicación basada en el ítem.

Usuario	Quiero encontrar un drama legal
Robot: Identificación de preferencias del usuario	¿Algún actor o director favorito?
Usuario	Quiero que una actriz tenga el papel principal
Robot: Recomendación	¿Qué tal *Suits*? Es un drama legal sobre un brillante desertor universitario
Usuario	¿Por qué me lo recomiendas?
Robot: Explicación	Es una película legal cuyo papel principal es una mujer. Y la historia es fantástica e interesante
Usuario	De acuerdo, ¿en qué periodo de tiempo transcurre?
Robot: Búsqueda de información del ítem	De 2011 a 2012
Usuario	¡Gracias! Me gusta esta película
Robot	Suena bien

Tabla 3. Ejemplo de un sistema de recomendación conversacional (modificada de Feng *et al*., 2023).

SISTEMAS DE RECOMENDACIÓN CONVERSACIONALES

Existen escenarios en los que las preferencias del usuario no pueden ser estimadas usando sus interacciones pasadas; por ejemplo, cuando se recomienda un celular o una casa. Hay recomendaciones que dependen altamente de su contexto y que pueden ser complejas de determinar debido a las necesidades y situaciones actuales del usuario. Otra suposición es que los usuarios ya conocen sus preferencias o las construyen en el proceso de decisión cuando observan las diferentes opciones de los ítems.

El objetivo de un sistema de recomendación conversacional es recomendar eficientemente ítems a los usuarios a través de una interfaz de diálogo. Esta tecnología contiene múltiples subtareas como la identificación de las preferencias del usuario, la recomendación, la explicación y la búsqueda de información del ítem (Feng *et al*., 2023). Además, los desafíos que enfrentan son:

1. ¿Cómo gestionar y resolver las subtareas?
2. ¿Cómo generar respuestas correctamente que interactúen con los usuarios?

En la Tabla 3 se presenta un ejemplo de un sistema de recomendación conversacional, en donde un CRS tiene la habilidad de gestionar y resolver las subtareas. Asimismo, una tecnología de este tipo necesita generar respuestas que interactúen con el usuario.

© **Raúl Chiquito**. *Tortilla y chile*. Campesino comiendo tortillas y chile.

CONCLUSIONES

La transformación digital es un nuevo paradigma económico y social en donde el principal activo es el conocimiento basado en la información para mejorar la toma de decisiones de las personas, las empresas, el gobierno u otras organizaciones. De esta manera, por la naturaleza de los sistemas de recomendación, se han convertido en una de las principales tecnologías de la era digital. Existen diversos enfoques de recomendación para aprovechar las diversas fuentes de información en diferentes dominios de aplicación. Otro fuerte impulso para los sistemas de recomendación además del comercio electrónico, es la Industria 4.0, que busca integrar la inteligencia artificial, la impresión 3D, el internet de las cosas (IoT) y otros avances tecnológicos en los procesos industriales y de fabricación. Por tal motivo, es importante generar nuevos recursos de recomendación en distintos idiomas y dominios de aplicación.

R E F E R E N C I A S

Castells P and Jannach D (2023). *Recommender Systems: A Primer*. Recuperado de: http://arxiv.org/abs/2302.02579.

Feng Y, Liu S, Xue Z, Cai Q *et al*. (2023). *A Large Language Model Enhanced Conversational Recommender System*. Recuperado de: http://arxiv.org/abs/2308.06212.

Jannach D, Manzoor A, Cai W and Chen L (2020). *A Survey on Conversational Recommender Systems*. Recuperado de: https://doi.org/10.1145/3453154.

Kemp S (2024). *Digital 2024*. Recuperado de: https://wearesocial.com/es/blog/2024/01/digital-2024/.

Morales VG, Pinto D, Rojas F y Gonzales J (2022). A Systematic Literature Review on the Hybrid Approaches for Recommender Systems. *Computación y Sistemas*, 26(1):357-372.

Morales-Murillo VG, Pinto D, Perez-Tellez F and Rojas-Lopez F (2024). A Transformer-Based Multi-Domain Recommender System for E-commerce. *International Journal of Combinatorial Optimization Problems and Informatics* 15(2):95-123.

Moreira GDSP, Rabhi S, Lee JM, Ak R and Oldridge E (2021). Transformers4Rec: Bridging the Gap between NLP and sequential/session-based recommendation. *RecSys 2021-15th ACM Conference on Recommender Systems*.

Vats A, Jain V, Raja R y Chadha A (2024). *Exploring the Impact of Large Language Models on Recommender Systems: An Extensive Review*. Recuperado de: http://arxiv.org/abs/2402.18590.

Wang S, Zhang Q, Hu L, Zhang X, Wang Y and Aggarwal C (2022). Sequential/Session-based Recommendations: Challenges, Approaches, Applications and Opportunities. *SIGIR 2022-Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*.

Zhang Y and Chen X (2020). Explainable recommendation: A survey and new perspectives. *Foundations and Trends in Information Retrieval*. 14(1):1-101.

Víctor Giovanni Morales Murillo
Karina Cancino Villatoro
Universidad Politécnica de Tapachula
vg055@hotmail.com